Workgroup: TEAS
Internet-Draft: draft-srld-teas-5g-slicing-09
Published: 23 May 2023
Intended Status: Informational
Expires: 24 November 2023
Authors: K. G. Szarkowicz, Ed.
Juniper Networks
R. Roberts, Ed.
Juniper Networks
J. Lucek
Juniper Networks
M. Boucadair, Ed.
Orange
L. M. Contreras
Telefonica

A Realization of IETF Network Slices for 5G Networks Using Current IP/MPLS Technologies

Abstract

5G slicing is a feature that was introduced by the 3rd Generation Partnership Project (3GPP) in mobile networks. This feature covers slicing requirements for all mobile domains, including the Radio Access Network (RAN), Core Network (CN), and Transport Network (TN).

This document describes a basic IETF Network Slice realization model in IP/MPLS networks with a focus on the Transport Network fulfilling 5G slicing connectivity requirements. This realization model reuses many building blocks currently commonly used in service provider networks.

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 24 November 2023.

Table of Contents

1. Introduction

[I-D.ietf-teas-ietf-network-slices] defines a framework for network slicing in the context of networks built using IETF technologies. The IETF network slicing framework introduces the concept of a Network Resource Partition (NRP), which is simply a collection of resources identified in the underlay network. There could be multiple realizations of IETF Network Slice and NRP concepts, where each realization might be optimized for the different network slicing use cases.

This document describes an IETF Network Slice realization model in IP/MPLS networks, using a single NRP and with a focus on fulfilling 5G slicing connectivity requirements. This IETF Network Slice realization model leverages many building blocks currently commonly used in service provider networks.

A brief 5G overview is provided in Appendix C for readers' convenience. The reader may refer to [TS-23.501] or [_5G-Book] for more details about 3GPP network architectures.

2. Definitions

The document uses the terms defined in [I-D.ietf-teas-ietf-network-slices]. See Section 3.3 for the contextualization of some of these terms.

An extended list of abbreviations used in this document is provided in Appendix B.

3. 5G Network Slicing Integration in Transport Networks

3.1. Scope of the Transport Network

Appendix C provides an overview of 5G network building blocks: the Radio Access Network (RAN), Core Network (CN), and Transport Network (TN). The Transport Network is defined by the 3GPP as the "part supporting connectivity within and between CN and RAN parts" (Section 1 of [TS-28.530]).

As discussed in Section 4.4.1 of [TS-28.530], the 3GPP managment system does not directly control the Transport Network: it is considered as a non-3GPP managed system.

> 'The non-3GPP part includes TN parts. The 3GPP management system provides the network slice requirements to the corresponding management systems of those non-3GPP parts, e.g. the TN part supports connectivity within and between CN and AN parts.' (Section 4.4.1 of [TS-28.530])

In practice, the TN may not map with a monolithic architecture and management domain. It is frequently segmented, non-uniform, and managed by different entities. For example, Figure 1 depicts a Network Function (NF) instance that is deployed in an edge data center (DC) connected to a NF located in a Public Cloud via a Wide Area Network (WAN) (e.g., MPLS-VPN service). In this example, the TN can be seen as an abstraction representing an end-to-end connectivity based upon three distinct domains: DC, WAN, and Public Cloud. A model for the Transport Network based on orchestration domains is introduced in Section 3.4. This model permits to define more precisely where the IETF Network Slices apply.

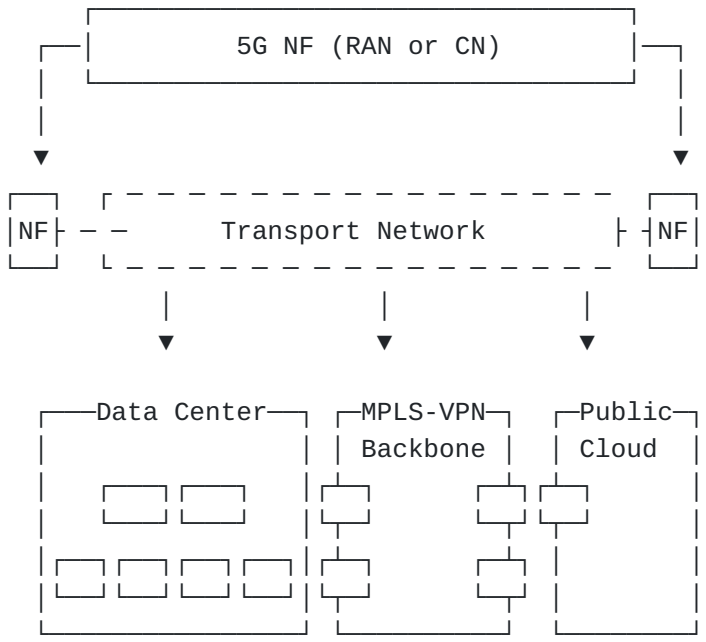


Figure 1: An Example of Transport Network Decomposition

The term "Transport Network" is used for disambiguation with 5G network (e.g., IP, packet-based forwarding vs RAN and CN). Consequently, the disambiguation applies to Transport Network Slicing vs. End-to-End 5G Network Slicing (Section 3.2) as well the management domains: RAN, CN, and TN domains.

3.2. 5G Network Slicing versus Transport Network Slicing

Network slicing has a different meaning in the 3GPP mobile and transport worlds. Hence, for the sake of precision and without seeking to be exhaustive, this section provides a brief description of the objectives of 5G Network Slicing and Transport Network Slicing:

*5G Network Slicing:

Is defined by the 3GPP as an appraoch where logical networks/ partitions are created, with appropriate isolation, resources and optimized topology to serve a purpose or service category or customers [TS-28.530]. These resources are from the TN, RAN, CN Network Functions, and the underlying infrastructure.

*TN Slicing:

The term "TN Slice" is used in this document to refer to a slice in the Transport Network domain of the overall 5G architecture.

The objective of TN Slicing is to isolate, guarantee, or prioritize Transport Network resources for slices. Examples of

such resources are: buffers, link capacity, or even Routing Information Base (RIB) and Forwarding Information Base (FIB).

TN Slicing provides various degrees of sharing of resources between slices. For example, the network capacity can be shared by all slices, usually with a guaranteed minimum per slice, or each individual slice can be allocated dedicated network capacity. Parts of a given network may use the former, while others use the latter. For example, in order to satisfy local engineering guidelines and specific service requirements, shared TN resources could be provided in the backhaul (or midhaul), and dedicated TN resources could be provided in the midhaul (or backhaul). The capacity partitioning strategy is deployment specific.

There are different options to implement TN slices based upon tools, such as Virtual Routing and Forwarding instances (VRFs) for logical separation, Quality of Service (QoS), or Traffic Engineering (TE).

3.3. Transport Network Reference Design

Figure 2 depicts the reference design used for modelling the Transport Network based on management perimeters (Customer vs. Provider).

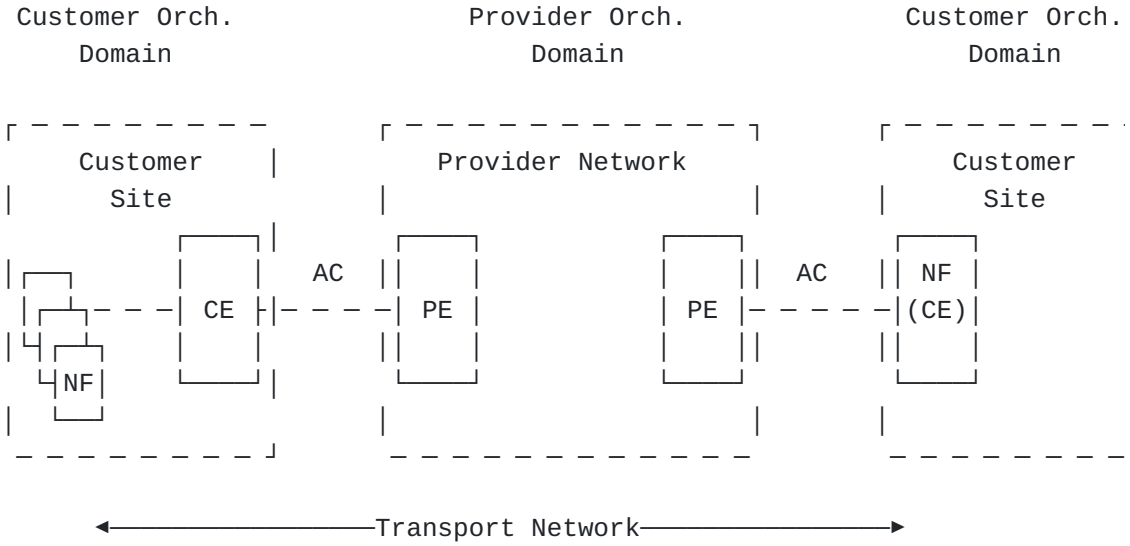


Figure 2: Reference Design: Customer Sites and Provider Network

The description of the main components shown in Figure 2 are:

Customer:

An entity that is responsible for managing and orchestrating the End-to-End 5G Mobile Network, notably RANs and CNs.

Customer Sites: A customer manages and deploys 5G Network Functions (RAN and CN) in Customer Sites. On top of 5G Network Functions (e.g., gNodeB (gNB), 5G Core (5GC)), a customer may manage additional TN elements (e.g., servers, routers, switches, or VPC Gateways) within a Customer Site. A Customer Site can be either a physical or a virtual location. Examples of Customer Sites are a customer private locations (Point of Presence (PoP), DC), a VPC in a Public Cloud, or servers hosted within provider or colocation service. The Orchestration of the TN within Customer Sites relies upon a set of controllers for automation purposes (e.g., Network Functions Virtualization Infrastructure (NFVI), Enhanced Container Network Interface (CNI), Fabric Managers, or Public Cloud APIs). The detail of these is out of the scope of this document.

Provider: An entity responsible for interconnecting Customer Sites. The provider orchestrates and manages a provider network.

Provider Network: A provider uses a provider network to interconnect Customer Sites. We assume in this document that the provider Network is based on IP or MPLS.

Customer Edge (CE): A device that provides logical connectivity to the provider network. The logical connectivity is enforced at Layer 2 and/or Layer 3 and is denominated an Attachment Circuit. Examples of CEs include TN components (e.g., router, switch, or firewalls) and also 5G Network Function (i.e., an element of 5G domain such as Centralized Unit (CU), Distributed Unit (DU), or User Plane Function (UPF)). This document generalizes the definition of a CE with the introduction of Distributed CEs in Section 3.3.1.

Provider Edge (PE): A device managed by a provider that is connected to a CE. The connectivity between a CE and a PE is achieved using one or multiple Attachment Circuit. This document generalizes the PE definition with the introduction of Distributed PEs in Section 3.3.1.

Attachment Circuit (AC): The logical connection that attaches a CE to a PE. A CE is connected to a PE via one or multiple ACs. An AC is technology-specific. For consistency with the AC data model terminology (e.g., [RFC9182]), we assume that an AC is configured on a “bearer”, which represents the underlying connectivity. Examples of ACs are VLANs (AC) configured on a physical interface

(bearer) or an Overlay VXLAN EVI (AC) configured on IP underlay (bearer).

In order to keep the figures simple, only one AC and single-homed CEs are represented. However, this document does not exclude the instantiation of multiple ACs between a CE and a PE nor the presence of CEs that are attached to more than one PE.

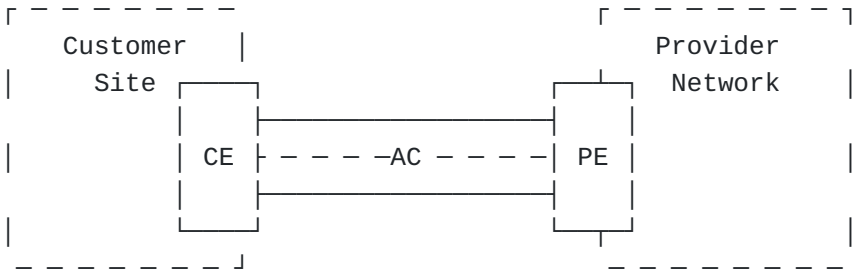
3.3.1. Distributed PE and CE

This document uses the concept of distributed CEs and PEs (e.g., Section 3.4.3 of [RFC4664]). This approach consolidates a definition of CE/PE/AC that is consistent with the orchestration perimeters. The CEs and PEs delimit respectively the customer and provider orchestration domains, while the AC interconnects these domains.

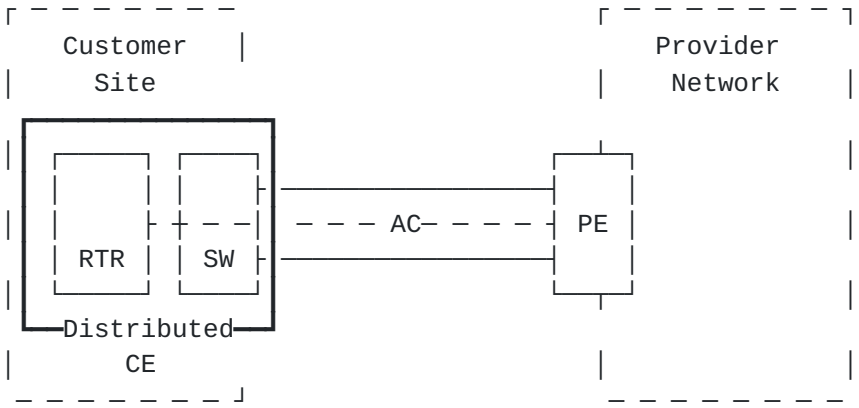
Distributed CE: The logical connectivity is realized by configuring multiple devices in the customer domain. The CE function is distributed. An example of such a distribution is the realization of an interconnection using a L3VPN service based on a distributed CE composed of a switch (Layer 2) and a router (Layer 3) (case (ii) in Figure 3).

Distributed PE: The logical connectivity is realized by configuring multiple devices in the Transport Network (provider orchestration domain). The PE function is distributed. An example of a distributed PE is the “Managed CE service”. For example, a provider delivers VPN services using CEs and PEs which are both managed by the provider (case (iii) in Figure 3). The managed CE can also be a Data Center Gateway as depicted in the example (iv) of Figure 3. A provider-managed CE may attach to CEs of multiple customers. However, this device is part of the provider network.

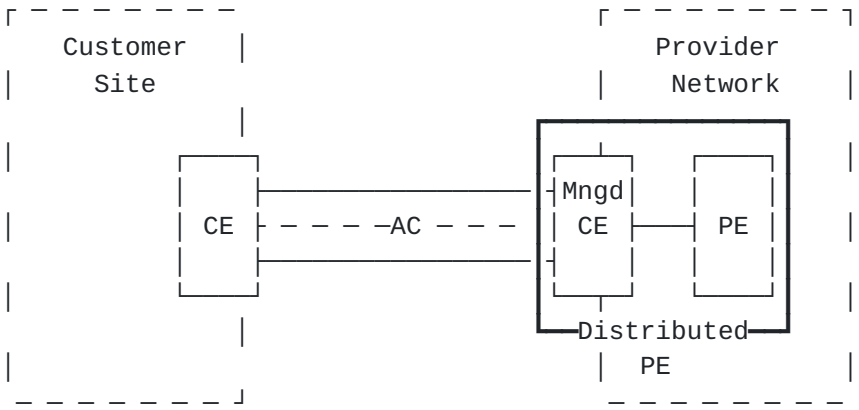
Figure 3 depicts the reference model together with examples of distributed CEs and PEs.



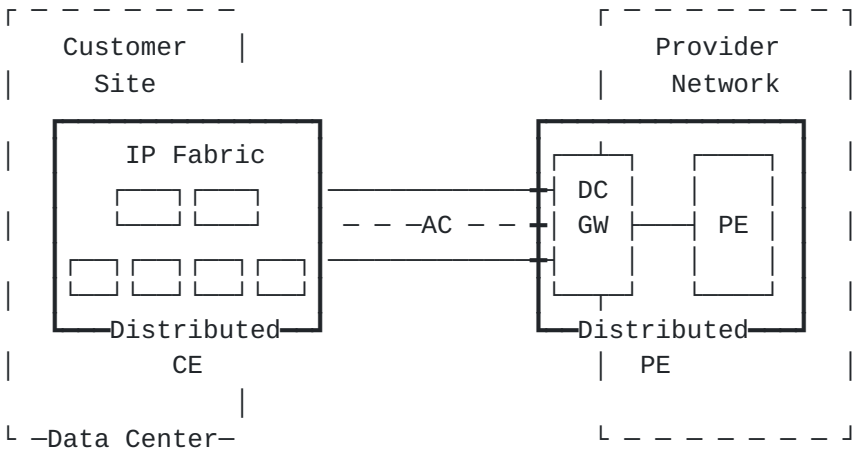
i) Reference Design



ii) Distributed CE



iii) Distributed PE



iv) Distributed PE

and CE

Figure 3: Generic Model vs Distributed CE and PE

In subsequent sections of this document, the terms CE and PE are used for both a single and a distributed devices.

3.3.2. Attachment Circuits for Inter-AS Options B/C

In some cases, a CE connects to the provider network using Inter-AS Option B or C as defined in Section 10 of [RFC4364] with the use of MPLS or SRv6 data planes. An example of such as an AC is depicted in Figure 4. The configuration of VRFs together with control plane identifiers, such as Route Targets (RTs) and Route Distinguishers (RDs), happens on the CE. This is a source of confusion since these configurations are typically enforced on PE devices. Notwithstanding, the reference design based on Orchestration scope prevails: the CE is managed by the customer and the AC is based on MPLS or SRv6 data plane technologies. Note that the complete termination of the AC within the provider network may happen on distinct routers: this is another example of distributed PE (e.g., in Inter-AS Option C, the Autonomous System Border Router (ASBR) and a remote PE in the provider network with VRF configuration form a distributed PE).

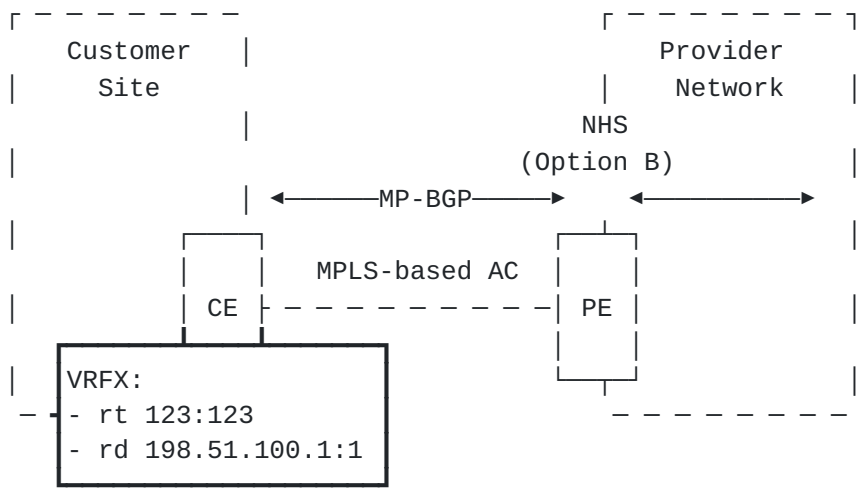


Figure 4: MPLS or SRv6 Attachment Circuit

This use case is also referred to in Section 4.3.2 and Section 4.3.3.

3.3.3. Co-Managed CE

A co-managed CE is orchestrated by both the customer and the provider. In this case, the customer and provider usually have control on distinct device configuration perimeters (e.g., the customer is responsible for the LAN interfaces, while the provider

is responsible for the WAN interfaces (including routing/forwarding policies)). Considering the generic model, a co-managed CE has both PE and CE functions and there is no strict AC connection, although we may consider that the AC stitching logic happens internally within the device. The provider manages the AC between the CE and the PE.

3.4. Orchestration Overview

3.4.1. End-to-End 5G Slice Orchestration Architecture

This section introduces a global framework for the orchestration of an end-to-end 5G Slice with a zoom on TN parts.

> This framework is consistent with the management coordination example shown in Figure 4.7.1 of [TS-28.530].

In reference to Figure 5, an end-to-end 5G Network Slice Orchestrator (5G NSO) is responsible for orchestrating end-to-end 5G Slices. The details of the 5G NSO is out of the scope of this document. The realization of the end-to-end 5G Slice spans RAN, CN, and TN. As mentioned in Section 3.1, the RAN and CN are under the responsibility of the 3GPP Management System, while the TN is not. The orchestration of the TN is split into two sub-domains in conformance with the reference design in {#sec-ref-design}:

* Provider Network Orchestration domain: as defined in [I-D.ietf-teas-ietf-network-slices], the provider relies on an IETF Network Slice Controller (NSC) to manage and orchestrate IETF Network Slices in the provider network. This framework permits to manage connectivity together with SLOs. Ultimately, the 5G NSO interfaces with an NSC for the management of IETF Network Slices using IETF APIs and data models.

* Customer Site Orchestration domain: the Orchestration of TN elements of the Customer Sites relies upon a variety of controllers (e.g., Fabric Manager, Element Management System, or VIM). The realization of this section does not involve the Transport Network Orchestration.

A TN Slice relies upon resources that can involve both the provider and customer TN domains. Therefore, a TN Slice has broader scope than an IETF Network Slice since the latter applies to the provider network only. More details are provided in the next section.

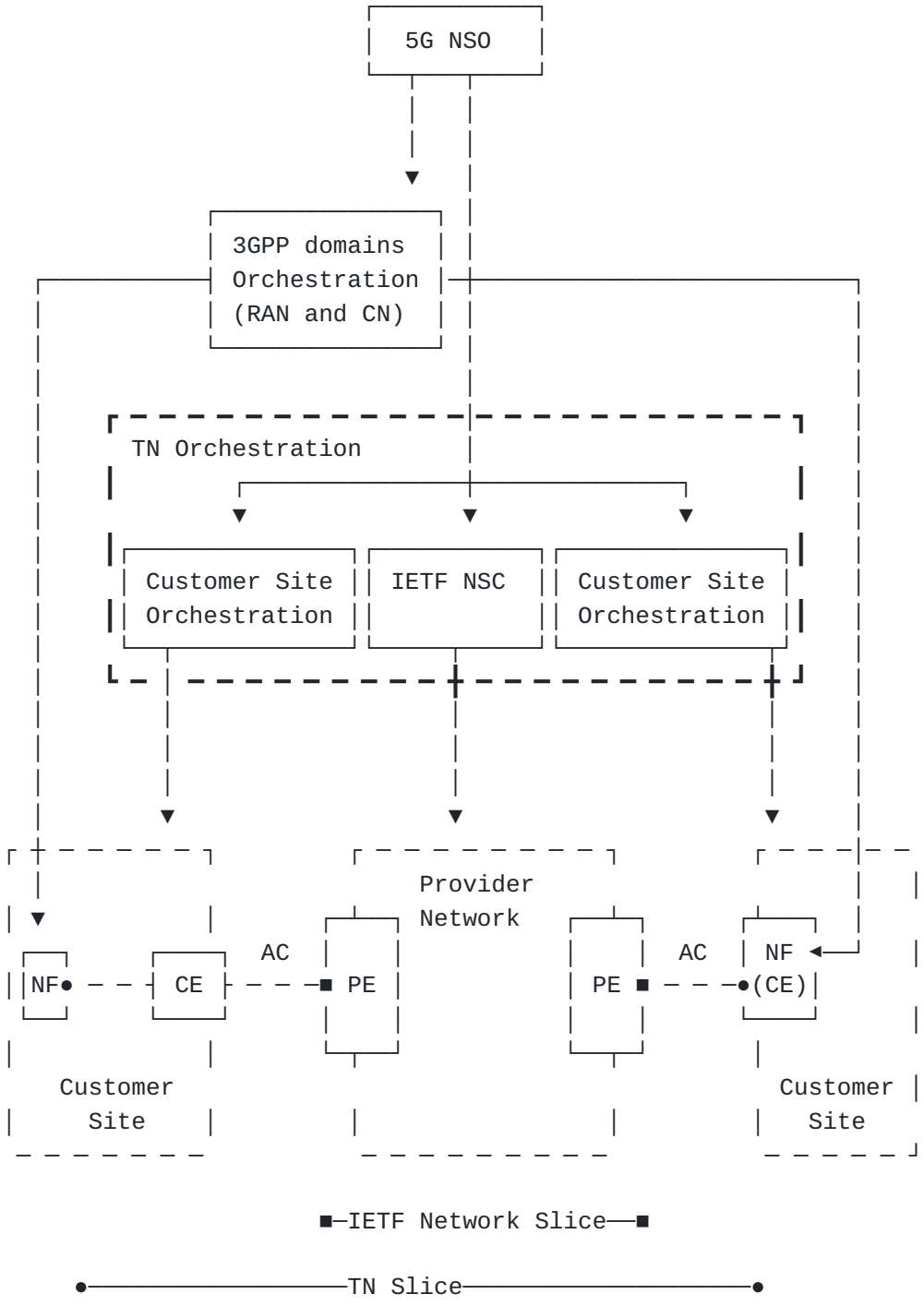


Figure 5: End-to-end 5G Slice Orchestration with TN

3.4.2. Transport Network Sections and Network Slice Instantiation

Based on the reference design, the connectivity between NFs can be decomposed into three main types of sections. Figure 6 depicts the different sections:

- Customer Site: Either connects two NFs located in the same Customer Site (e.g., NF1-NF2) or it connects a NF to a CE (e.g., NF1-CE). This section may not be present if the NF is the CE (e.g., NF3): in this case the AC connects the NF to the PE. The realization of this section is driven by the 5G Network Orchestration and potentially the Customer Site Orchestration (e.g., Fabric Manager, Element Management System, or VIM). The realization of this section does not involve the Transport Network Orchestration.

- Provider Network: Represents the connectivity between two PEs (e.g., PE1-PE2).The realization of this section is controlled by an IETF NSC.

- Attachment Circuit: Represents the connectivity between CEs and PEs (e.g., CE-PE1 and PE2-NF3). The orchestration of this section relies partially upon an IETF NSC for the configuration of the AC on the PE customer-facing interfaces and the Customer Site Orchestration for the configuration of the AC on the CE.

As depicted in Figure 6, the realization of an IETF Network Slice (i.e., connectivity with performance commitments) involves the provider network and partially the AC (the PE-side of the AC). Note that the provisioning of a new network slice may rely on a partial or full pre-provisioned section (e.g., a network slice may rely on an existing AC). Notwithstanding, a framework for the automation of both sections is proposed in this document. The Customer Site section is considered as an extension of the connectivity of the RAN/CN domain without complex slice-specific performances requirements: the Customer Site infrastructure is usually over-provisioned with short distances (low latency) where basic QoS/ Scheduling logic is sufficient to comply with the target SLOs. In other words, the main focus for the enforcement of end-to-end SLOs is managed at the network slice between PE interfaces connected to the AC.

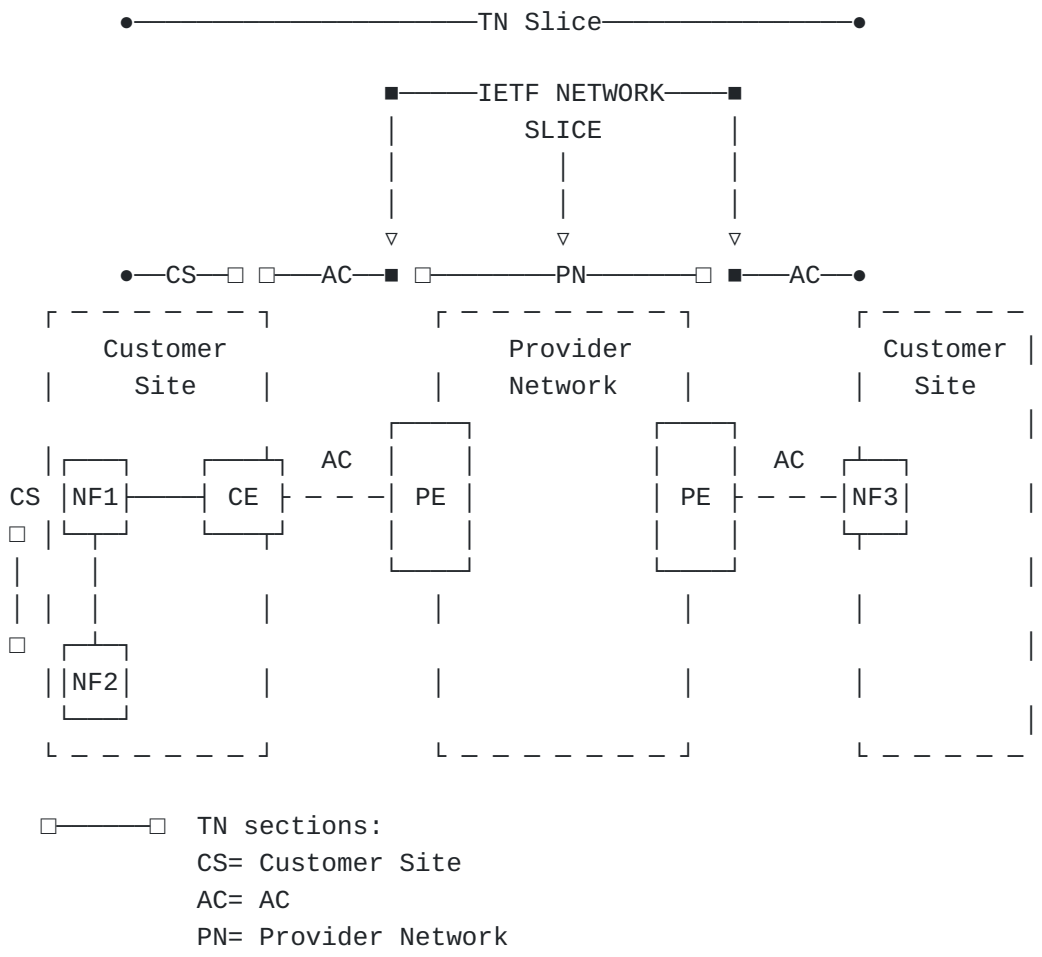


Figure 6: Segmentation of the Transport Network

Resource synchronization for AC provisioning: The realization of the Attachment Circuit is made up of TN resources shared between the Customer Site Orchestration and the provider network orchestration (e.g., NSC). More precisely, a PE and a CE connected via an AC must be provisioned with consistent data plane and control plane information (e.g., VLAN- IDs, IP addresses/subnets, or BGP AS number). Hence, the realization of this interconnection is technology-specific and requires a coordination between the Customer Site Orchestration and an NSC. Automating the provisioning and management of the AC is recommended. Aligned with [RFC8969], we assume that this coordination is based upon standard YANG data models and APIs (more details in further sections).

Figure 7 is a basic example of a Layer 3 CE-PE link realization with shared network resources (such as VLAN-IDs and IP prefixes) which are passed between Orchestrators via a dedicated interface. This document proposes to rely upon IETF service data models: the IETF Network Slice Service Interface

[I-D.ietf-teas-ietf-network-slice-nbi-yang] or the Attachment Circuit Service Interface ([I-D.boro-opsawg-teas-attachment-circuit].

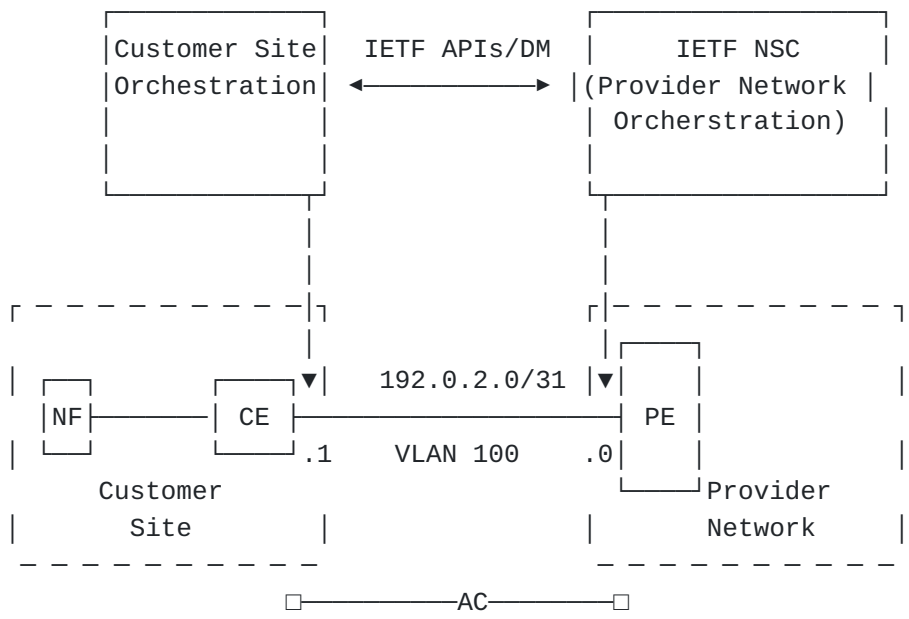


Figure 7: Coordination of TN Resources for the AC Provisioning

3.4.3. Additional Segmentation and Domains

More complex scenarios can happen with extra segmentation of the TN and additional TN Orchestration domains. It is not realistic to describe any design flavor, however the main concepts presented here in terms of segmentation (provider/customer) and stitching (AC) can be reused for the integration of more complex integrations.

3.5. 5G Slice to IETF Network Slice Mapping

Editor Note: This section is intended to focus on the realization implications of the mappings. Will reassess in future versions whether this section should be maintained or moved to [I-D.ietf-teas-5g-network-slice-application].

There are multiple options to map a 5G network slice to IETF Network Slices:

*1 to N: A single 5G Network Slice can map to multiple IETF Network Slices (1 to N). One example of such a case is the separation of the 5G Control Plane and User Plane: this use case is represented in Figure 8 where a slice (eMBB) is deployed with a separation of User Plane and Control Plane at the TN.

*N to 1: Multiple 5G Network Slices may rely upon the same IETF Network Slice. In such a case, the Service Level Agreement (SLA) differentiation of slices would be entirely controlled at 5G Control Plane, for example, with appropriate placement strategies: this use case is represented in Figure 9, where a User Plane Function (UPF) for the URLLC slice is instantiated at the edge cloud close to the gNB CU-UP User Plane for better latency/jitter control, while the 5G Control Plane and the UPF for eMBB slice are instantiated in the regional cloud.

*N to M: The 5G to IETF Network Slice mapping combines both approaches with a mix of shared and dedicated associations.

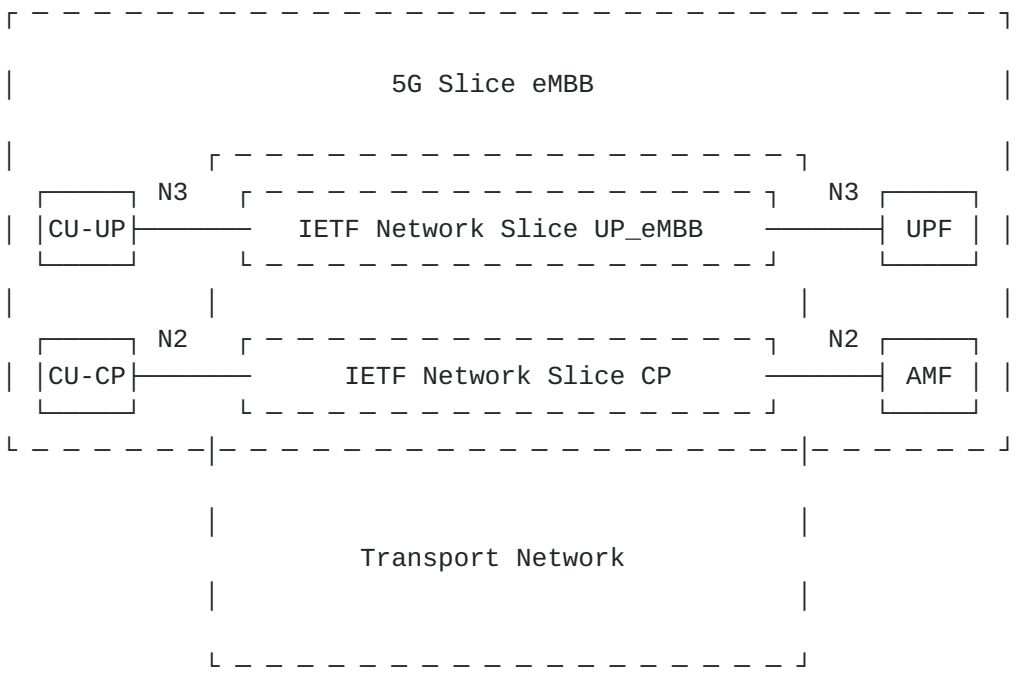


Figure 8: 1 (5G Slice) to N (IETF Network Slice) Mapping

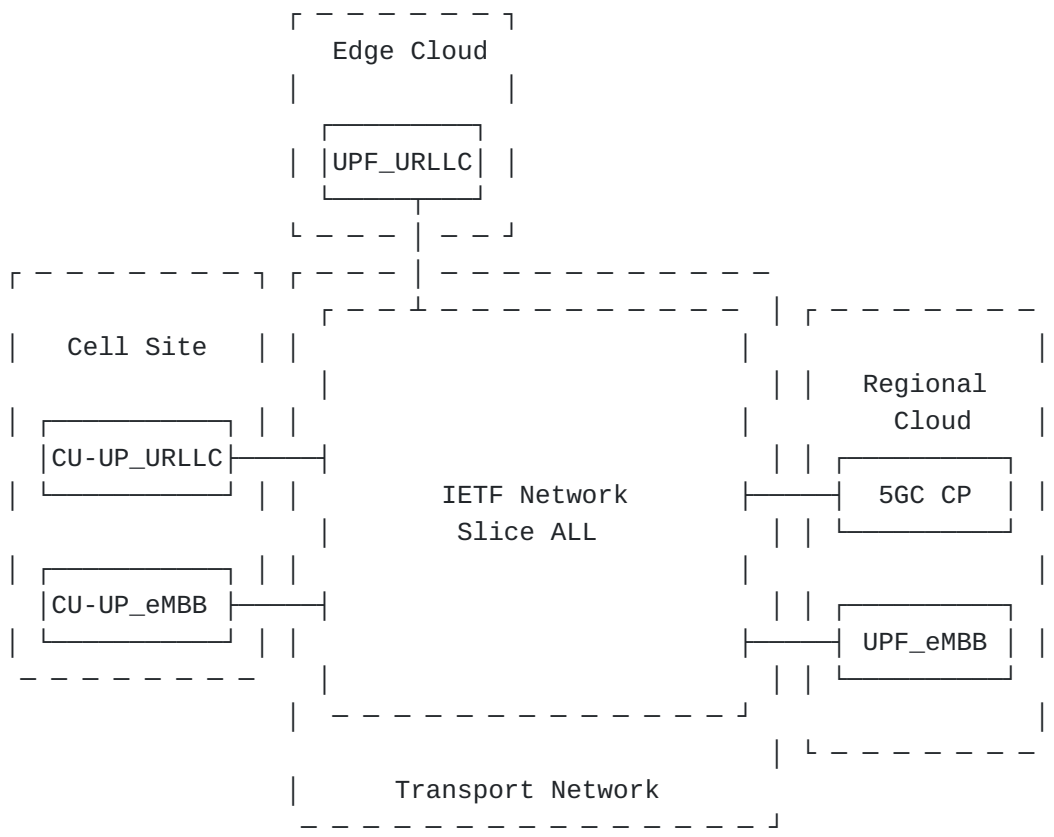


Figure 9: N (5G Slice) to 1 (IETF Network Slice) Mapping

Note that the actual realization of the mapping depends on several factors, such as the actual business cases, the NF vendor capabilities, the NF vendor reference designs, as well as service provider or even legal requirements.

Specifically, the actual mapping is a design choice of service operators that may be a function of, e.g., the number of instantiated slices, requested services, or local engineering capabilities and guidelines. However, operators should carefully consider means to ease slice migration strategies. For example, a provider may initially adopt a 1-to-1 mapping if it has to instantiate few network slices and accommodate the need of few customers. That provider may decide to move to a N-to-1 mapping for aggregation/scalability purposes if sustained increased slice demand is observed. Putting in place adequate automation means to realize network slices (including the adjustment of slice services to network slices mapping) would ease slice migration operations.

3.6. First 5G Slice versus Subsequent Slices

A 5G Network Slice is fully functional with both 5G Control Plane and User Plane capabilities (i.e., dedicated NF functions or contexts). In this regard, the creation of the "first slice" is

subject to a specific logic since it must deploy both CP and UP. This is not the case for the deployment of subsequent slices because they can share the same CP of the first slice, while instantiating dedicated UP. An example of an incremental deployment is depicted in Figure 10.

At the time of writing (2023), Section 6.2 of [NG.113] specifies that the eMBB slice (SST=1 and no Slice Differentiator (SD)) should be supported globally. This 5G slice would be the first slice in any 5G deployment.

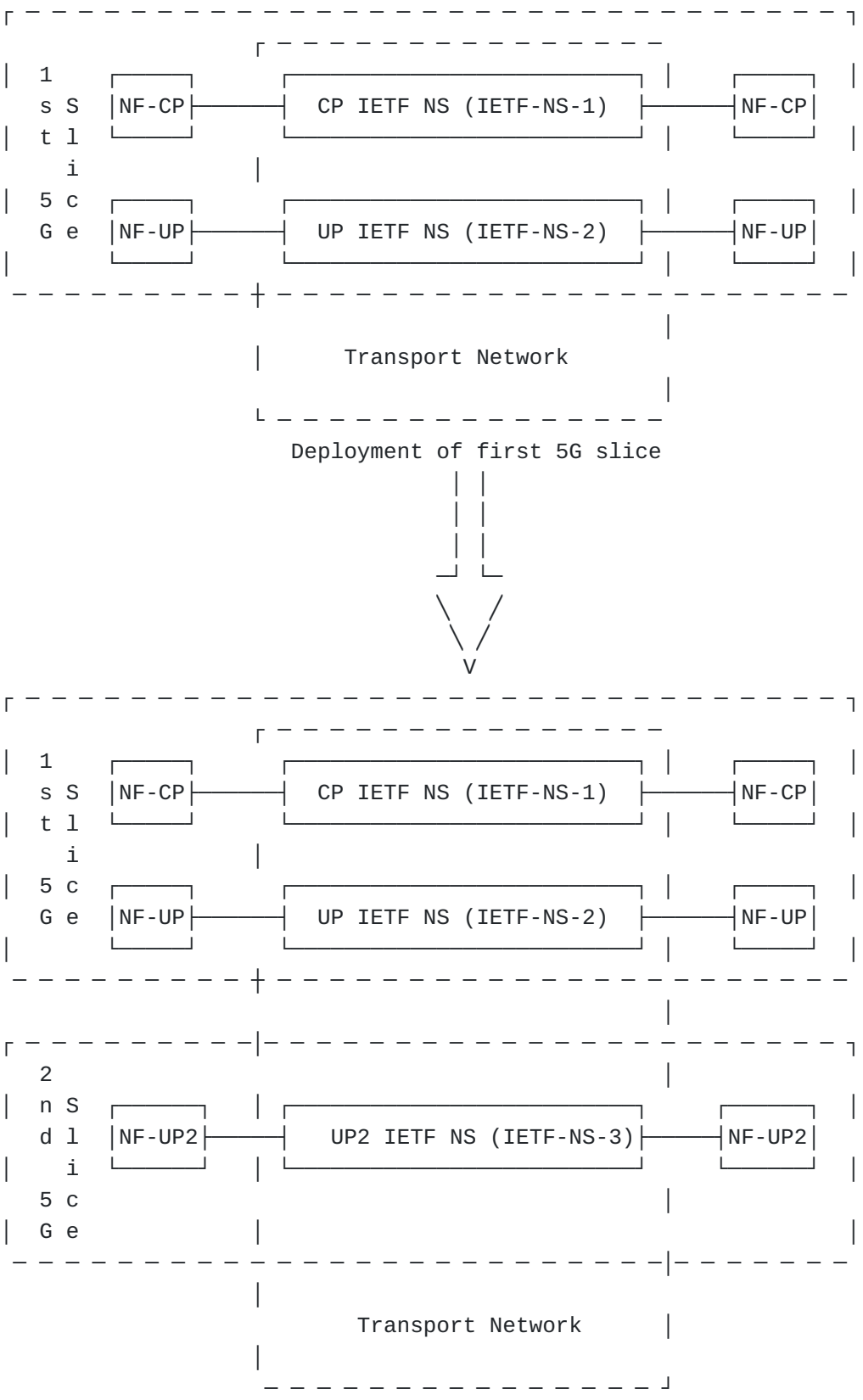


Figure 10: First and Subsequent Slice Deployment

4. Overview of the Realization Model

[I-D.ietf-teas-ietf-network-slices] introduces the concept of the Network Resource Partition (NRP), which is defined as a collection of resources identified in the underlay network. In the basic realization model described in this document, depicted in Figure 11, a single NRP is used with the following characteristics:

*L2VPN/L3VPN service instances for logical separation:

This realization model of transport for 5G slices assumes Layer 3 delivery for midhaul and backhaul transport connections, and a Layer 2 or Layer 3 for fronthaul connections. eCPRI supports both delivery models. L2VPN/L3VPN service instances might be used as a basic form of logical slice separation. Furthermore, using service instances results in an additional outer header (as packets are encapsulated/decapsulated at the nodes hosting service instances) providing clean discrimination between 5G QoS and TN QoS, as explained in Section 5.

*Fine-grained resource control at the PE:

This is sometimes called 'admission control' or 'traffic conditioning'. The main purpose is the enforcement of the bandwidth contract for the slice right at the edge of the provider network where the traffic is handed-off between the customer site and the provider network.

The toolset used here is granular ingress policing (rate limiting) to enforce contracted bandwidths per slice and, potentially, per traffic class within the slice. Out-of-contract traffic might be immediately dropped, or marked as high drop-probability traffic, which is more likely to be dropped somewhere inside the provider network if congestion occurs. In the egress direction at the PE node, hierarchical schedulers/shapers can be deployed, providing guaranteed rates per slice, as well as guarantees per traffic class within each slice.

For managed CEs, edge admission control can be distributed between CEs and PEs, where a part of the admission control is implemented on the CE and other part of the admission control is implemented on the PE.

*Coarse resource control at the transit (non-attachment circuits) links in the provider network, using a single NRP, spanning the entire provider network. Transit nodes in the provider network do not maintain any state of individual slices. Instead, only a flat (non-hierarchical) QoS model is used on transit links in the provider network, with up to 8 traffic classes. At the PE, traffic-flows from multiple slice services are mapped to the

limited number of traffic classes used on provider network transit links.

*Capacity planning/management for efficient usage of provider network resources:

The role of capacity management is to ensure the provider network capacity can be utilized without causing any bottlenecks. The toolset used here can range from careful network planning, to ensure more less equal traffic distribution (i.e., equal cost load balancing), to advanced traffic engineering techniques, with or without bandwidth reservations, to force more consistent load distribution even in non-ECMP friendly network topologies.

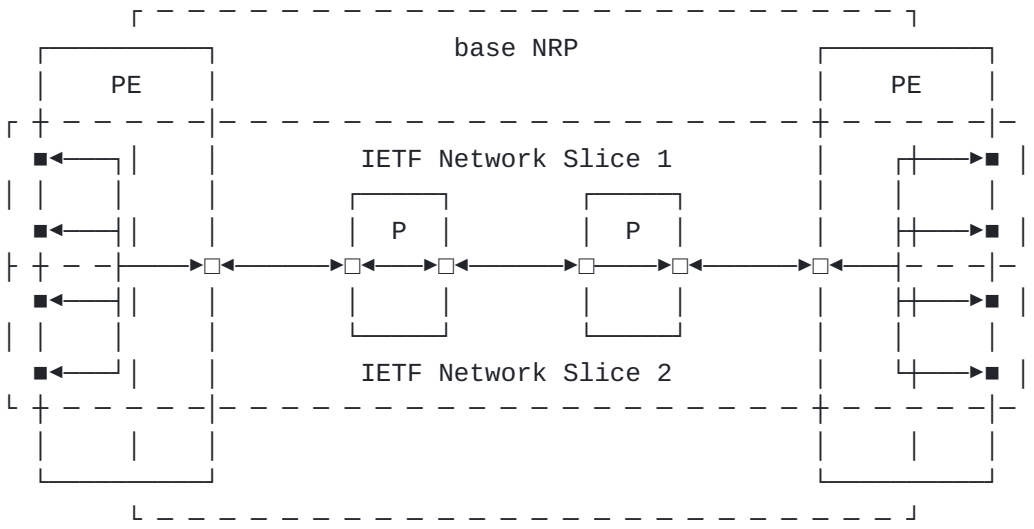


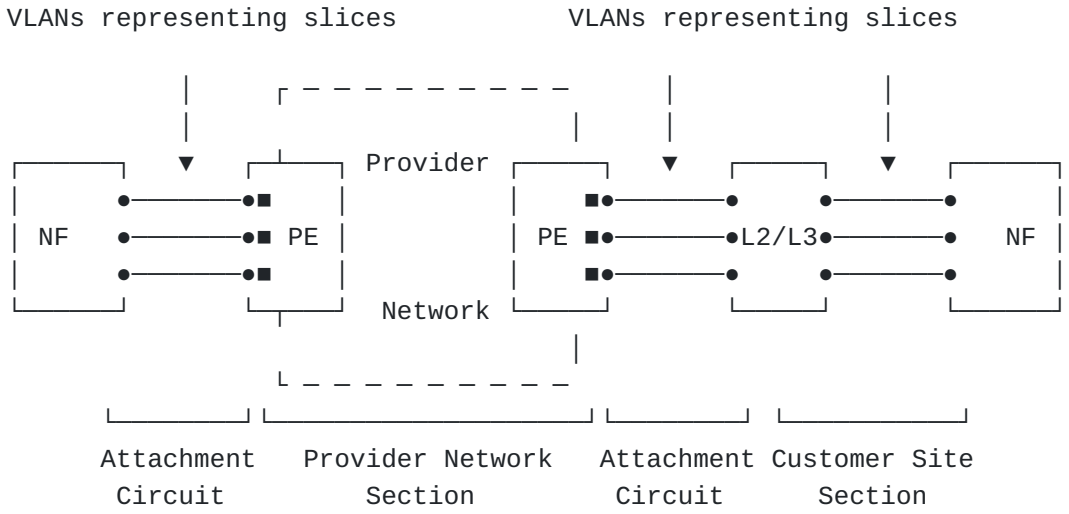
Figure 11: Resource Allocation Slicing Model with a Single NRP

The 5G control plane relies upon the Single Network Slice Selection Assistance Information (S-NSSAI) 32-bit slice identifier for slice identification. The S-NSSAI is not visible to the transport domain. So instead, 5G functions can expose the 5G slices to the transport domain by mapping to explicit Layer 2 or Layer 3 identifiers, such as VLAN-IDs, IP addresses, or Differentiated Services Code Point (DSCP). More details about the mapping between 3GPP and IETF network slices is provided in [I-D.ietf-teas-5g-network-slice-application].

4.1. VLAN Hand-off

In this option, the IETF Network Slice, fulfilling connectivity requirements between NFs of some 5G slice, is represented at the SDP by a VLAN ID (or double VLAN IDs, commonly known as QinQ), as depicted in Figure 12. Each VLAN represents a distinct logical

interface on the attachment circuits, hence it provides the possibility to place these logical interfaces in distinct L2 or L3 service instances and implement separation between slices via service instances. Since the 5G interfaces are IP based interfaces (the only exception could be the F2 fronthaul- interface, where eCPRI with Ethernet encapsulation is used), this VLAN is typically not transported across the provider network. Typically, it has only local significance at a particular SDP. For simplification it is recommended to rely on the same VLAN identifier for all ACs, when possible. However, SDPs for a same slice at different locations may also use different VLAN values. Therefore, a VLAN to IETF Network Slice mapping table is maintained for each AC, and the VLAN allocation is coordinated between customer orchestration and provider orchestration. Thus, while VLAN hand-off is simple from the NF point of view, it adds complexity due to the requirement of maintaining mapping tables for each SDP.



● – logical interface represented by VLAN on physical interface
■ - Service Demarcation Point

Figure 12: 5G Slice with VLAN Hand-off

4.2. IP Hand-off

In this option, the slices in the TN domain are instantiated by IP tunnels (for example, IPsec or GTP-U tunnels) established between NFs, as depicted in Figure 13. The transport for a single 5G slice might be constructed with multiple such tunnels, since a typical 5G slice contains many NFs - especially DUs and CUs. If a shared NF (i.e., an NF that serves multiple slices, for example a shared DU) is deployed, multiple tunnels from shared NF are established, each tunnel representing a single slice. As opposed to the VLAN hand-off case, there is no logical interface representing a slice on the PE,

hence all slices are handled within single service instance. On the other hand, similarly to the VLAN hand-off case, a mapping table tracking IP to IETF Network Slice mapping is required.

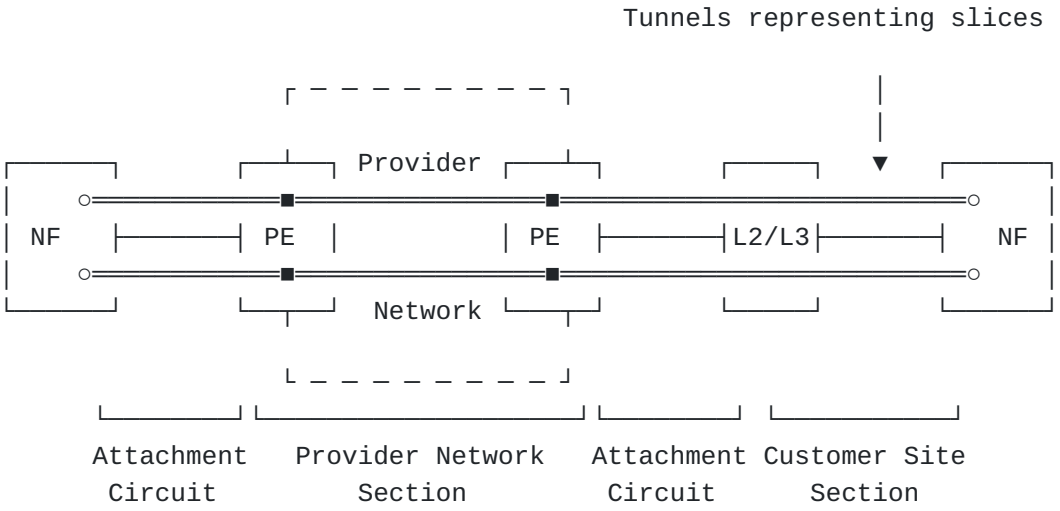


Figure 13: 5G Slice with IP Hand-off

The mapping table can be simplified if, for example, IPv6 addressing is used to address NFs. An IPv6 address is a 128-bit long field, while the S-NSSAI is a 32-bit field: Slice/Service Type (SST): 8 bits, Slice Differentiator (SD): 24 bits. 32 bits, out of 128 bits of the IPv6 address, may be used to encode the S-NSSAI, which makes an IP to Slice mapping table unnecessary. This mapping is simply a local allocation method to allocate IPv6 addresses to NF loopbacks, without redefining IPv6 semantics. Different IPv6 address allocation schemes following this mapping approach may be used, with one example allocation showed in Figure 14.

Note that this addressing scheme is local to an ingress or egress NF; intermediary nodes are not required to associate any additional semantic with IPv6 address.

One benefit of embedding the S-NSSAI in the IPv6 address is that it provides a very easy way of identifying the packet as belonging to given S-NSSAI at any place in the TN domain. This might be used, for example, to selectively enable per S-NSSAI monitoring, or any other per S-NSSAI handling, if required.

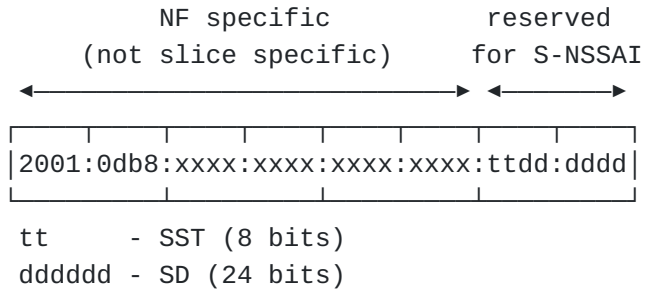


Figure 14: An Example of S-NSSAI embedded into IPv6

In the example shown in Figure 14, the most significant 96 bits of the IPv6 address are unique to the NF, but do not carry any slice-specific information, while the least significant 32 bits are used to embed the S-NSSAI information. The 96-bit part of the address may be further divided based, for example, on the geographical location or the DC identification.

Figure 15 shows an example of a slicing deployment, where the S-NSSAI is embedded into IPv6 addresses used by NFs. NF-A has a set of tunnel termination points, with unique per-slice IP addresses allocated from the 2001:db8::a:0:0/96 prefix, while NF-B uses a set of tunnel termination points with per-slice IP addresses allocated from 2001:db8::b:0:0/96. This example shows two slices: eMBB (SST=1) and MIoT (SST=3). Therefore, for eMBB the tunnel IP addresses are auto- derived (without the need for a mapping table) as {2001:db8::a:100:0, 2001:db8::b:100:0}, while for MIoT (SST=3) tunnel uses {2001:db8::a:300:0, 2001:db8::b:300:0}.

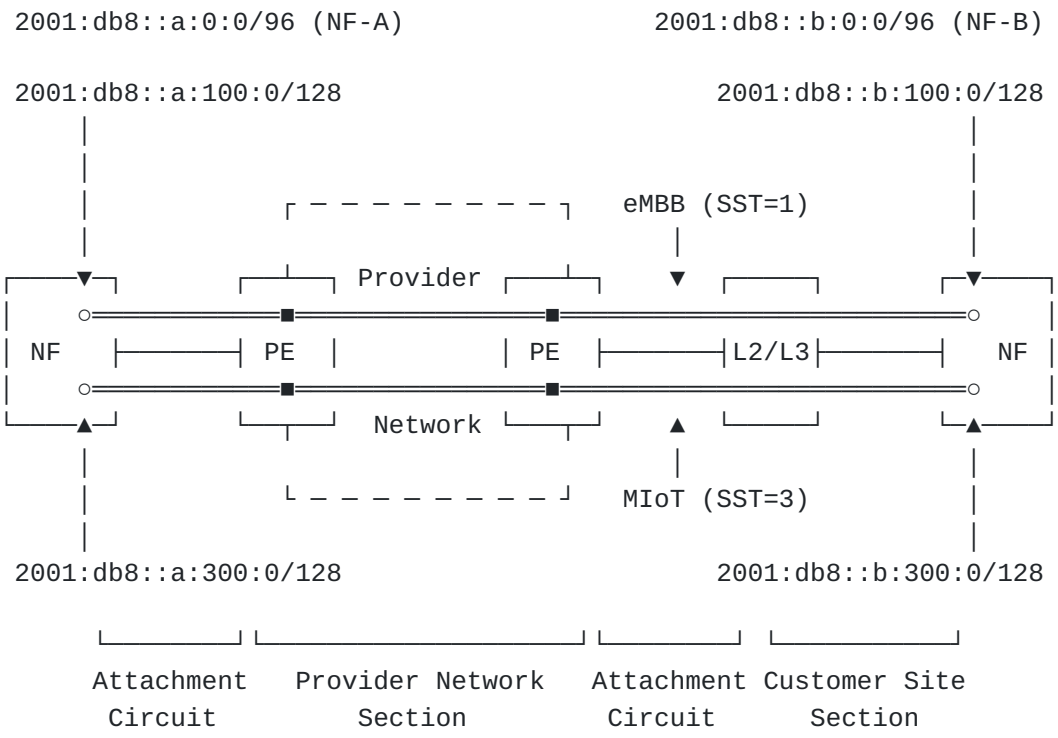


Figure 15: Deployment example with S-NSSAI embedded into IPv6

4.3. MPLS Label Hand-off

In this option, the service instances representing different slices are created directly on the NF, or within the customer site hosting the NF, and attached to the provider network. Therefore, the packet is MPLS encapsulated outside the provider network with native MPLS encapsulation, or MPLSoUDP encapsulation, depending on the capability of the customer site, with the service label depicting the slice.

There are three major methods (based upon Section 10 of [RFC4364]) for interconnecting MPLS services over multiple service domains:

*Option A (Section 4.3.1): VRF-to-VRF connections.

*Option B (Section 4.3.2): redistribution of labeled VPN routes with next-hop change at domain boundaries.

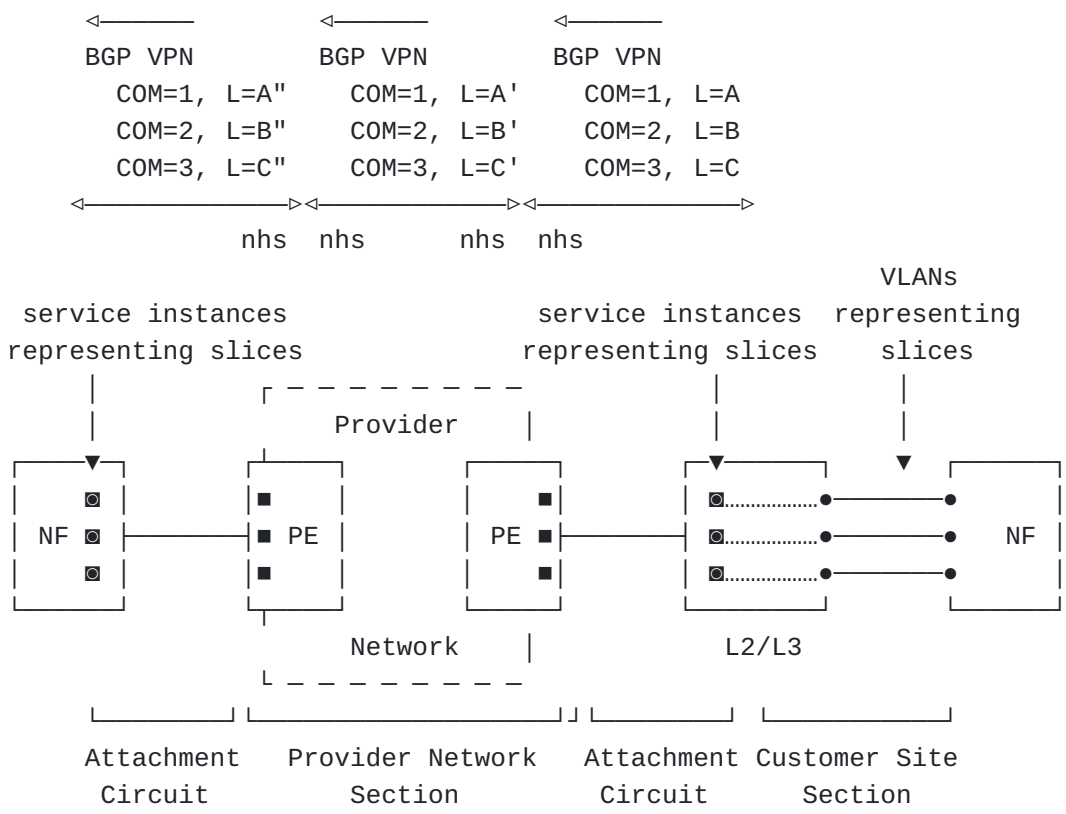
*Option C (Section 4.3.3): redistribution of labeled VPN routes without next-hop change + redistribution of labeled transport routes with next-hop change at domain boundaries.

4.3.1. Option A

This option is not based on MPLS label hand-off, but VLAN hand-off, described in Section 4.1.

4.3.2. Option B

In this option, L3VPN service instances are instantiated outside the provider network. These L3VPN service instances are instantiated in the customer site, which could be for example either on the compute, hosting mobile network functions (Figure 16, left hand side), or within the DC/cloud infrastructure itself (e.g., on the top of the rack or leaf switch within cloud IP fabric (Figure 16, right hand side)). On the attachment circuit connected to PE, packets are already MPLS encapsulated (or MPLSoUDP/MPLSoIP encapsulated, if cloud or compute infrastructure don't support native MPLS encapsulation). Therefore, the PE uses neither a VLAN nor an IP address for slice identification at the SDP, but instead uses the MPLS label.



- ● – logical interface represented by VLAN on physical interface
- ◙ - service instances (with unique MPLS label)
- ■ - Service Demarcation Point

Figure 16: MPLS Hand-off: Option B

MPLS labels are allocated dynamically in Option 10B deployments, where at the domain boundaries service prefixes are reflected with next-hop self, and new label is dynamically allocated, as visible in Figure 16 (e.g., labels A, A' and A" for the first depicted slice). Therefore, for any slice-specific per hop behavior at the provider network edge, the PE must be able to determine which label represents which slice. In the BGP control plane, when exchanging service prefixes over attachment circuit, each slice might be represented by a unique BGP community, so tracking label assignment to the slice is possible. For example, in Figure 16, for the slice identified with COM=1, PE advertises a dynamically allocated label A". Since, based on the community, the label to slice association is known, PE can use this dynamically allocated label A" to identify incoming packets as belonging to slice 1, and execute appropriate edge per hop behavior.

It is worth noting that slice identification in the BGP control plane might be with per-prefix granularity. In extreme case, each prefix can have different community representing a different slice. Depending on the business requirements, each slice could be represented by a different service instance, as outlined in Figure 16. In that case, the route target extended community might be used as slice differentiator. In another deployment, all prefixes (representing different slices) might be handled by single 'mobile' service instance, and some other BGP attribute (e.g., a standard community) might be used for slice differentiation. Or there could be a deployment that groups multiple slices together into a single service instance, resulting in a handful of service instances. In any case, fine-grained per-hop behavior at the edge of provider network is possible.

4.3.3. Option C

for further study

5. QoS Mapping Realization Models

5.1. QoS Layers

The resources are managed via various QoS policies deployed in the network. QoS mapping models to support 5G slicing connectivity implemented over packet switched provider network uses two layers of QoS that are discussed in Section 5.1.

5.1.1. 5G QoS Layer

QoS treatment is indicated in the 5G QoS layer by the 5QI (5G QoS indicator), as defined in [TS-23.501]. A 5QI is an identifier (ID) that is used as a reference to 5G QoS characteristics (e.g., scheduling weights, admission thresholds, queue management

thresholds, and link layer protocol configuration) in the RAN domain. Given that 5QI applies to the RAN domain, it is not visible to the provider network. Therefore, if 5QI-aware treatment is desired in the provider network as well, 5G network functions might set DSCP with a value representing 5QI so that differentiated treatment can implemented in the provider network as well. Based on these DSCP values, at SDP of each provider network section used to construct transport for given 5G slice, very granular QoS enforcement might be implemented.

The exact mapping between 5QI and DSCP is out of scope for this document. Mapping recommendations are documented, e.g., in [I-D.henry-tsvwg-diffserv-to-qci].

Each slice service might have flows with multiple 5QIs, thus there could be many different 5QIs being deployed. 5QIs (or, more precisely, corresponding DSCP values) are visible to the provider network at SDP (i.e., at the edge of the provider network).

In this document, this layer of QoS will be referred as '5G QoS Class' ('5G QoS' in short), or '5G DSCP'.

5.1.2. TN QoS Layer

Control of the TN resources on provider network transit links, as well as traffic scheduling/prioritization on provider network transit links, is based on a flat (non-hierarchical) QoS model in this IETF Network Slice realization. That is, IETF Network Slices are assigned dedicated resources (e.g., QoS queues) at the edge of the provider network (at SDPs), while all IETF Network Slices are sharing resources (sharing QoS queues) on the transit links of the provider network. Typical router hardware can support up to 8 traffic queues per port, therefore the architecture assumes 8 traffic queues per port support in general.

At this layer, QoS treatment is indicated by QoS indicator specific to the encapsulation used in the provider network, and it could be DSCP or MPLS Traffic Class (TC). This layer of QoS will be referred as 'TN QoS Class', or 'TN QoS' for short, in this document.

5.2. QoS Realization Models

While 5QI might be exposed to the provider network, via the DSCP value (corresponding to specific 5QI value) set in the IP packet generated by NFs, some 5G deployments might use 5QI in the RAN domain only, without requesting per 5QI differentiated treatment from the provider network. This can be due to an NF limitation (e.g., no capability to set DSCP), or it might simply depend on the overall slicing deployment model. The O-RAN Alliance, for example, defines a phased approach to the slicing, with initial phases

utilizing only per slice, but not per 5QI, differentiated treatment in the TN domain (Annex F of [O-RAN.WG9.XPSAAS]).

Therefore, from a QoS perspective, the 5G slicing connectivity realization architecture defines two high-level realization models for slicing in the TN domain: a 5QI-unaware model and a 5QI- aware model. Both slicing models in the TN domain could be used concurrently within the same 5G slice. For example, the TN segment for 5G midhaul (F1-U interface) might be 5QI-aware, while at the same time the TN segment for 5G backhaul (N3 interface) might follow the 5QI-unaware model.

These models are further elaborated in the following two subsections.

5.2.1. 5QI-unaware Model

In 5QI-unaware mode, the DSCP values in the packets received from NF at SDP are ignored. In the provider network, there is no QoS differentiation at the 5G QoS Class level. The entire IETF Network Slice is mapped to single TN QoS Class, and, therefore, to a single QoS queue on the routers in the provider network. With a small number of deployed 5G slices (for example only two 5G slices: eMBB and MIoT), it is possible to dedicate a separate QoS queue for each slice on transit routers in the provider network. However, with introduction of private/enterprises slices, as the number of 5G slices (and thus corresponding IETF Network Slices) increases, a single QoS queue on transit links in the provider network serves multiple slices with similar characteristics. QoS enforcement on transit links is fully coarse (single NRP, sharing resources among all IETF Network Slices), as displayed in Figure 17.

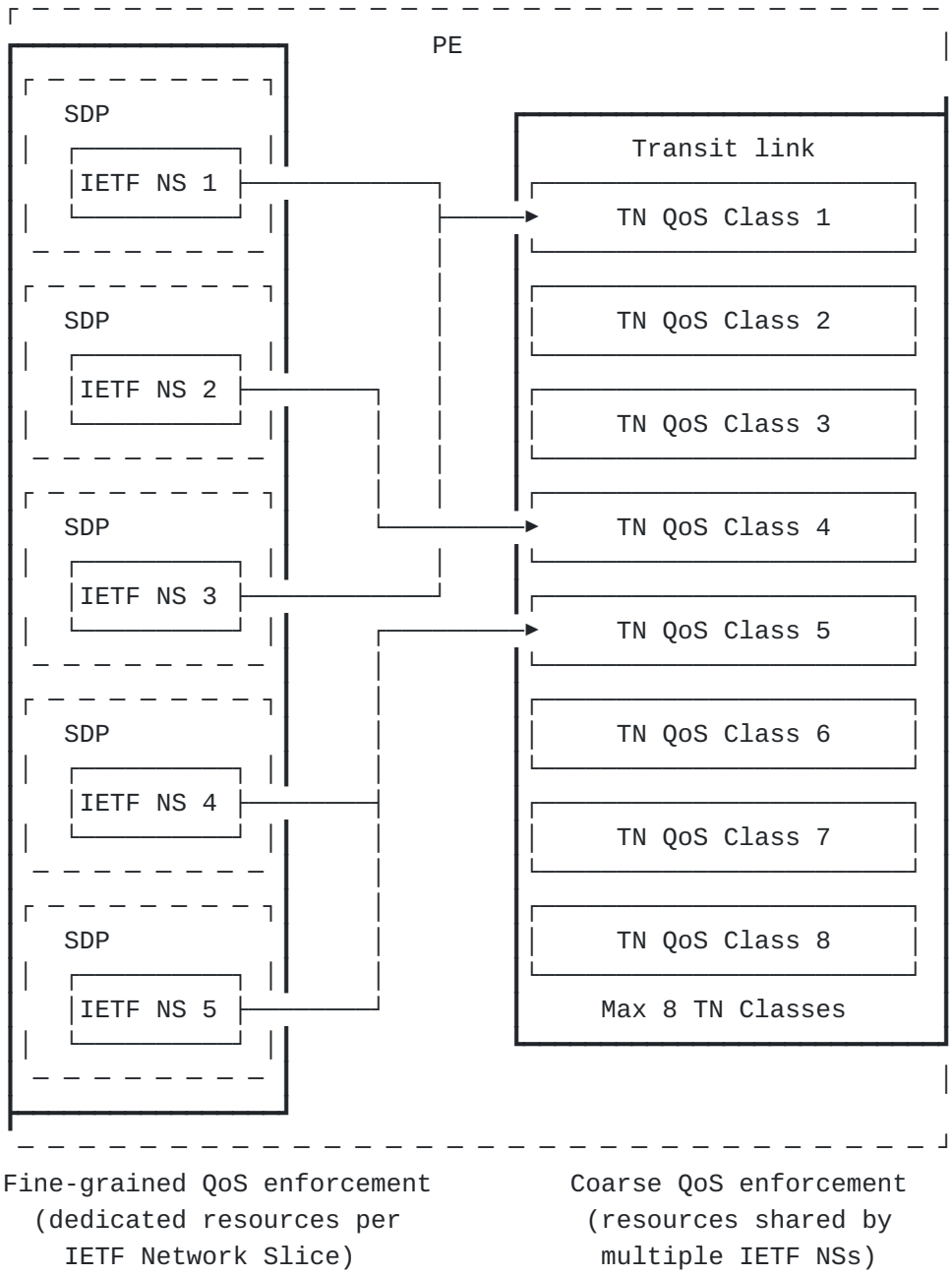


Figure 17: Slice to TN QoS Mapping (5QI-unaware Model)

When the IP traffic is handed over at the SDP from the attachment circuit to the provider network, the PE encapsulates the traffic into MPLS (if MPLS transport is used in the provider network), or IPv6 - optionally with some additional headers (if SRv6 transport is used in the provider network), and sends out the packets on the provider network transit link.

The original IP header retains the DCSP marking (which is ignored in 5QI-unaware model), while the new header (MPLS or IPv6) carries QoS marking (MPLS Traffic Class bits for MPLS encapsulation, or DSCP for SRv6/IPv6 encapsulation) related to TN CoS. Based on TN QoS Class

marking, per hop behavior for all IETF Network Slices is executed on provider network transit links. Provider network transit routers do not evaluate the original IP header for QoS-related decisions. This model is outlined in Figure 18 for MPLS encapsulation, and in Figure 19 for SRv6 encapsulation.

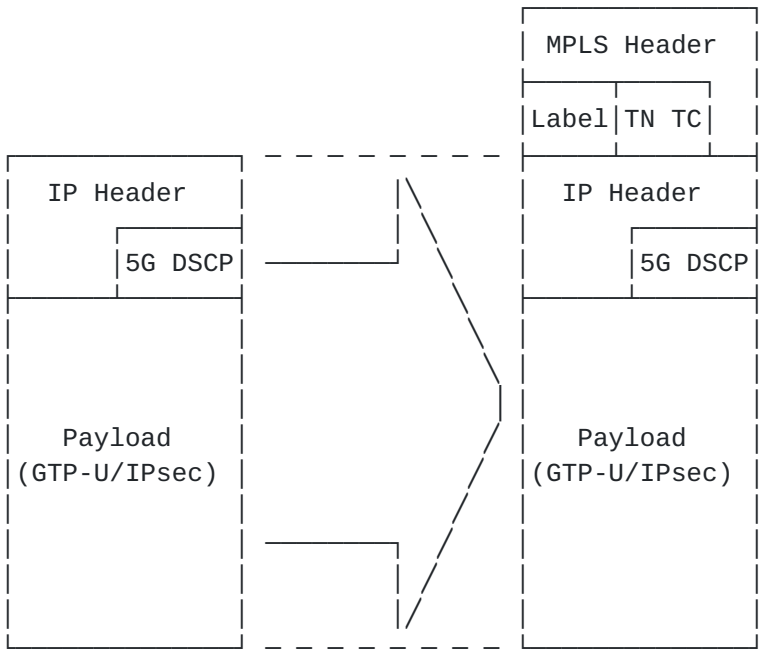


Figure 18: QoS with MPLS Encapsulation

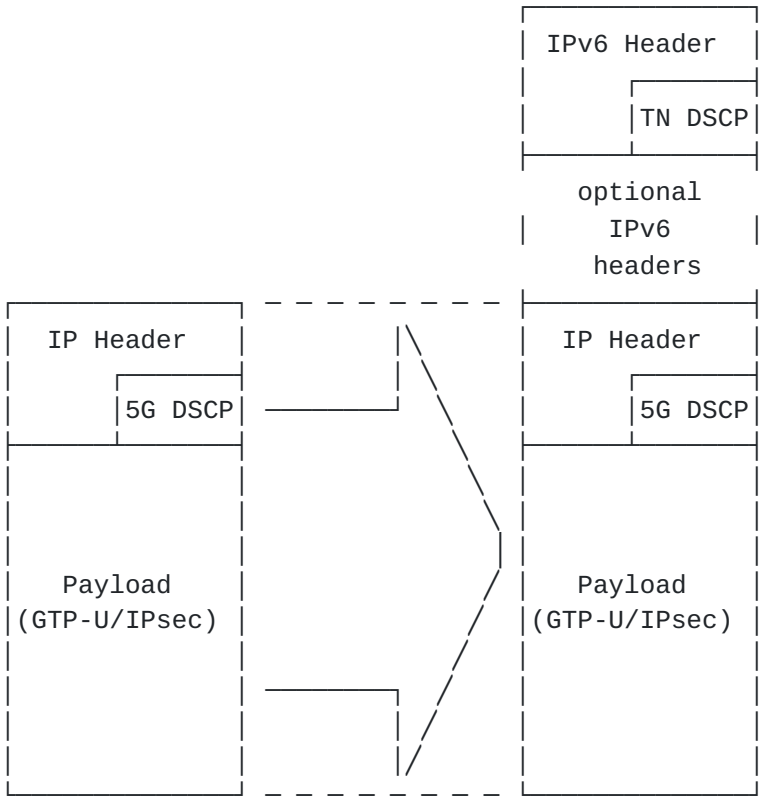


Figure 19: QoS with IPv6 Encapsulation

From the QoS perspective, both options are similar. However, there is one difference between the two options. The MPLS TC is only 3 bits (8 possible combinations), while DSCP is 6 bits (64 possible combinations). Hence, SRv6 [RFC8754] provides more flexibility for TN CoS design, especially in combination with soft policing with in-profile/ out-profile traffic, as discussed in Section 5.2.1.1.

Provider network edge resources are controlled in a granular, fine-grained manner, with dedicated resource allocation for each IETF Network Slice. The resource control/enforcement happens at each SDP in two directions: inbound and outbound.

5.2.1.1. Inbound Edge Resource Control

The main aspect of inbound provider network edge resource control is per-slice traffic capacity enforcement. This kind of enforcement is often called 'admission control' or 'traffic conditioning'. The goal of this inbound enforcement is to ensure that the traffic above the contracted rate is dropped or deprioritized, depending on the business rules, right at the edge of provider network. This, combined with appropriate network capacity planning/management (Section 7) is required to ensure proper isolation between slices in a scalable manner. As a result, traffic of one slice has no influence on the traffic of other slices, even if the slice is

misbehaving (e.g., DDoS attacks or node/link failures) and generates traffic volumes above the contracted rates.

The slice rates can be characterized with following parameters [I-D.ietf-teas-ietf-network-slice-nbi-yang]:

*CIR: Committed Information Rate (i.e., guaranteed bandwidth)

*PIR: Peak Information Rate (i.e., maximum bandwidth)

These parameters define the traffic characteristics of the slice and are part of SLO parameter set provided by the 5G NSO to IETF NSC. Based on these parameters the provider network inbound policy can be implemented using one of following options:

*1r2c (single-rate two-color) rate limiter

This is the most basic rate limiter, which meters at the SDP a traffic stream of given slice and marks its packets as in-contract (below contracted CIR) or out-of-contract (above contracted CIR). In-contract packets are accepted and forwarded. Out-of contract packets are either dropped right at the SDP (hard rate limiting), or remarked (with different MPLS TC or DSCP TN markings) to signify 'this packet should be dropped in the first place, if there is a congestion' (soft rate limiting), depending on the business policy of the provider network. In the second case, while packets above CIR are forwarded at the SDP, they are subject to being dropped during any congestion event at any place in the provider network.

*2r3c (two-rate three-color) rate limiter

This was initially defined in [RFC2698], and its improved version in [RFC4115]. In essence, the traffic is assigned to one of the these three categories:

-Green, for traffic under CIR

-Yellow, for traffic between CIR and PIR

-Red, for traffic above PIR

An inbound 2r3c meter implemented with [RFC4115], compared to [RFC2698], is more 'customer friendly' as it doesn't impose outbound peak-rate shaping requirements on customer edge (CE) devices. 2r3c meters in general give greater flexibility for provider network edge enforcement regarding accepting the traffic (green), de- prioritizing and potentially dropping the traffic on transit during congestion (yellow), or hard dropping the traffic (red).

Inbound provider network edge enforcement model for 5QI-unaware model, where all packets belonging to the slice are treated the same way in the provider network (no 5Q QoS Class differentiation in the provider) is outlined in Figure 20.

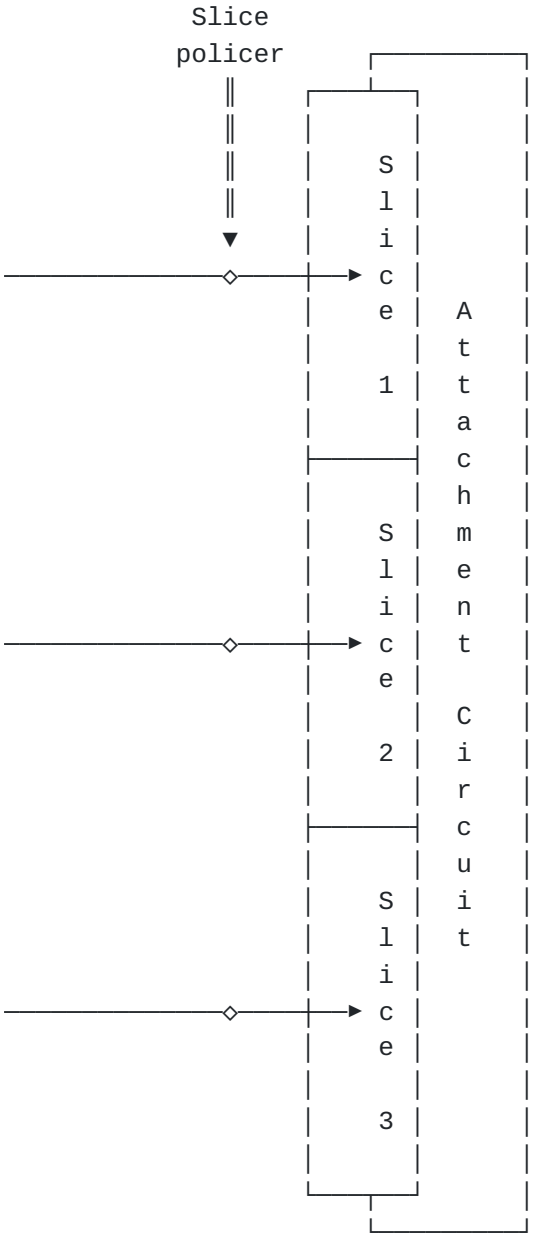


Figure 20: Ingress Slice Admission Control (5QI-unware Model)

5.2.1.2. Outbound Edge Resource Control

While inbound slice admission control at the provider network edge is mandatory in the architecture described in this document, outbound provider network edge resource control might not be

required in all use cases. Use cases that specifically call for outbound provider network edge resource control are:

* Slices use both CIR and PIR parameters, and provider network edge links (attachment circuits) are dimensioned to fulfil the aggregate of slice CIRs. If at any given time, some slices send the traffic above CIR, congestion in outbound direction on the provider network edge link (attachment circuit) might happen. Therefore, fine-grained resource control to guarantee at least CIR for each slice is required.

* Any-to-Any (A2A) connectivity constructs are deployed, again resulting in potential congestion in outbound direction on the provider network edge links, even if only slice CIR parameters are used. This again requires fine-grained resource control per slice in outbound direction at the provider network edge links.

As opposed to inbound provider network edge resource control, typically implemented with rate-limiters/policers, outbound resource control is typically implemented with a weighted/priority queuing, potentially combined with optional shapers (per slice). A detailed analysis of different queuing mechanisms is out of scope for this document, but is provided in [RFC7806].

Figure 21 outlines the outbound provider network edge resource control model for 5QI-unaware slices. Each slice is assigned a single egress queue. The sum of slice CIRs, used as the weight in weighted queueing model, must not exceed the physical capacity of the attachment circuit. Slice requests above this limit must be rejected by the IETF NSC, unless an already established slice with lower priority, if such exists, is preempted.

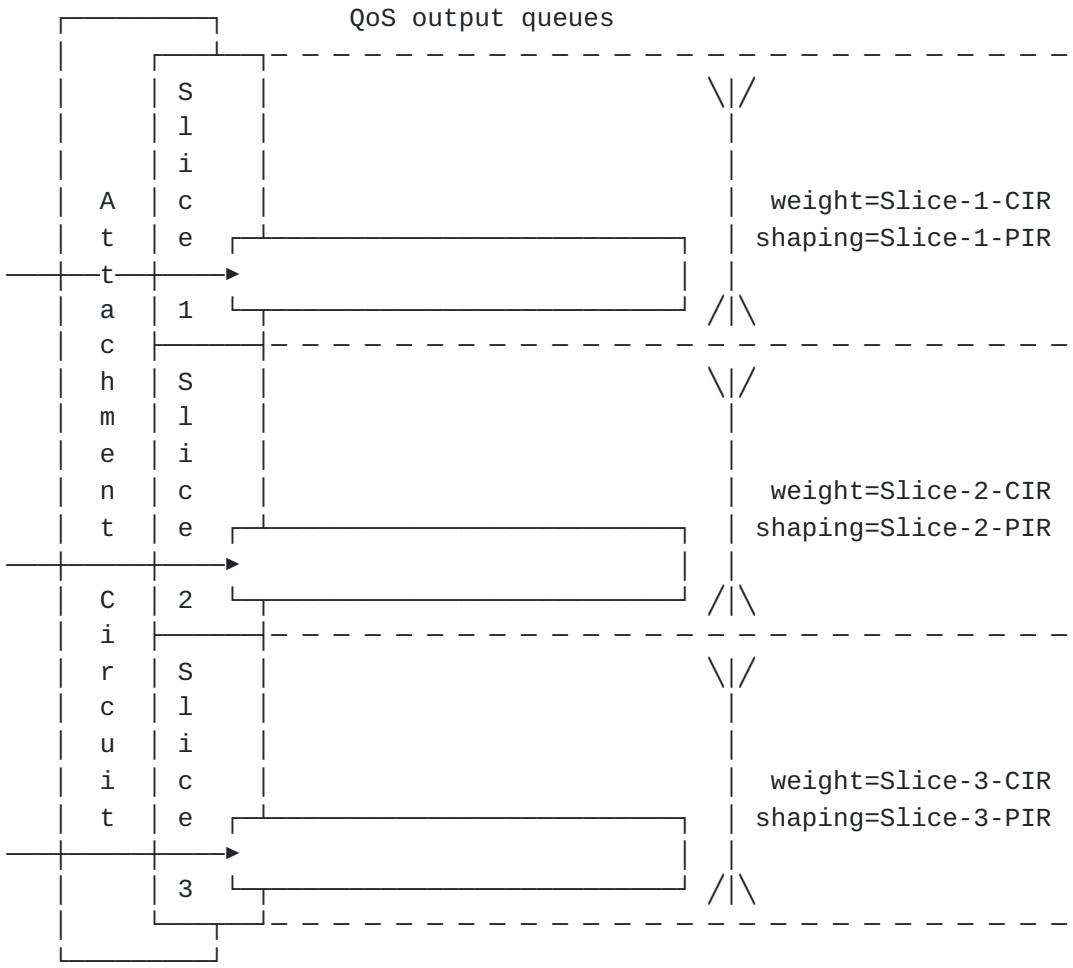


Figure 21: Ingress Slice Admission control (5QI-unaware Model)

5.2.2. 5QI-aware Model

In the 5QI-aware model, potentially a large number of 5G QoS Classes, represented via DSCP set by NFs (the architecture scales to thousands of 5G slices) is mapped (multiplexed) to up to 8 TN QoS Classes used in provider network transit equipment, as outlined in Figure 22.

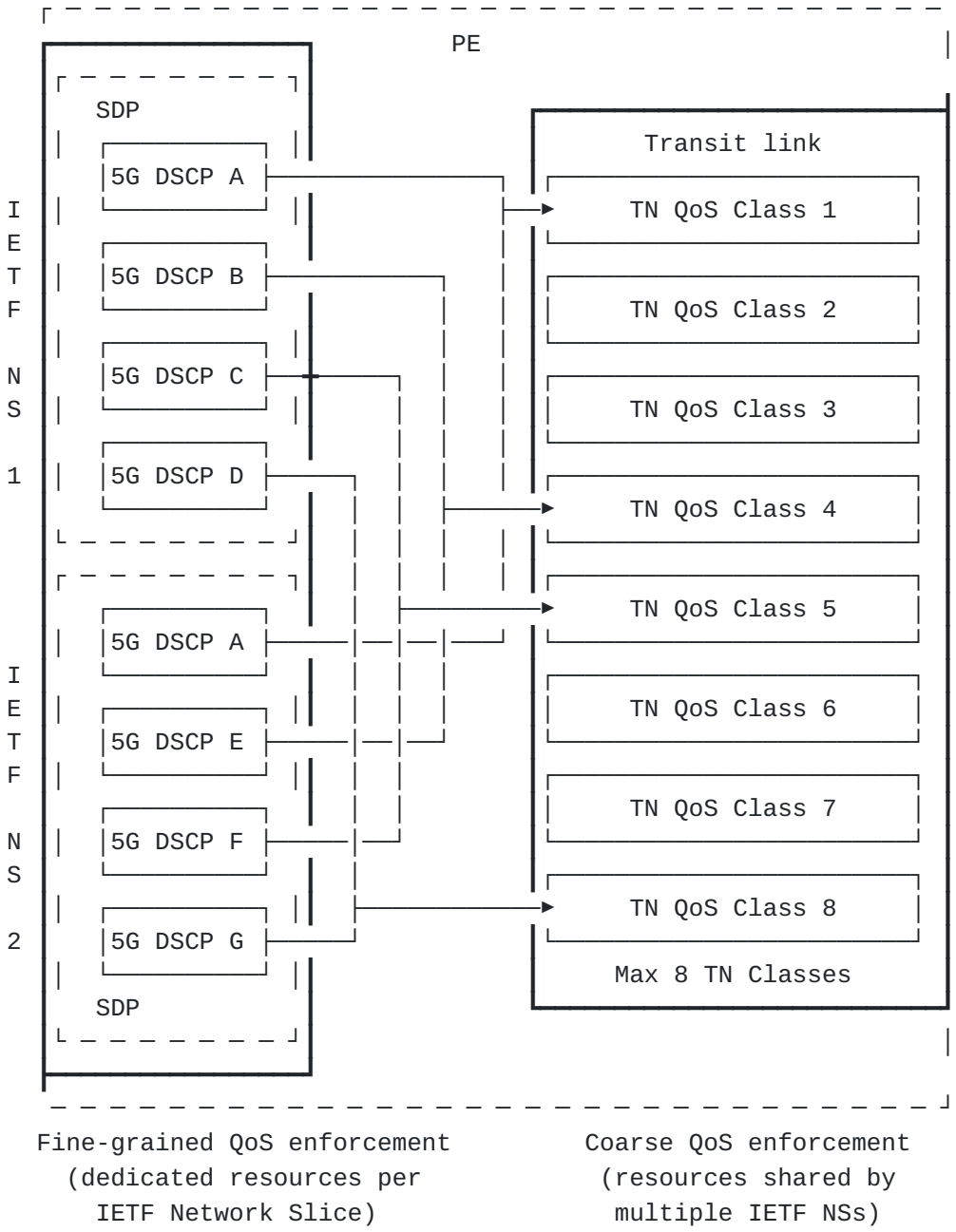


Figure 22: Slice 5Q QoS to TN QoS Mapping (5QI-aware Model)

Given that in large scale deployments (large number of 5G slices), the number of potential 5G QoS Classes is much higher than the number of TN QoS Classes, multiple 5G QoS Classes with similar characteristics - potentially from different slices - would be grouped with common operator-defined TN logic and mapped to a same TN QoS Class when transported in the provider network. That is, common per hop behavior (PHB) is executed on transit provider network routers for all packets grouped together. An example of this approach is outlined in Figure 23.

Note:

The numbers indicated in Figure 23 (S-NSSAI, 5QI, DSCP, queue, etc.) are provided for illustration purposes only and should not be considered as deployment guidance.

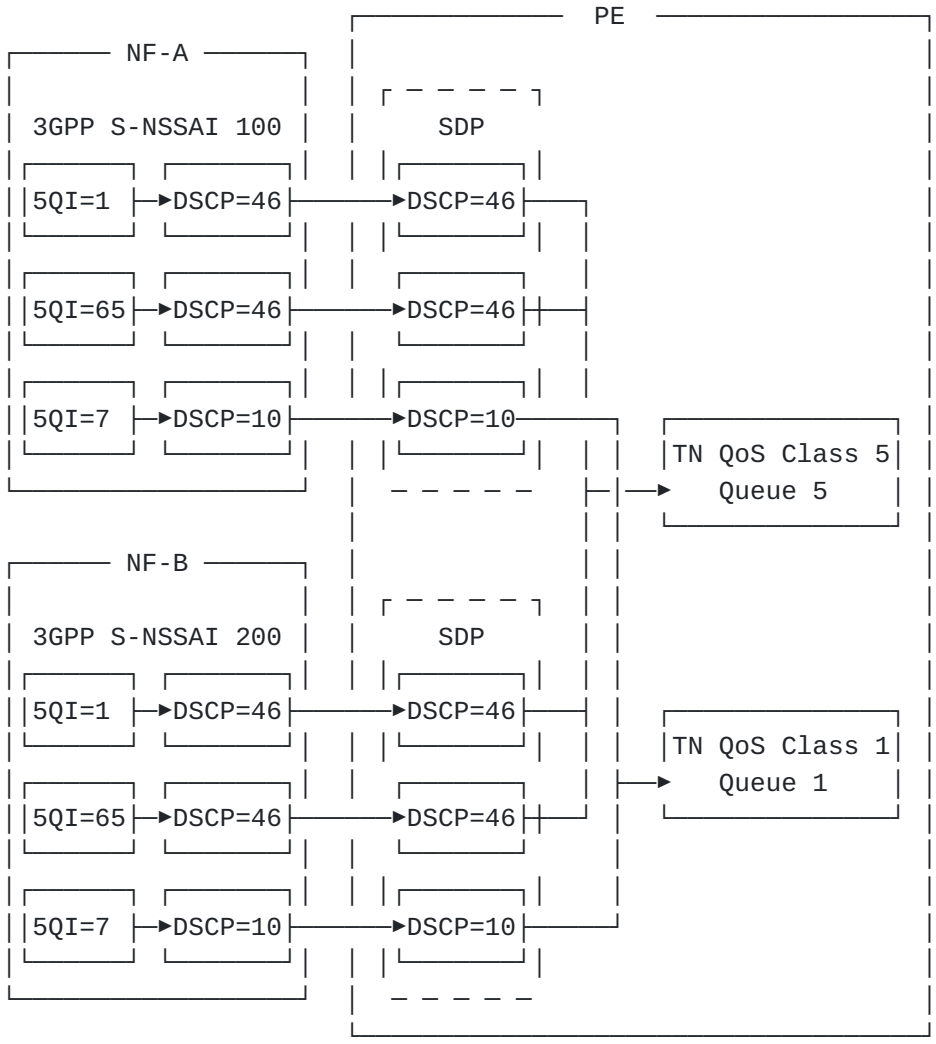


Figure 23: Example of 3GPP QoS Mapped to TN QoS

In current SDO progress of 3GPP (Rel.17) and O-RAN the mapping of 5QI to DSCP is not expected in per-slice fashion, where 5QI to DSCP mapping may vary from 3GPP slice to 3GPP slice, hence the mapping of 5G QoS DSCP values to TN QoS Classes may be rather common.

Like in 5QI-unaware model, the original IP header retains the DCSP marking corresponding to 5QI (5G QoS Class), while the new header (MPLS or IPv6) carries QoS marking related to TN QoS Class. Based on TN QoS Class marking, per hop behavior for all aggregated 5G QoS Classes from all IETF Network Slices is executed on provider network transit links. Provider network transit routers do not evaluate original IP header for QoS related decisions. The original DSCP marking retained in the original IP header is used at the PE for

fine-grained per slice and per 5G QoS Class inbound/outbound enforcement on the AC.

In 5QI-aware model, compared to 5QI-unware model, provider network edge resources are controlled in an even more granular, fine-grained manner, with dedicated resource allocation for each IETF Network Slice and dedicated resource allocation for number of traffic classes (most commonly up 4 or 8 traffic classes, depending on the HW capability of the equipment) within each IETF Network Slice.

5.2.2.1. Inbound Edge Resource Control

Compared to the 5QI-unware model, admission control (traffic conditioning) in the 5QI-aware model is more granular, as it enforces not only per slice capacity constraints, but may as well enforce the constraints per 5G QoS Class within each slice.

5G slice using multiple 5QIs can potentially specify rates in one of the following ways:

* Rates per traffic class (CIR or CIR+PIR), no rate per slice (sum of rates per class gives the rate per slice).

* Rate per slice (CIR or CIR+PIR), and rates per prioritized (premium) traffic classes (CIR only). Best effort traffic class uses the bandwidth (within slice CIR/PIR) not consumed by prioritized classes.

In the first option, the slice admission control is executed with traffic class granularity, as outlined in Figure 24. In this model, if a premium class doesn't consume all available class capacity, it cannot be reused by non-premium (i.e., Best Effort) class.

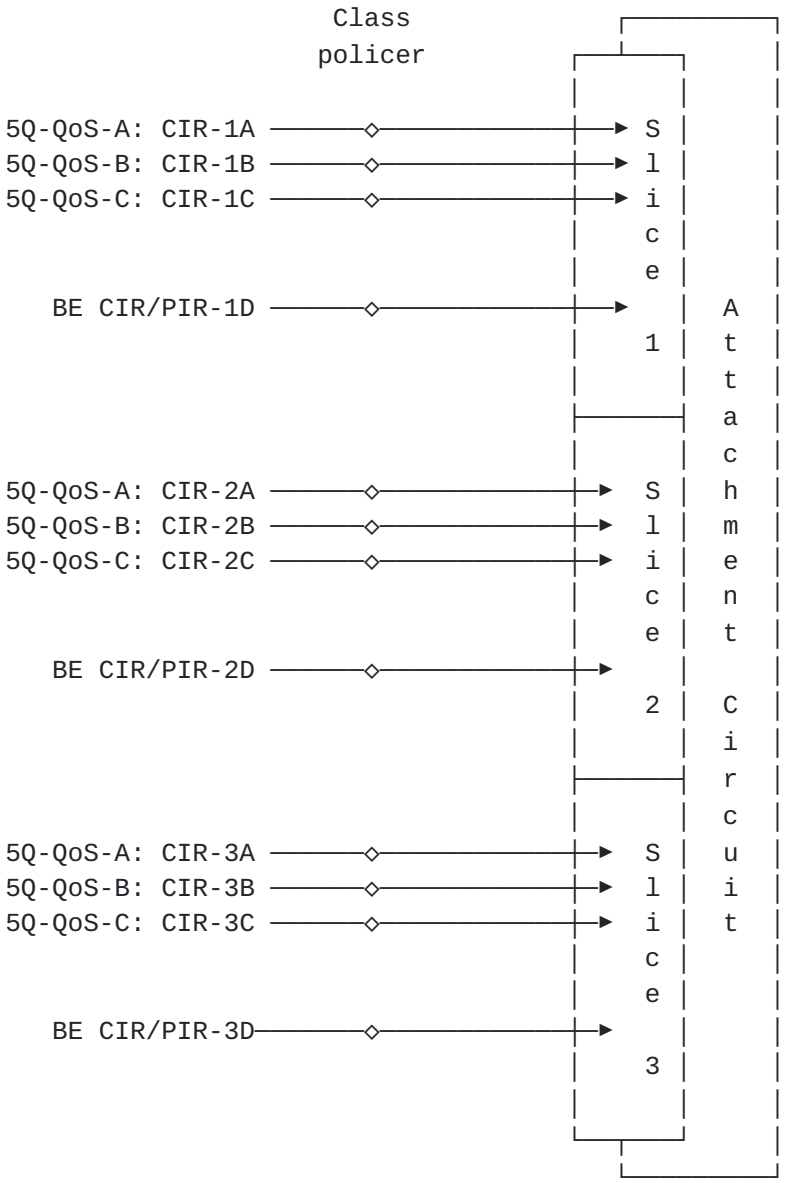


Figure 24: Ingress Slice Admission Control (5QI-aware Model)

The second model combines the advantages of 5QI-unaware model (per slice admission control) with the per traffic class admission control, as outlined in Figure 24. Ingress admission control is at class granularity for premium classes (CIR only). Non-premium class (i.e., Best Effort) has no separate class admission control policy, but it is allowed to use the entire slice capacity, which is available at any given moment. I.e., slice capacity, which is not consumed by premium classes. It is a hierarchical model, as depicted in Figure 25.

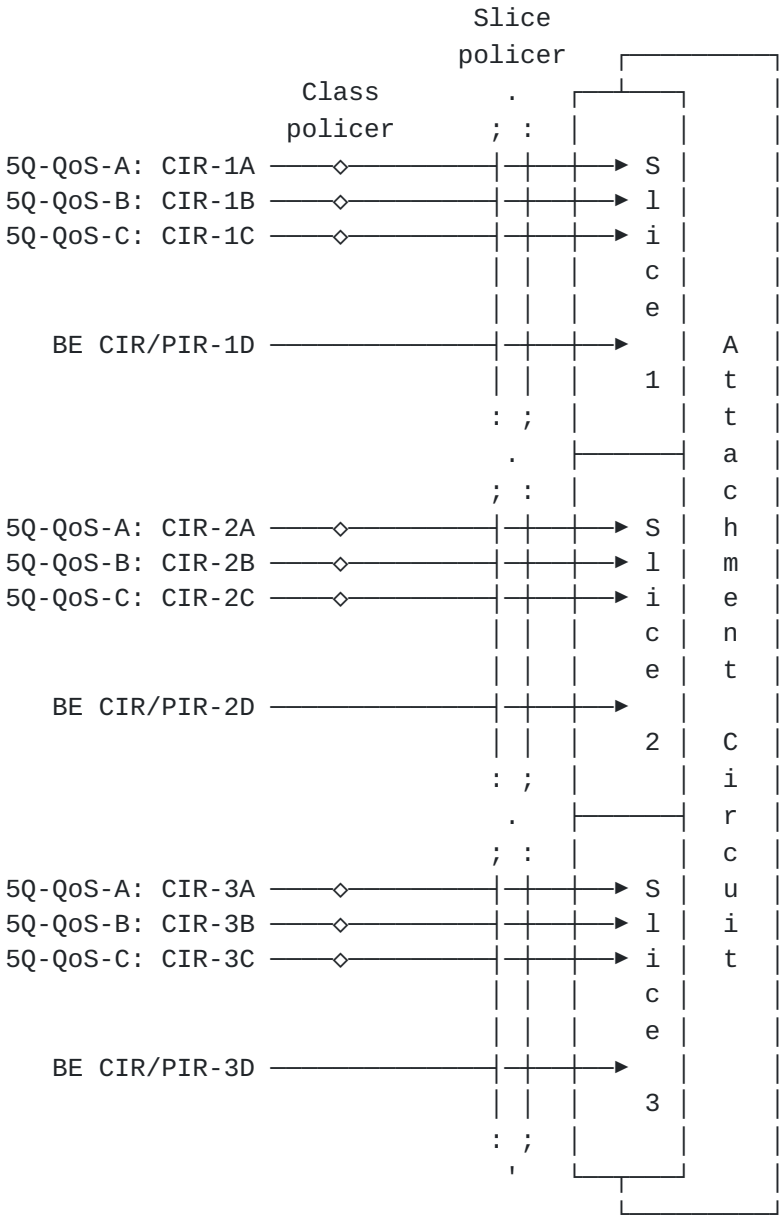


Figure 25: Ingress Slice Admission Control (5QI-aware) - Hierarchical

5.2.2.2. Outbound Edge Resource Control

Figure 26 outlines the outbound edge resource control model at the transport network layer for 5QI-aware slices. Each slice is assigned multiple egress queues. The sum of queue weights (equal to 5Q QoS CIRs within the slice) CIRs must not exceed the CIR of the slice itself. And, similarly to the 5QI-aware model, the sum of slice CIRs must not exceed the physical capacity of the attachment circuit.

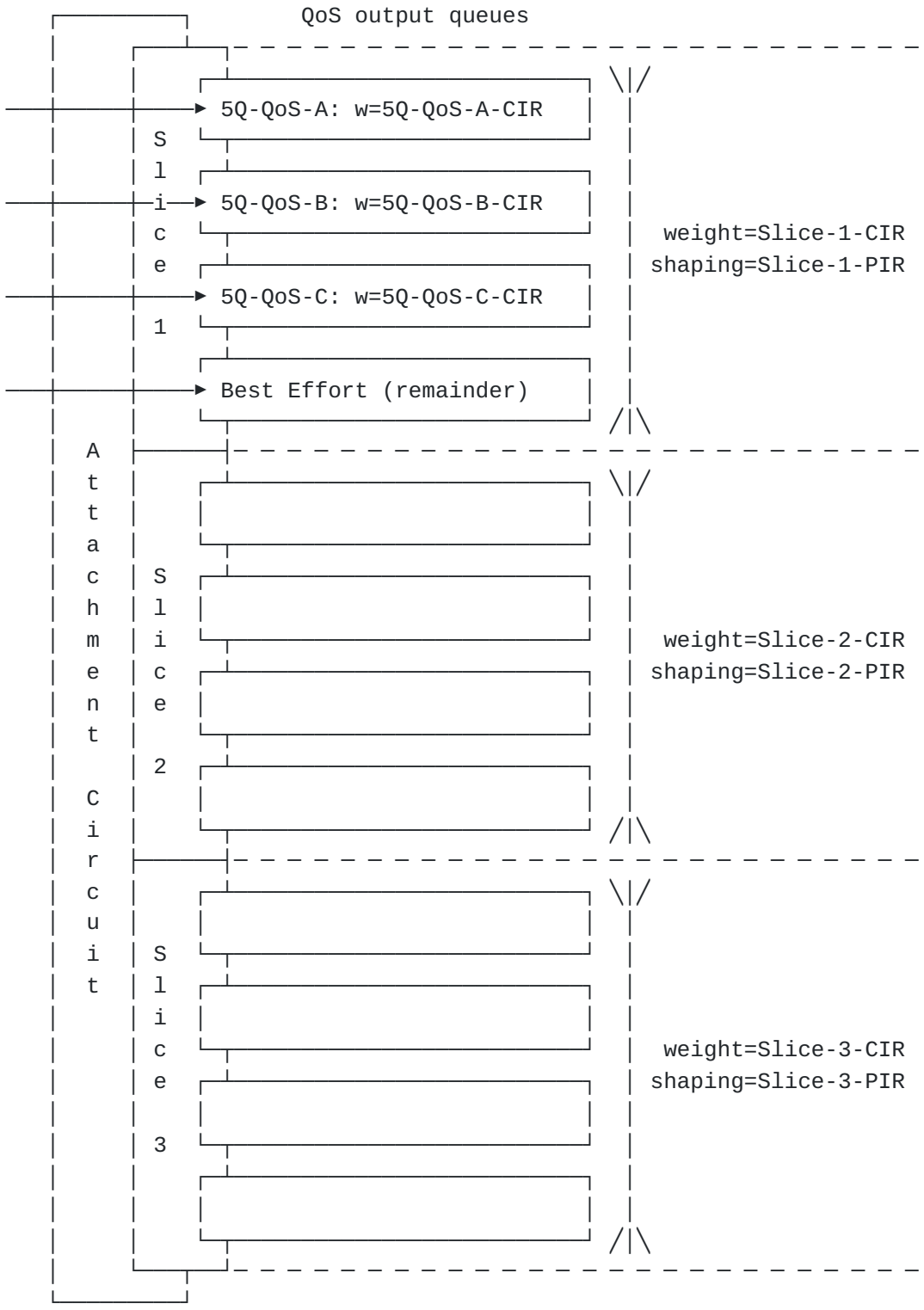


Figure 26: Egress Slice Admission Control (5QI-aware)

5.3. Transit Resource Control

Transit resource control is much simpler than Edge resource control in the provider network. As outlined in Figure 22, at the provider network edge, 5Q QoS Class marking (represented by DSCP related to

5QI set by mobile network functions in the packets handed off to the TN) is mapped to the TN QoS Class. Based in TN QoS Class, when the packet is encapsulated with outer header (MPLS or IPv6), TN QoS Class marking (MPLS TC or IPv6 DSCP in outer header, as depicted in Figure 18 and Figure 19) is set in the outer header. PHB in provider network transit routers is based exclusively on that TN QoS Class marking, i.e., original 5G QoS Class DSCP is not taken into consideration on transit.

Provider network transit resource control does not use any inbound interface policy, but only outbound interface policy, which is based on priority queue combined with weighted or deficit queuing model, without any shaper. The main purpose of transit resource control is to ensure that during network congestion events, for example caused by network failures and temporary rerouting, premium classes are prioritized, and any drops only occur in traffic that was de-prioritized by ingress admission control Section 5.2.1.1 or in non-premium (best-effort) classes. Capacity planning and management, as described in Section 7, ensures that enough capacity is available to fulfill all approved slice requests.

6. Transport Planes Mapping Models

A network operator might define various tunnel groups, where each tunnel group is created with specific optimization criteria and constraints. This document refers to such tunnel groups as 'transport planes'. For example, a transport plane "A" might represent tunnels optimized for latency, and transport plane "B" might represent tunnels optimized for high capacity.

Figure 27 depicts an example of a simple network with two transport planes. These transport planes might be realized via various IP/MPLS techniques, for example Flex-Algo or RSVP/SR traffic engineering tunnels with or without PCE, and with or without bandwidth reservations.

Section 7 discusses in detail different bandwidth models that can be deployed in the provider network. However, discussion about how to realize or orchestrate transport planes is out of scope for this document.

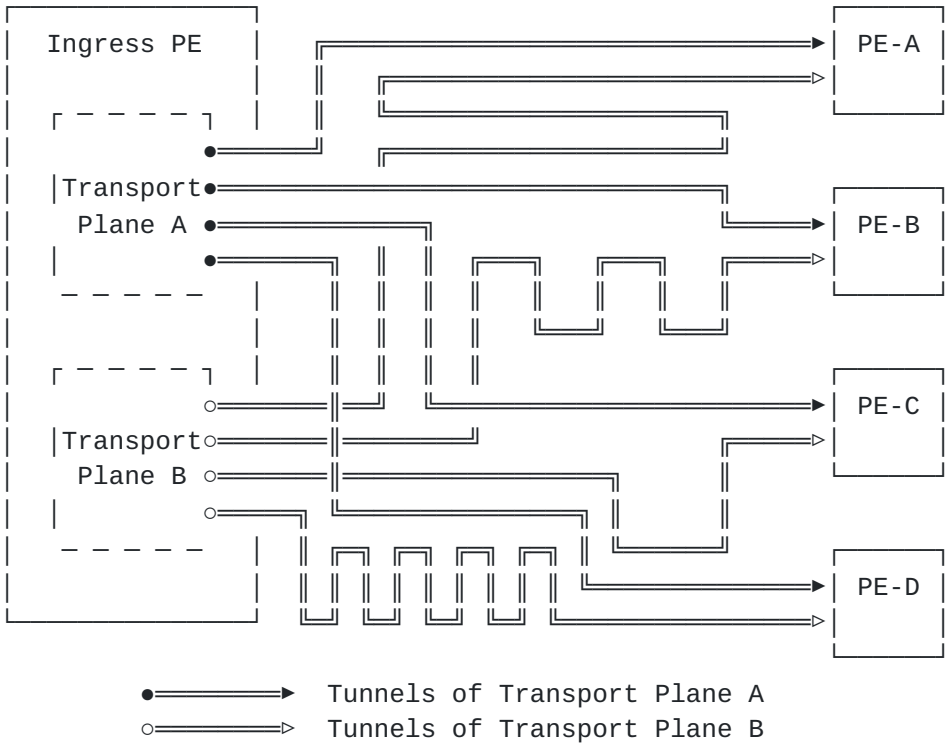


Figure 27: Transport Planes

Note that there could be multiple tunnels within a single transport plane between any pair of PEs. For readability, Figure 27 shows only single tunnel per transport plane for (ingress PE, egress PE) pair.

Similar to the QoS mapping models discussed in Section 5, for mapping to transport planes at the ingress PE, both 5QI-unaware and 5QI-aware models are defined. In essence, entire slices can be mapped to transport planes without 5G QoS consideration (5QI-unaware model), or flows with different 5G QoS Classes, even if they are from the same slice, might be mapped to different transport planes (5QI-aware model).

6.1. 5QI-unaware Model

As discussed in Section 5.2.1, in the 5QI-unware model, the provider network doesn't take into account 5G QoS during execution of per-hop behavior. The entire slice is mapped to single TN QoS Class, therefore the entire slice is subject to the same per-hop behavior. Similarly, in 5QI-unaware transport plane mapping model, the entire slice is mapped to a single transport plane, as depicted in Figure 28.

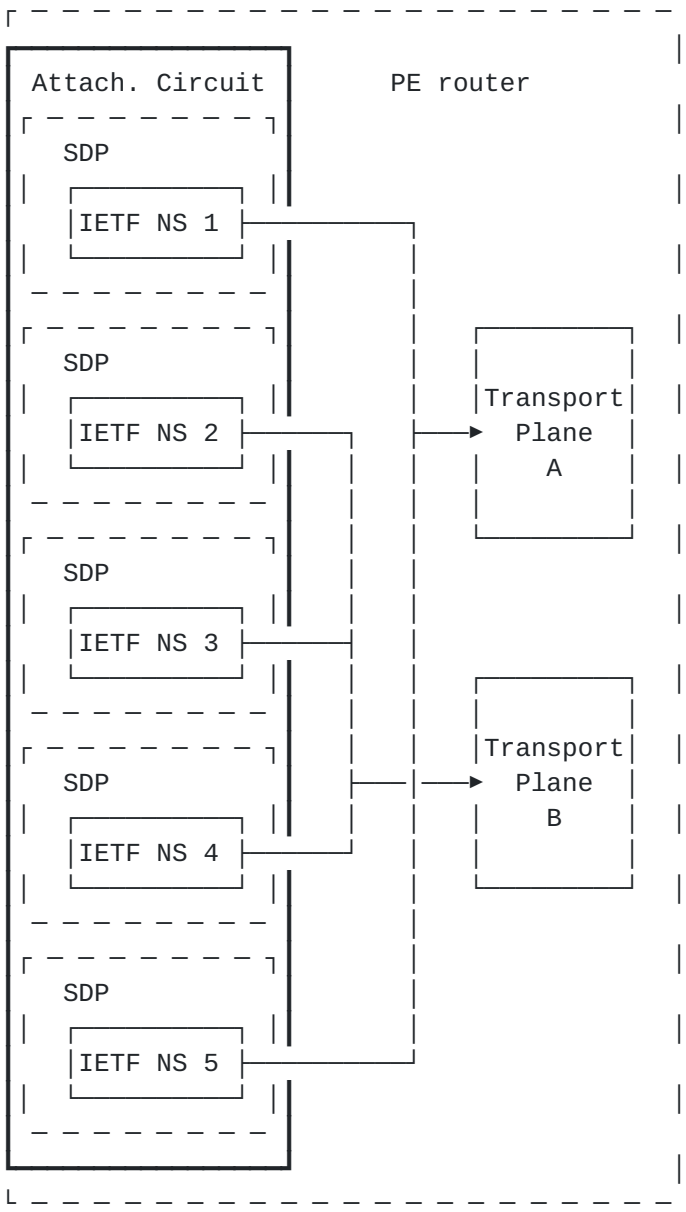


Figure 28: Slice to Transport Plane Mapping (5QI-unaware Model)

It is worth noting that there is no strict correlation between TN QoS Classes and Transport Planes. The TN domain can be operated with e.g., 8 TN QoS Classes (representing 8 hardware queues in the routers), and 2 Transport Planes (e.g., latency optimized transport plane using link latency metrics for path calculation, and transport plane following IGP metrics). TN QoS Class determines the per-hop behavior when the packets are transiting through the provider network, while Transport Plane determines the path, optimized or constrained based on operator's business criteria, that the packets use to transit through the provider network.

6.2. 5QI-aware Model

In 5QI-aware model, the traffic can be mapped to transport planes at the granularity of 5G QoS Class. Given that the potential number of transport planes is limited, packets from multiple 5G QoS Classes with similar characteristics are mapped to a common transport plane, as depicted in Figure 29.

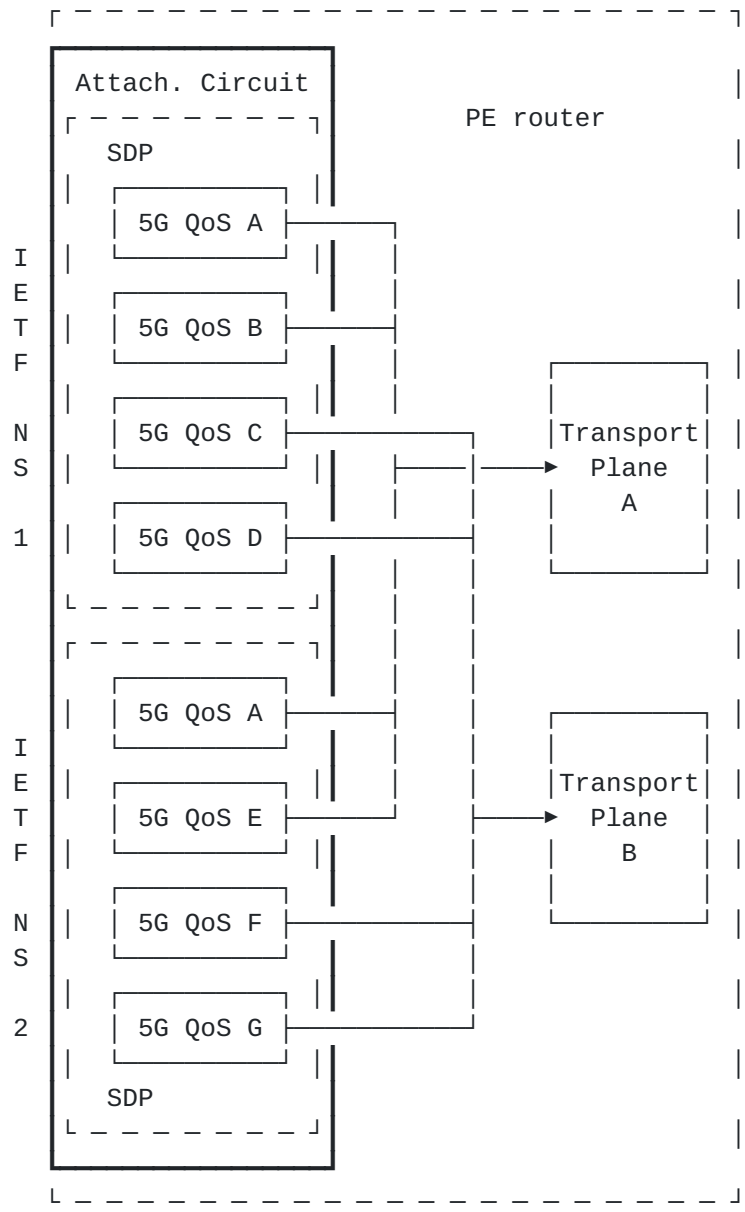


Figure 29: Slice to Transport Plane mapping (5QI-aware Model)

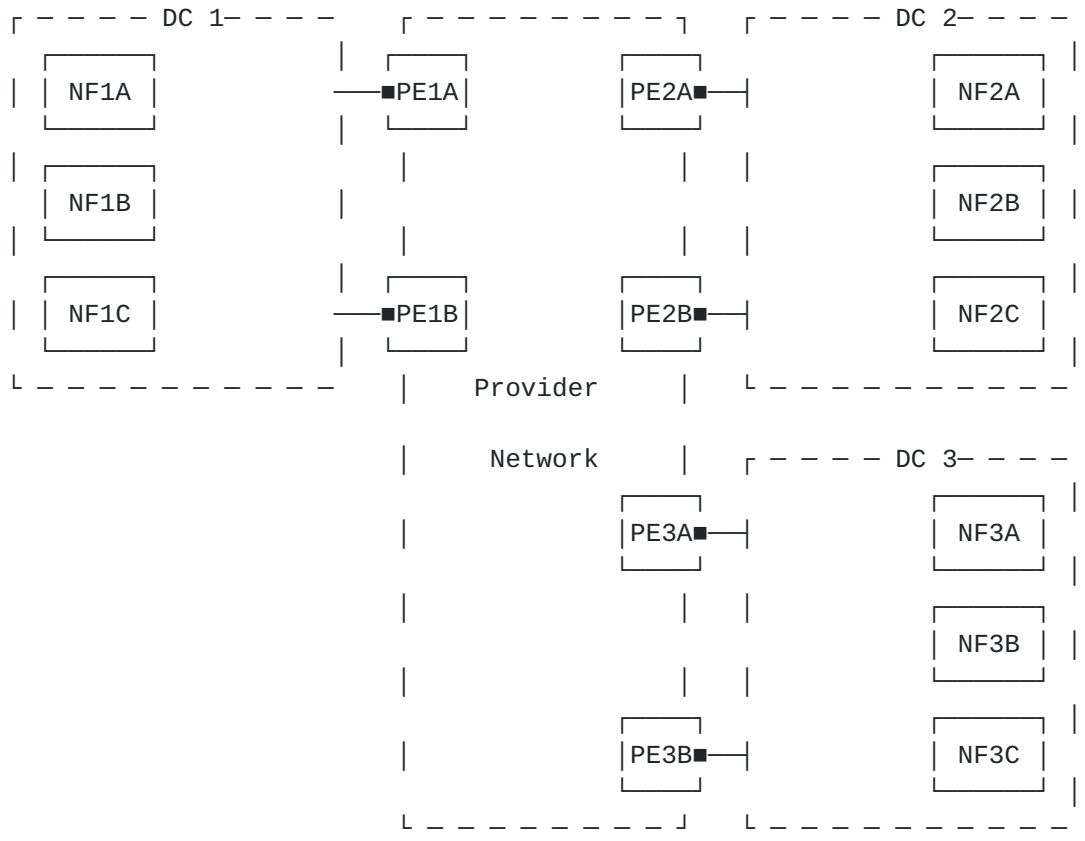
7. Capacity Planning/Management

7.1. Bandwidth Requirements

This section describes the information conveyed by the 5G NSO to the transport controller with respect to slice bandwidth requirements.

Figure 30 shows three DCs that contain instances of network functions. Also shown are PEs that have links to the DCs. The PEs belong to the provider network. Other details of the provider network, such as P-routers and transit links are not shown. Also details of the DC infrastructure in customer sites, such as switches and routers, are not shown.

The 5G NSO is aware of the existence of the network functions and their locations. However, it is not aware of the details of the provider network. The transport controller has the opposite view - it is aware of the provider network infrastructure and the links between the PEs and the DCs, but is not aware of the individual network functions at customer sites.



■ - SDP, with fine-grained QoS (dedicated resources per IETF NS)

Figure 30: An Example of Multi-DC Architecture

Let us consider 5G Slice "X" that uses some of the network functions in the three DCs. If this slice has latency requirements, the 5G NSO will have taken those into account when deciding which NF instances in which DC are to be invoked for this slice. As a result of such a placement decision, the three DCs shown are involved in 5G Slice "X", rather than other DCs. For its decision-making, the 5G NSO needs information from the NSC about the observed latency between DCs. Preferably, the NSC would present the topology in an abstracted form, consisting of point-to-point abstracted links between pairs of DCs and associated latency and, optionally, delay variation and link loss values. It would be valuable to have a mechanism for the 5G NSO to inform the NSC which DC-pairs are of interest for these metrics - there may be of order thousands of DCs, but the 5G NSO will only be interested in these metrics for a small fraction of all the possible DC-pairs, i.e. those in the same region of the provider network. The mechanism for conveying the information is out of scope for this document.

Figure 31 shows the matrix of bandwidth demands for 5G slice "X". Within the slice, multiple network function instances might be sending traffic from DCi to DCj. However, the 5G NSO sums the associated demands into one value. For example, NF1A and NF1B in DC1 might be sending traffic to multiple NFs in DC2, but this is expressed as one value in the traffic matrix: the total bandwidth required for 5G Slice X from DC1 to DC2 (8 units). Each row in the right-most column in the traffic matrix shows the total amount of traffic going from a given DC into the transport network, regardless of the destination DC. Note that this number can be less than the sum of DC-to-DC demands in the same row, on the basis that not all the network functions are likely to be sending at their maximum rate simultaneously. For example, the total traffic from DC1 for Slice X is 11 units, which is less than the sum of the DC-to-DC demands in the same row (13 units). Note, as described in Section 5, a slice may have per-QoS class bandwidth requirements, and may have CIR and PIR limits. This is not included in the example, but the same principles apply in such cases.

From \ To	DC 1	DC 2	DC 3	Total from DC
DC 1	n/a	8	5	11.0
DC 2	1	n/a	2	2.5
DC 3	4	7	n/a	10.0

Slice X

From \ To	DC 1	DC 2	DC 3	Total from DC
DC 1	n/a	4	2.5	6.0
DC 2	0.5	n/a	0.8	1.0
DC 3	2.6	3	n/a	5.1

Slice Y

Figure 31: Inter-DC Traffic Demand Matrix

[I-D.ietf-teas-ietf-network-slice-nbi-yang] can be used to convey all of the information in the traffic matrix to the IETF NSC. The IETF NSC applies policers corresponding to the last column in the traffic matrix to the appropriate PE routers, in order to enforce the bandwidth contract. For example, it applies a policer of 11 units to PE1A and PE1B that face DC1, as this is the total bandwidth that DC1 sends into the provider network corresponding to Slice X. Also, the controller may apply shapers in the direction from the TN to the DC, if otherwise there is the possibility of a link in the DC being oversubscribed. Note that a peer NF endpoint of an AC can be identified using 'peer-sap-id' as defined in [I-D.ietf-opsawg-sap].

Depending on the bandwidth model used in the provider network (Section 7.2), the other values in the matrix, i.e., the DC-to-DC demands, may not be directly applied to the provider network. Even so, the information may be useful to the IETF NSC for capacity planning and failure simulation purposes. If, on the other hand, the DC-to-DC demand information is not used by the IETF NSC, the IETF YANG Data Model for L3VPN Service Delivery [RFC8299] or the IETF YANG Data Model for L2VPN Service Delivery [RFC8466] could be used instead of [I-D.ietf-teas-ietf-network-slice-nbi-yang], as they support conveying the bandwidth information in the right-most column of the traffic matrix.

The provider network may be implemented in such a way that it has various types of paths, for example low-latency traffic might be mapped onto a different transport path to other traffic (for example a particular flex-algo or a particular set of TE LSPs), as discussed in Section 5. The 5G NSO can use [I-D.ietf-teas-ietf-network-slice-nbi-yang] to request low-latency transport for a given slice if required. However, [RFC8299] or [RFC8466] do not support requesting a particular transport-type, e.g., low-latency. One option is to augment these models to convey this information. This can be achieved by reusing the 'underlay-transport' construct defined in [RFC9182] and [RFC9291].

7.2. Bandwidth Models

This section describes three bandwidth management schemes that could be employed in the provider network. Many variations are possible, but each example describes the salient points of the corresponding scheme. Schemes 2 and 3 use TE; other variations on TE are possible as described in [I-D.ietf-teas-rfc3272bis].

7.2.1. Scheme 1: Shortest Path Forwarding (SPF)

Shortest path forwarding is used according to the IGP metric. Given that some slices are likely to have latency SLOs, the IGP metric on each link can be set to be in proportion to the latency of the link. In this way, all traffic follows the minimum latency path between endpoints.

In Scheme 1, although the operator provides bandwidth guarantees to the slice customers, there is no explicit end-to-end underpinning of the bandwidth SLO, in the form of bandwidth reservations across the provider network. Rather, the expected performance is achieved via capacity planning, based on traffic growth trends and anticipated future demands, in order to ensure that network links are not over-subscribed. This scheme is analogous to that used in many existing business VPN deployments, in that bandwidth guarantees are provided to the customers but are not explicitly underpinned end to end across the provider network.

A variation on the scheme is that Flex-Algo [I-D.ietf-lsr-flex-algo] is used. For example one Flex-Algo could use latency-based metrics and another Flex-Algo could use the IGP metric. There would be a many-to-one mapping of network slices to Flex- Algos.

While Scheme 1 is technically feasible, it is vulnerable to unexpected changes in traffic patterns and/or network element failures resulting in congestion. This is because, unlike Schemes 2 and 3 that employ TE, traffic cannot be diverted from the shortest path.

7.2.2. Scheme 2: TE LSPs with Fixed Bandwidth Reservations

Scheme 2 uses RSVP-TE or SR-TE LSPs with fixed bandwidth reservations. By "fixed", we mean a value that stays constant over time, unless the 5G NSO communicates a change in slice bandwidth requirements, due to the creation or modification of a slice. Note that the "reservations" would be in the mind of the transport controller - it is not necessary (or indeed possible for SR-TE) to reserve bandwidth at the network layer. The bandwidth requirement acts as a constraint whenever the controller (re)computes the path of an LSP. There could be a single mesh of LSPs between endpoints that carry all of the traffic types, or there could be a small handful of meshes, for example one mesh for low-latency traffic that follows the minimum latency path and another mesh for the other traffic that follows the minimum IGP metric path, as described in Section 5. There would be a many-to-one mapping of slices to LSPs.

The bandwidth requirement from DCi to DCj is the sum of the DCi-DCj demands of the individual slices. For example, if only Slice X and Slice Y are present, then the bandwidth requirement from DC1 to DC2 is 12 units (8 units for Slice X and 4 units for Slice Y). When the 5G NSO requests a new slice, the transport controller, in its mind, increments the bandwidth requirement according to the requirements of the new slice. For example, in Figure 30, suppose a new slice is instantiated that needs 0.8 Gbps from DC1 to DC2. The transport controller would increase its notion of the bandwidth requirement from DC1 to DC2 from 12 Gbps to 12.8 Gbps to accommodate the additional expected traffic.

In the example, each DC has two PEs facing it for reasons of resilience. The transport controller needs to determine how to map the DC1 to DC2 bandwidth requirement to bandwidth reservations of TE LSPs from DC1 to DC2. For example, if the routing configuration is arranged such that in the absence of any network failure, traffic from DC1 to DC2 always enters PE1A and goes to PE2A, the controller reserves 12.8 Gbps of bandwidth on the LSP from PE1A to PE2A. If, on the other hand, the routing configuration is arranged such that in the absence of any network failure, traffic from DC1 to DC2 always enters PE1A and is load-balanced across PE2A and PE2B, the controller reserves 6.4 Gbps of bandwidth on the LSP from PE1A to PE2A and 6.4 Gbps of bandwidth on the LSP from PE1A to PE2B. It might be tricky for the transport controller to be aware of all conditions that change the way traffic lands on the various PEs, and therefore know that it needs to change bandwidth reservations of LSPs accordingly. For example, there might be an internal failure within DC1 that causes traffic from DC1 to land on PE1B, rather than PE1A. The transport controller may not be aware of the failure and therefore may not know that it now needs to apply bandwidth reservations to LSPs from PE1B to PE2A/PE2B.

7.2.3. Scheme 3: TE LSPs without Bandwidth Reservation

Like Scheme 2, Scheme 3 uses RSVP-TE or SR-TE LSPs. There could be a single mesh of LSPs between endpoints that carry all of the traffic types, or there could be a small handful of meshes, for example one mesh for low-latency traffic that follows the minimum latency path and another mesh for the other traffic that follows the minimum IGP metric path, as described in Section 5. There would be a many-to-one mapping of slices to LSPs.

The difference between Scheme 2 and Scheme 3 is that Scheme 3 does not have fixed bandwidth reservations for the LSPs. Instead, actual measured data-plane traffic volumes are used to influence the placement of TE LSPs. One way of achieving this is to use distributed RSVP-TE with auto-bandwidth. Alternatively, the transport controller can use telemetry-driven automatic congestion avoidance. In this approach, when the actual traffic volume in the data plane on given link exceeds a threshold, the controller, knowing how much actual data plane traffic is currently travelling along each RSVP or SR-TE LSP, can tune the paths of one or more LSPs using the link such that they avoid that link.

It would be undesirable to move a minimum-latency LSP rather than another type of LSP in order to ease the congestion, as the new path will typically have a higher latency, if the minimum-latency LSP is currently following the minimum-latency path. This can be avoided by designing the algorithms described in the previous paragraph such that they avoid moving minimum-latency LSPs unless there is no alternative.

8. Network Slicing OAM

The deployment and maintenance of slices within a network imply a set OAM functions ([RFC6291]) to be deployed by the providers, e.g.:

* Providers should be able to execute OAM tasks on a per network slice basis. These tasks can cover the "full" slice within a domain or a portion of that slice (for troubleshooting purposes, for example).

  For example, per-slice OAM tasks can consist in (but not limited to):

  - tracing resources that are bound to a given network 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 network slice service requirements, or

-assessing the compliance of the allocated network slice resources against flow/ customer service requirements.

[RFC7276] provides an overview of 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 or specific feature that would address their needs.

SFC OAM [I-D.ietf-sfc-oam-packet] should also be supported for slices that make uses of service function chaining [RFC7665]. An example of SFC OAM technique to Continuity Check, Connectivity Verification, or tracing service functions is specified in [I-D.ietf-sfc-multi-layer-oam].

*Providers may want to enable differentiated failure detect and repair features for a subset of network slices. For example, a given network slice may require fast detect and repair mechanisms, while others may not be engineered with such means. The provider can use techniques such as [RFC5286], [RFC5714], or [RFC8355].

*Providers may deploy means to dynamically discover the set of network slices that are enabled within its network. Such dynamic discovery capability facilitates the detection of any mismatch between the view maintained by the control/management plane and the actual network configuration. When mismatches are detected, corrective actions must be undertaken accordingly. For example, a provider may rely upon L3NM [RFC9182] or L2NM [RFC9291] to maintain the full set of L3VPN/L2VPNs that are used to deliver network slice services. The correlation between an LxVPN instance and a network slice service is maintained using "parent-service-id" attribute (Section 7.3 of [RFC9182].

*Means to report a set of network performance metrics to assess whether the agreed slice service objectives are honored. These means are used for SLO monitoring and violation detect purposes. For example, [RFC9375] can be used to report links' one-way delay, one-way delay variation, etc. Both conventional active/ passive measurement methods [RFC7799] and more recent telemetry methods (e.g. YANG Push [RFC8641]) can be used.

*Means to report and expose observed performance metrics and other OAM state to customer. For example, [I-D.ietf-teas-ietf-network-slice-nbi-yang] exposes a set of statistics per SDP, connectivity construct, and connection group.

9. IANA Considerations

This document does not make any IANA request.

10. Security Considerations

IETF Network Slices considerations are discussed in Section 6 of [I-D.ietf-teas-ietf-network-slices].

Many of the YANG modules cited in this document define schema for data that is designed to be accessed via network management protocols such as NETCONF [RFC6241] or RESTCONF [RFC8040]. The lowest NETCONF layer is the secure transport layer, and the mandatory-to-implement secure transport is Secure Shell (SSH) [RFC6242]. The lowest RESTCONF layer is HTTPS, and the mandatory-to-implement secure transport is TLS [RFC8446].

The NETCONF access control model [RFC8341] provides the means to restrict access for particular NETCONF or RESTCONF users to a preconfigured subset of all available NETCONF or RESTCONF protocol operations and content.

Security considerations specific to each of the technologies and protocols listed in the document are discussed in the specification documents of each of these protocols.

Adequate admission control policies should be configured in the edge of the provider network to control access to specific slice resources. Likewise, access to classification and mapping tables must be controlled to prevent misbehaviors (an unauthorized entity may modify the table to bind traffic to a random slice, redirect the traffic, etc.). Network devices must check that a required access privilege is provided before granting access to specific data or performing specific actions.

11. References

11.1. Normative References

[I-D.ietf-teas-ietf-network-slices]
Farrel, A., Drake, J., Rokui, R., Homma, S., Makhijani, K., Contreras, L. M., and J. Tantsura, "A Framework for IETF Network Slices", Work in Progress, Internet-Draft, draft-ietf-teas-ietf-network-slices-19, 21 January 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-ietf-network-slices-19>.

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

[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/rfc/rfc6241>.

[RFC6242] Wasserman, M., "Using the NETCONF Protocol over Secure Shell (SSH)", RFC 6242, DOI 10.17487/RFC6242, June 2011, <https://www.rfc-editor.org/rfc/rfc6242>.

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

[RFC8341] Bierman, A. and M. Bjorklund, "Network Configuration Access Control Model", STD 91, RFC 8341, DOI 10.17487/RFC8341, March 2018, <https://www.rfc-editor.org/rfc/rfc8341>.

[RFC8446] Rescorla, E., "The Transport Layer Security (TLS) Protocol Version 1.3", RFC 8446, DOI 10.17487/RFC8446, August 2018, <https://www.rfc-editor.org/rfc/rfc8446>.

11.2. Informative References

[I-D.boro-opsawg-teas-attachment-circuit] Boucadair, M., Roberts, R., de Dios, O. G., Barguil, S., and B. Wu, "YANG Data Models for 'Attachment Circuits'-as-a-Service (ACaaS)", Work in Progress, Internet-Draft, draft-boro-opsawg-teas-attachment-circuit-06, 3 May 2023, <https://datatracker.ietf.org/doc/html/draft-boro-opsawg-teas-attachment-circuit-06>.

[I-D.henry-tsvwg-diffserv-to-qci] Henry, J., Szigeti, T., and L. M. Contreras, "Diffserv to QCI Mapping", Work in Progress, Internet-Draft, draft-henry-tsvwg-diffserv-to-qci-04, 13 April 2020, <https://datatracker.ietf.org/doc/html/draft-henry-tsvwg-diffserv-to-qci-04>.

[I-D.ietf-lsr-flex-algo] Psenak, P., Hegde, S., Filsfils, C., Talaulikar, K., and A. Gulko, "IGP Flexible Algorithm", Work in Progress, Internet-Draft, draft-ietf-lsr-flex-algo-26, 17 October 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-lsr-flex-algo-26>.

[I-D.ietf-opsawg-sap] Boucadair, M., de Dios, O. G., Barguil, S., Wu, Q., and V. Lopez, "A YANG Network Model for Service Attachment Points (SAPs)", Work in Progress, Internet-Draft, draft-ietf-opsawg-sap-15, 18 January 2023,

<https://datatracker.ietf.org/doc/html/draft-ietf-opsawg-sap-15>.

[I-D.ietf-sfc-multi-layer-oam] Mirsky, G., Meng, W., Ao, T., Khasnabish, B., Leung, K., and G. S. Mishra, "Active OAM for Service Function Chaining (SFC)", Work in Progress, Internet-Draft, draft-ietf-sfc-multi-layer-oam-23, 26 March 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-sfc-multi-layer-oam-23>.

[I-D.ietf-sfc-oam-packet]
Boucadair, M., "OAM Packet and Behavior in the Network Service Header (NSH)", Work in Progress, Internet-Draft, draft-ietf-sfc-oam-packet-03, 26 March 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-sfc-oam-packet-03>.

[I-D.ietf-teas-5g-network-slice-application]
Geng, X., Contreras, L. M., Rokui, R., Dong, J., and I. Bykov, "IETF Network Slice Application in 3GPP 5G End-to-End Network Slice", Work in Progress, Internet-Draft, draft-ietf-teas-5g-network-slice-application-00, 4 May 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-5g-network-slice-application-00>.

[I-D.ietf-teas-ietf-network-slice-nbi-yang]
Wu, B., Dhody, D., Rokui, R., Saad, T., Han, L., and J. Mullooly, "A YANG Data Model for the IETF Network Slice Service", Work in Progress, Internet-Draft, draft-ietf-teas-ietf-network-slice-nbi-yang-04, 13 March 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-ietf-network-slice-nbi-yang-04>.

[I-D.ietf-teas-rfc3272bis]
Farrel, A., "Overview and Principles of Internet Traffic Engineering", Work in Progress, Internet-Draft, draft-ietf-teas-rfc3272bis-22, 27 October 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-rfc3272bis-22>.

[NG.113] GSMA, "NG.113: 5GS Roaming Guidelines Version 4.0", May 2021, <https://www.gsma.com/newsroom/wp-content/uploads//NG.113-v4.0.pdf>.

[O-RAN.WG9.XPSAAS] O-RAN Alliance, "O-RAN.WG9.XPSAAS: O-RAN WG9 Xhaul Packet Switched Architectures and Solutions Version

03.00", February 2022, <https://www.o-ran.org/specifications>.

[RFC2698] Heinanen, J. and R. Guerin, "A Two Rate Three Color Marker", RFC 2698, DOI 10.17487/RFC2698, September 1999, <https://www.rfc-editor.org/rfc/rfc2698>.

[RFC4115] Aboul-Magd, O. and S. Rabie, "A Differentiated Service Two-Rate, Three-Color Marker with Efficient Handling of in-Profile Traffic", RFC 4115, DOI 10.17487/RFC4115, July 2005, <https://www.rfc-editor.org/rfc/rfc4115>.

[RFC4664] Andersson, L., Ed. and E. Rosen, Ed., "Framework for Layer 2 Virtual Private Networks (L2VPNs)", RFC 4664, DOI 10.17487/RFC4664, September 2006, <https://www.rfc-editor.org/rfc/rfc4664>.

[RFC5286] Atlas, A., Ed. and A. Zinin, Ed., "Basic Specification for IP Fast Reroute: Loop-Free Alternates", RFC 5286, DOI 10.17487/RFC5286, September 2008, <https://www.rfc-editor.org/rfc/rfc5286>.

[RFC5714] Shand, M. and S. Bryant, "IP Fast Reroute Framework", RFC 5714, DOI 10.17487/RFC5714, January 2010, <https://www.rfc-editor.org/rfc/rfc5714>.

[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, <https://www.rfc-editor.org/rfc/rfc6291>.

[RFC6459] Korhonen, J., Ed., Soininen, J., Patil, B., Savolainen, T., Bajko, G., and K. Iisakkila, "IPv6 in 3rd Generation Partnership Project (3GPP) Evolved Packet System (EPS)", RFC 6459, DOI 10.17487/RFC6459, January 2012, <https://www.rfc-editor.org/rfc/rfc6459>.

[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, <https://www.rfc-editor.org/rfc/rfc7276>.

[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/rfc/rfc7665>.

[RFC7799] Morton, A., "Active and Passive Metrics and Methods (with Hybrid Types In-Between)", RFC 7799, DOI 10.17487/RFC7799, May 2016, <https://www.rfc-editor.org/rfc/rfc7799>.

[RFC7806] Baker, F. and R. Pan, "On Queuing, Marking, and Dropping", RFC 7806, DOI 10.17487/RFC7806, April 2016, <https://www.rfc-editor.org/rfc/rfc7806>.

[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/rfc/rfc8299>.

[RFC8355] Filsfils, C., Ed., Previdi, S., Ed., Decraene, B., and R. Shakir, "Resiliency Use Cases in Source Packet Routing in Networking (SPRING) Networks", RFC 8355, DOI 10.17487/RFC8355, March 2018, <https://www.rfc-editor.org/rfc/rfc8355>.

[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/rfc/rfc8466>.

[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/rfc/rfc8641>.

[RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <https://www.rfc-editor.org/rfc/rfc8754>.

[RFC8969] Wu, Q., Ed., Boucadair, M., Ed., Lopez, D., Xie, C., and L. Geng, "A Framework for Automating Service and Network Management with YANG", RFC 8969, DOI 10.17487/RFC8969, January 2021, <https://www.rfc-editor.org/rfc/rfc8969>.

[RFC9182] Barguil, S., Gonzalez de Dios, O., Ed., Boucadair, M., Ed., Munoz, L., and A. Aguado, "A YANG Network Data Model

for Layer 3 VPNs", RFC 9182, DOI 10.17487/RFC9182, February 2022, <https://www.rfc-editor.org/rfc/rfc9182>.

[RFC9291] Boucadair, M., Ed., Gonzalez de Dios, O., Ed., Barguil, S., and L. Munoz, "A YANG Network Data Model for Layer 2 VPNs", RFC 9291, DOI 10.17487/RFC9291, September 2022, <https://www.rfc-editor.org/rfc/rfc9291>.

[RFC9375] Wu, B., Ed., Wu, Q., Ed., Boucadair, M., Ed., Gonzalez de Dios, O., and B. Wen, "A YANG Data Model for Network and VPN Service Performance Monitoring", RFC 9375, DOI 10.17487/RFC9375, April 2023, <https://www.rfc-editor.org/rfc/rfc9375>.

[TR-GSTR-TN5G] ITU-T, "Technical Report GSTR-TN5G", February 2018, <https://www.itu.int/dms_pub/itu-t/opb/tut/T-TUT-HOME-2018-PDF-E.pdf>.

[TS-23.501] 3GPP, "TS 23.501: System architecture for the 5G System (5GS)", 2021, <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144>.

[TS-28.530] 3GPP, "TS 23.530: Management and orchestration; Concepts, use cases and requirements)", 2023, <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3273>.

[_5G-Book] Peterson, L., Sunay, O., and B. Davie, "5G Mobile Networks: A Systems Approach", 2022, <https://5g.systemsapproach.org/>.

Appendix A. Open Issues

The following issues should be resolved prior to the WGLC:

1. Assess which/whether some the material in the "5G Slice to IETF Network Slice Mapping" Section should be maintained in this draft or moved to [I-D.ietf-teas-5g-network-slice-application] (Adrian)

   *This issue is tracked at https://github.com/boucadair/5g-slice-realization/issues/40.

2. Assess whether we need to mainatin the "First 5G Slice vs Subsequent Slices" Section:

   *Unless we explain how this ss important for realization, this section should be deleted (Med)

*The motivation of this section is not clear (from Reza)

*Need to describe the implications to the realization of IETF network slices (Jie)

*The issue is tracked at https://github.com/boucadair/5g-slice-realization/issues/19

3. Clarify the use of inter-AS option B/C to model the AC between CE and PE (Jie)

*The issue is tracked at https://github.com/boucadair/5g-slice-realization/issues/52

4. Further discuss whether the TN slice in the customer site is covered or is out of the scope of IETF network slice (Jie)

*The issue is tracked at https://github.com/boucadair/5g-slice-realization/issues/53

Active issues can be tracked at: https://github.com/boucadair/5g-slice-realization/issues

Appendix B. Acronyms and Abbreviations

3GPP: 3rd Generation Partnership Project

5GC: 5G Core

5QI: 5G QoS Indicator

A2A: Any-to-Any

AC: Attachment Circuit

AMF: Access and Mobility Management Function

AUSF: Authentication Server Function

BBU: Baseband Unit

BH: Backhaul

BS: Base Station

CE: Customer Edge

CIR: Committed Information Rate

CN: Core Network

CoS: Class of Service

CP: Control Plane

CSP: Communication Service Provider

CU: Centralized Unit

CU-CP: Centralized Unit Control Plane

CU-UP: Centralized Unit User Plane

DC: Data Center

DDoS: Distributed Denial of Services

DN: Data Network

DSCP: Differentiated Services Code Point

DU: Distributed Unit

eCPRI: enhanced Common Public Radio Interface

FH: Fronthaul

FIB: Forwarding Information Base

GPRS: Generic Packet Radio Service

gNB: gNodeB

GTP: GPRS Tunneling Protocol

GTP-U: GPRS Tunneling Protocol User plane

HW: Hardware

ID: Identifier

IGP: Interior Gateway Protocol

IP: Internet Protocol

L2VPN: Layer 2 Virtual Private Network

L3VPN: Layer 3 Virtual Private Network

LSP: Label Switched Path

MH: Midhaul

MIoT: Massive Internet of Things

MPLS: Multiprotocol Label Switching

NF: Network Function

NR: New Radio

NRF: Network Function Repository

NRP: Network Resource Partition

NSC: Network Slice Controller

PE: Provider Edge

PIR: Peak Information Rate

PLMN: Public Land Mobile Network

PSTN: Public Switched Telephony Network

QoS: Quality of Service

RAN: Radio Access Network

RF: Radio Frequency

RIB: Routing Information Base

RSVP: Resource Reservation Protocol

RU: Radio Unit

SD: Slice Differentiator

SDP: Service Demarcation Point

SLA: Service Level Agreement

SLO: Service Level Objective

SMF: Session Management Function

S-NSSAI: Single Network Slice Selection Assistance Information

SST: Slice/Service Type

SR: Segment Routing

SRv6: Segment Routing version 6

TC: Traffic Class

TE: Traffic Engineering

TN: Transport Network

TS: Technical Specification

UDM: Unified Data Management

UE: User Equipment

UP: User Plane

UPF: User Plane Function

URLLC: Ultra Reliable Low Latency Communication

VLAN: Virtual Local Area Network

VNF: Virtual Network Function

VPN: Virtual Private Network

VRF: Virtual Routing and Forwarding

VXLAN: Virtual Extensible Local Area Network

Appendix C. An Overview of 5G Networking

This section provides a brief introduction to 5G mobile networking with a perspective on the Transport Network. This section does not intend to replace or define 3GPP architecture, instead its objective is to provide an overview for readers that do not have a mobile background. For more comprehensive information, refer to [TS-23.501].

C.1. Key Building Blocks

[TS-23.501] defines the Network Functions (UPF, AMF, etc.) that compose the 5G System (5GS) Architecture together with related interfaces (e.g., N1, N2). This architecture has native Control and User Plane separation, and the Control Plane leverages a service-based architecture. Figure 32 outlines an example 5GS architecture with a subset of possible network functions and network interfaces.

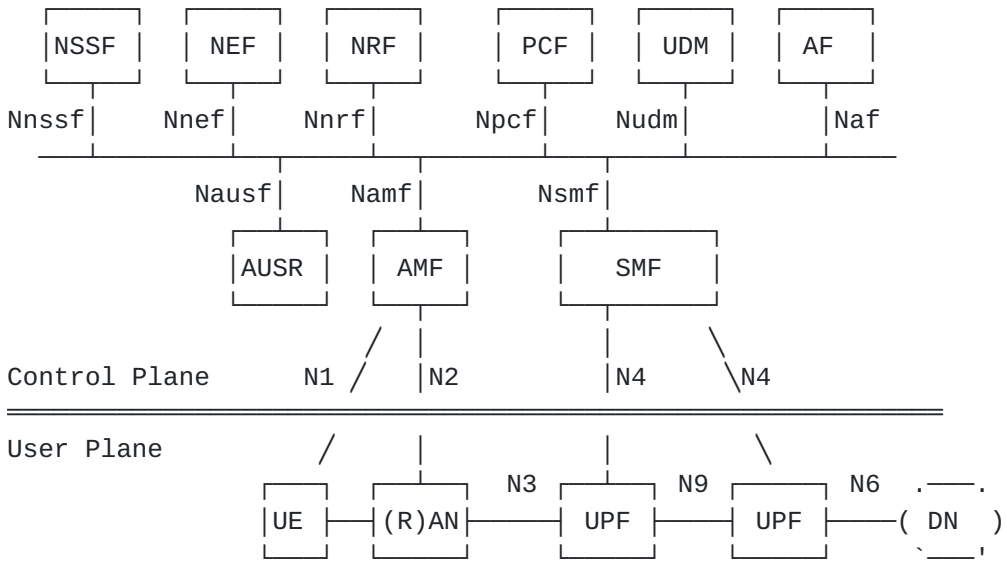


Figure 32: 5GS Architecture and Service-based Interfaces

Similar to previous versions of 3GPP mobile networks [RFC6459], a 5G mobile network is split into the following four major domains (Figure 33):

*UE, MS, MN, and Mobile:

The terms UE (User Equipment), MS (Mobile Station), MN (Mobile Node), and mobile refer to the devices that are hosts with the ability to obtain Internet connectivity via a 3GPP network. An MS is comprised of the Terminal Equipment (TE) and a Mobile Terminal (MT). The terms UE, MS, MN, and mobile are used interchangeably within this document.

*Radio Access Network (RAN):

Provides wireless connectivity to the UE devices via radio. It is made up of the Antenna that transmits and receives signals to the UE and the Base Station that digitizes the signal and converts the RF data stream to IP packets.

*Core Network (CN):

Controls the CP of the RAN and provides connectivity to the Data Network (e.g., the Internet or a private VPN). The Core Network hosts dozens of services such as authentication, phone registry, charging, access to PSTN and handover.

*Transport Network (TN):

Provides connectivity between 5G Network Functions. The TN may provide connectivity from the RAN to the Core Network, as well as

within the RAN or within the CN. The traffic generated by NFs is - mostly - based on IP or Ethernet.

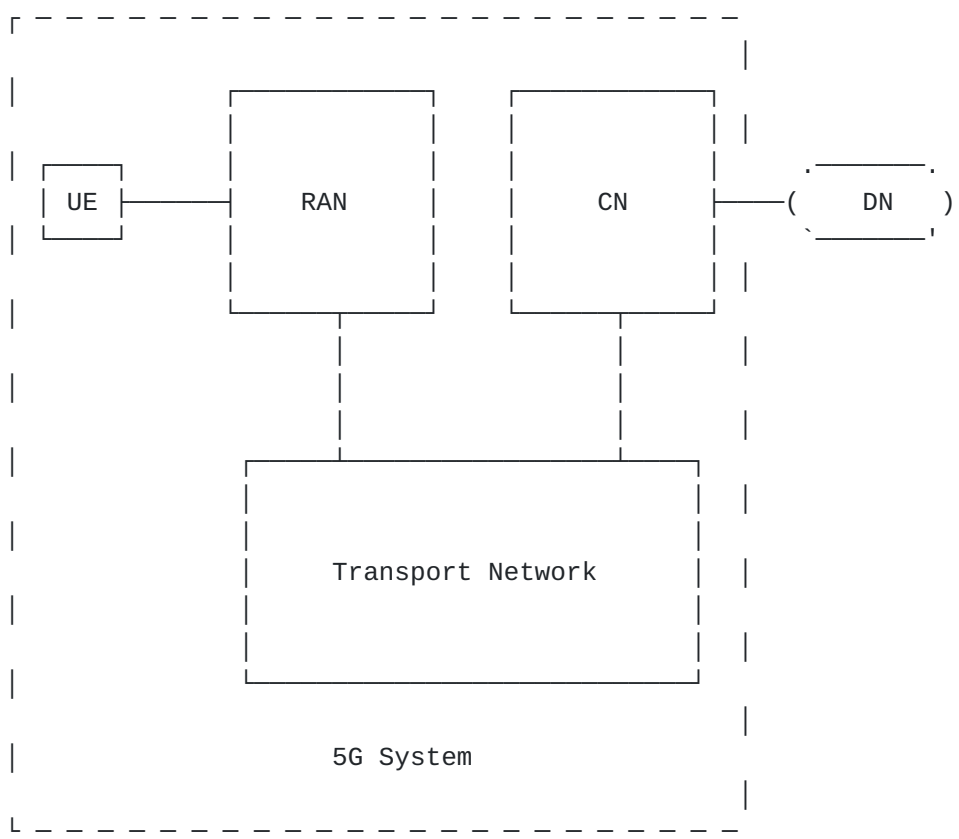


Figure 33: Building Blocks of 5G Architecture (A High-Level Representation)

C.2. Core Network (CN)

The 5G Core Network (5GC) is made up of a set of NFs which fall into two main categories (Figure 34):

*5GC User Plane:

The User Plane Function (UPF) is the interconnect point between the mobile infrastructure and the Data Network (DN). It interfaces with the RAN via the N3 interface by encapsulating/ decapsulating the User Plane Traffic in GTP Tunnels (aka GTP-U or Mobile User Plane).

*5GC Control Plane:

The 5G Control Plane is made up of a comprehensive set of Network Functions. An exhaustive list and description of these entities is out of the scope of this document. The following NFs and

interfaces are worth mentioning, since their connectivity may rely on the Transport Network:

-the AMF (Access and Mobility Function) connects with the RAN control plane over the N2 interface

-the SMF controls the 5GC UPF via the N4 interface

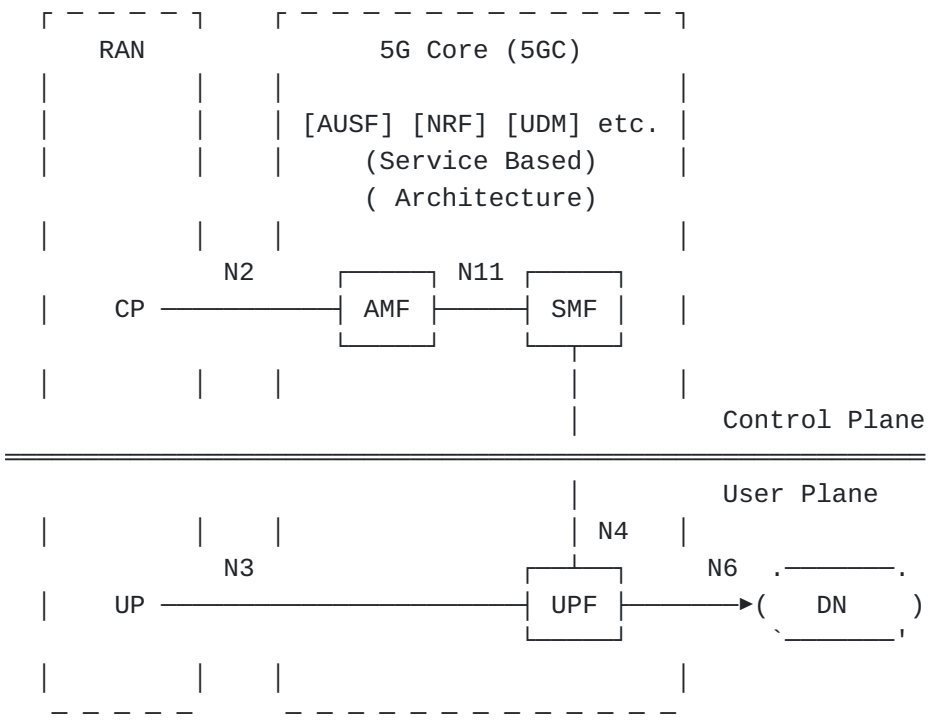


Figure 34: 5G Core Network (CN)

C.3. Radio Access Network (RAN)

The RAN connects cellular wireless devices to a mobile Core Network. The RAN is made up of three components, which form the Radio Base Station:

*The Baseband Unit (BBU) provides the interface between the Core Network and the Radio Network. It connects to the Radio Unit and is responsible for the baseband signal processing to packet.

*The Radio Unit (RU) is located close to the Antenna and controlled by the BBU. It converts the Baseband signal received from the BBU to a Radio frequency signal.

*The Antenna converts the electric signal received from the RU to radio waves

The 5G RAN Base Station is called a gNodeB (gNB). It connects to the Core Network via the N3 (User Plane) and N2 (Control Plane) interfaces.

The 5G RAN architecture supports RAN disaggregation in various ways. Notably, the BBU can be split into a DU (Distributed Unit) for digital signal processing and a CU (Centralized Unit) for RAN Layer 3 processing. Furthermore, the CU can be itself split into Control Plane (CU-CP) and User Plane (CU-UP).

Figure 35 depicts a disaggregated RAN with NFs and interfaces.

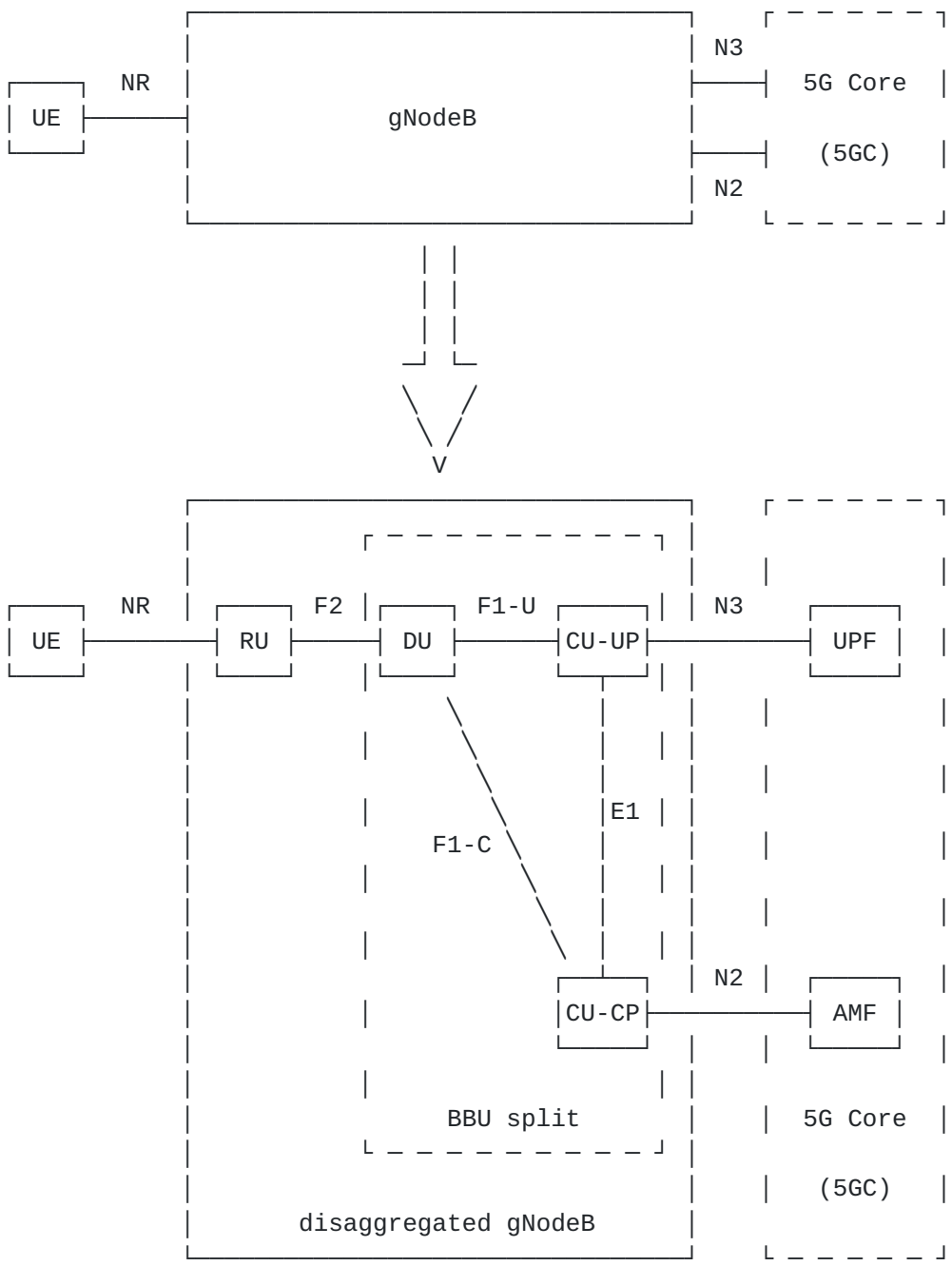


Figure 35: RAN Disaggregation

C.4. Transport Network (TN)

The 5G transport architecture defines three main segments for the Transport Network, which are commonly referred to as Fronthaul (FH), Midhaul (MH), and Backhaul (BH) [TR-GSTR-TN5G]:

- *Fronthaul happens before the BBU processing. In 5G, this interface is based on eCPRI (Enhanced CPRI) with native Ethernet or IP encapsulation.

- *Midhaul is optional: this segment is introduced in the BBU split presented in Appendix B.3, where Midhaul network refers to the DU- CU interconnection (i.e., F1 interface). At this level, all traffic is encapsulated in IP (signaling and user plane).

- *Backhaul happens after BBU processing. Therefore, it maps to the interconnection between the RAN and the Core Network. All traffic is also encapsulated in IP.

Figure 36 illustrates the different segments of the Transport Network with the relevant Network Functions.

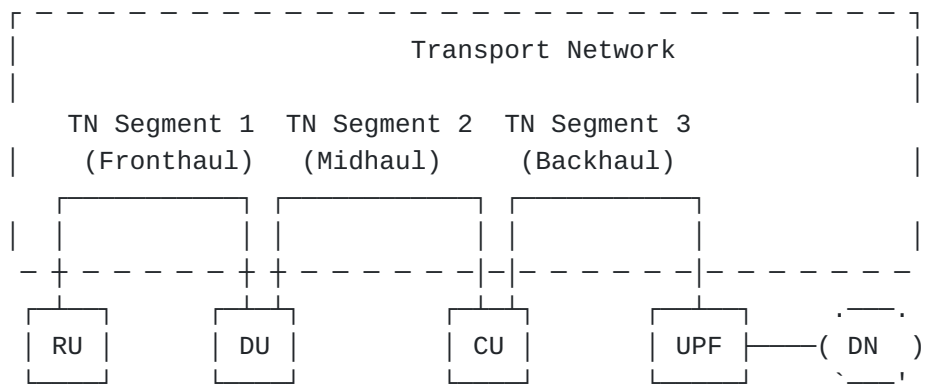


Figure 36: 5G Transport Segments

It is worth mentioning that a given part of the transport network can carry several 5G transport segments concurrently, as outlined in Figure 37. This is because different types of 5G network functions might be placed in the same location (e.g., the UPF from one slice might be placed in the same location as the CU-UP from another slice).

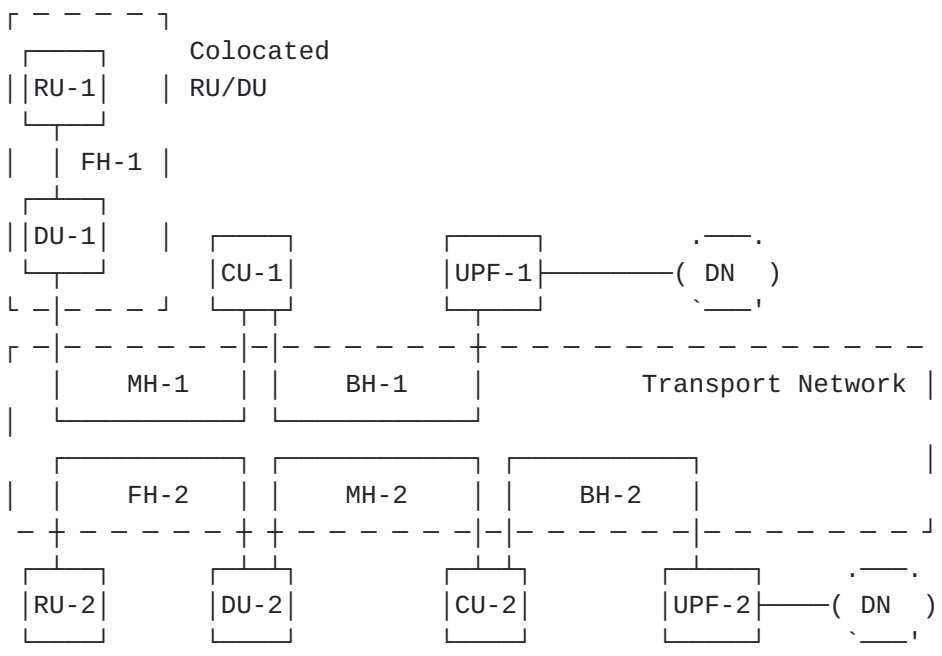


Figure 37: Concurrent 5G Transport Segments

Acknowledgments

The authors would like to thank Adrian Farrel, Joel Halpern, Tarek Saad, and Jie Dong for their reviews of this document and for providing valuable comments.

Contributors

John Drake
Juniper Networks
Sunnyvale,
United States of America

Email: jdrake@juniper.net

Ivan Bykov
Ribbon Communications
Tel Aviv
Israel

Email: ivan.bykov@rbbn.com

Reza Rokui
Ciena
Ottawa
Canada

Email: rrokui@ciena.com

Luay Jalil
Verizon
Dallas, TX,
United States of America

Email: luay.jalil@verizon.com

Beny Dwi Setyawan
XL Axiata
Jakarta
Indonesia

Email: benyds@xl.co.id

Amit Dhamija
Rakuten
Bangalore
India

Email: amit.dhamija@rakuten.com

Mojdeh Amani
British Telecom
London
United Kingdom

Email: mojdeh.amani@bt.com

Authors' Addresses

Krzysztof G. Szarkowicz (editor)
Juniper Networks
Wien
Austria

Email: kszarkowicz@juniper.net

Richard Roberts (editor)
Juniper Networks
Rennes
France

Email: rroberts@juniper.net

Julian Lucek
Juniper Networks
London
United Kingdom

Email: jlucek@juniper.net

Mohamed Boucadair (editor)
Orange
France

Email: mohamed.boucadair@orange.com

Luis M. Contreras
Telefonica
Ronda de la Comunicacion, s/n
Madrid
Spain

Email: luismiguel.contrerasmurillo@telefonica.com
URI: http://lmcontreras.com/