none K. Makhijani

Internet-Draft J. Qin

Intended status: Informational R. Ravindran

Expires: November 25, 2017 Huawei Technologies

L. Geng

China Mobile

L. Qiang

S. Peng

Huawei Technologies

X. de Foy

A. Rahman

InterDigital Inc.

1. Galis

University College London

May 24, 2017

Network Slicing Use Cases: Network Customization for different services

draft-makhijani-netslices-usecase-customization-02

Abstract

Network Slicing paradigm enables creation of end-to-end partitioned

network infrastructure which includes the UE, Access/Core transport

networks, edge and central data center resources to provide isolated

network platforms for various service verticals. The motivation

behind Network Slicing (NS) is to allow deployment of new services

with heterogeneous SLA requirements without causing or experiencing

any degradation due to other already running or deployed services in

the network. The purpose of this document is to focus on use cases

that benefit from the usefulness of network slicing.

Status of This Memo

This Internet-Draft is submitted in full conformance with the

provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering

Task Force (IETF). Note that other groups may also distribute

working documents as Internet-Drafts. The list of current Internet-

Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months

and may be updated, replaced, or obsoleted by other documents at any

time. It is inappropriate to use Internet-Drafts as reference

material or to cite them other than as "work in progress."

This Internet-Draft will expire on November 25, 2017.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the

document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal

Provisions Relating to IETF Documents

(http://trustee.ietf.org/license-info) in effect on the date of

publication of this document. Please review these documents

carefully, as they describe your rights and restrictions with respect

to this document. Code Components extracted from this document must

include Simplified BSD License text as described in Section 4.e of

the Trust Legal Provisions and are provided without warranty as

described in the Simplified BSD License.

Table of Contents

1. Introduction

1.1. Requirements Language

1.2. Terminology

2. Network customization for diverse services

2.1. Overview

2.2. Strict Resource Demand Concept

2.3. Network Customization Concept

2.4. Scope of use cases

3. A Generalized Network Slice as a Service

3.1. Virtualized Customer Premise Equipment

3.1.1. Traditional CPEs

3.1.2. Trends in CPE infrastructure

3.1.3. vCPE as a network slice

4. Network Slicing in 3GPP Mobile Network

4.1. Network Slices in 3GPP Systems

4.2. Challenges

4.3. Creating 3GPP Network Slices

4.4. Managing 3GPP Network Slices

4.5. Operating 3GPP Network Slices

5. Services with Resource Assurance

5.1. Enhanced Broadband

5.1.1. Media delivery networks

5.1.2. Enhanced Media Streaming Description

5.1.3. eMBB Type Slices

5.1.4. Network Operator's View

5.2. Massive machine to machine communication

5.2.1. Wireless Sensor Networks

5.2.2. Massive Internet of Things Description

5.2.3. mMTC Type Slices

5.2.4. Network Operator's View

5.3. Ultra-reliable low latency communication

5.3.1. Brief introduction

5.3.2. Challenges

5.3.3. Network Operator's View

5.4. Critical Communications

5.4.1. Public Safety Infrastructure

5.4.2. Enhanced Critical Service Type Slices

6. Network Infrastructure for new technologies

6.1. ICN as a Network Slice

6.1.1. Information Centric Networks Description

6.1.2. ICN Type Slices Asks

6.1.3. Network Operator's View

6.2. Network Slices in a Communication Endpoint

6.2.1. Connected Vehicle

6.2.2. Sliced User Equipment

6.2.3. Network operator view

7. Overall Use case Analysis

7.1. Requirements

7.2. Considertations

8. Conclusion

9. Security Considerations

10. IANA Considerations

11. Acknowledgements

12. References

12.1. Normative References

12.2. Informative References

Authors' Addresses

1. Introduction

Network Slicing (NS) is a mandatory requirement in 5G to meet the

diverse service requirements in different 5G service scenarios. NS

refers to the managed partitions of physical and/or virtual network

resources, network physical/virtual and service functions [RFC7665]

that can act as an independent instance of a connectivity network

and/or as a network cloud [I-D.gdmb-netslices-intro-and-ps]. In 3rd

Generation Partnership Project (3GPP) [TR23.799] defines "Network

slicing enables the operator to create networks customized to provide

optimized solutions for different market scenarios which demands

diverse requirements, e.g. in the areas of functionality, performance

and isolation". Draft [I-D.gdmb-netslices-intro-and-ps] defines

network slicing in a broad context and suggests related problems and

work areas. Other organizations like Next Generation Mobile Networks

(NGMN) [Network-Slicing-Concept] and ITU-T FG IMT-2020

[FG-IMT2020-Gaps] also present their separate definitions of NS. To

maximize resource utilization and minimize infrastructure cost,

services will need to operate over a shared network infrastructure,

as against the traditional monolithic model operated either as

dedicated network or as an overlay. Service operators can utilize or

benefit from Network Slicing through multi-tenancy, enabling

different customized network infrastructures for different group of

services across different network segments and operating them

independently. Moreover, NS should also guarantee the isolation

between different network slices both from resource and security

perspective. That is, the operation of the data packets traversing

one network slice must not adversely affect the service operation in

other network slices sharing the same underlying packet network.

This document describes use cases, where service operators can

utilize or benefit from Network Slicing through multi-tenancy,

enabling different customized infrastructures for different group of

services across different network segments and operating them

independently. Although 5G will drive NS based deployments, the

scope of the document is not limited to 5G; it covers example

scenarios specified from 5G vision as well as generalized scenarios

that can be applied to existing infrastructures.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",

"SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this

document are to be interpreted as described in RFC 2119.

Additionally, the key words "MIGHT", "COULD", "MAY WISH TO", "WOULD

PROBABLY", "SHOULD CONSIDER", and "MUST (BUT WE KNOW YOU WON'T)" in

this document are to interpreted as described in RFC 6919.

1.2. Acronysms and Abbreviations

o NS: Network Slcinig. This acronym will be used as short-form for

network slicing all across the document.

o CSP: Communication Service provider.

o CPE: Customer premise equiment. It is a device of set of devices

managed by CSP to provide remote connectivity functions.

o SD-WAN controller: A software defined WAN controller, a network

service function that manages connectivity across multiple sites

via on-site CPEs.

o mMTC: Massive Machine Type Communications.

o uRLLC: Ultra-Reliable and Low Latency Communication

o V2X (Vehicle-to-everything): Is a communication of information

from a vehicle to any other entity that may be a user end-device,

network element or application end point.

o RAN: Radio Access Network

o ITS (Intelligent Transportation Systems): Considered as an aspect

of how using Internet of Things resource like road sensors can

creates a smart transport network. The network provides services

relating to transport and traffic management systems through flow

of information between sensors, smart devices and humans.

o Over-the-top (OTT): A service, e.g., content delivery using a CDN

or a social networking service, operated by a different service

providers to which the users of the NSP service are attached to,

and to whom it serves as a communication (or bit pipe) provider

o Industry vertical: A collection of services or tools specific to

an industry, trade or market sector. also, referred to as Service

Verticals in this document.

o TETRA: Terrestrial trunked radio is a digital trunked mobile radio

standard to meet needs of public safety, transportation and

utilities like organizations.

2. Network customization for diverse services

2.1. Overview

[Note: To discuss diverse set of service attributes] Often services

specify a broader resource requirements to offer desired QoE to it

consumers ,while the underlying infrastructure is generally best-

effort. Traditionally, basic service guarantees (is this with

reference to the best effort point earlier ?)are associated with

resource attributes such as: \* Throughput \* Packet loss \* Latency \*

Network Bandwidth/Burst or other bit rates. \* Security In addition,

other attributes as mentioned below are embedded into the

infrastructure to improve over all quality of experience

o Redundancy

o Reliability

o Authentication (this should be part of security ?)

More recently, other service attributes have become significant such

as \* Continuity during mobility \* Purpose-built network functions. \*

Opportunistic service placement

It should be possible for the providers of any service to

continuously evolve, adapt, and differentiate themselves through

purpose-built infrastructures with minimal impact on network

deployment and operations. The motivation behind Network slicing

paradigm is exactly that . By creating logically partitioned network

infrastructures, isolated platforms for various industry verticals

can be provided. NS is envisioned to enable new service deployments

without having to build new network infrastructures or causing

disruptions to other already deployed services in the network. In

regards to NS, there are two primary characteristics, a) Strict

demand for network resource , b) Network Customization.

2.2. Strict Resource Demand Concept

Several services are sensitive to response times and/or amount of

bandwidth, e.g. realtime interactive multimedia, high bandwidth video

feed or remote access to an enterprise network . Failure to meet

these criteria leads to service degradation. Moreover, newer

scenarios from different industries are evolving due to these factors

- a) everything connected, b) technological advancements in sensors,

IoT, robotics and multi-media, c) innovations in social network

interactions (including both human-human or human-machine). These

may impose even stricter and more specific set of resource and

connectivity requirements on per service basis. The challenge lies

in utilizing common network infrastructure and judiciously allocating

available infrastructure resources.

2.3. Network Customization Concept

Network slicing is enabled through customization. Customization

gives control to the operator (of a slice) to create, provision,

change and network resources to suit their service demands. It

requires ability to decompose resources from an underlying network

infrastructure and logically aggregate them as part of a slice.

These customizations are not only in the context of the network

resources but also include placement and logical connection of the

network functions based on the service requirements.

2.4. Scope of use cases

Resource slicing is defined as to logically partition network

resources among multiple users, allowing multiple sub-networks to

exist simultaneously (planetlab - <https://www.planet-lab.org/>).

Network slice is a managed group of network resources, network functions

and service instances. While writing use cases following were considered o A Network Slicing aware infrastructure allows operators to use part of the

network resources to meet stringent resource requirement as in

Section 2.2 or exploit dynamic customizations as described in [#ns-

customization]. Finally, there will be scenarios that require both

customization and strict resource requirements.

o The document doesn't specify whether a network slice consists of a single

or multiple service(s). At the simplest level one service may

correspond to a slice, however, it is possible that many services may

become part of the same slice for the purpose of isolated data or

context sharing. o Use cases below are discussed from 2 perspectives

- a. Newer scenarios: that should absolutely meet strict resource

requirements, as if they use a dedicated infrastructure. The example

use cases are categorized in Section 5.

Use cases for each use case the answers to the following questions: o

What is the use case? o How is it addressed today? o How would you

like it to be addressed in the future? o What do you want the IETF to

deliver?

3. A Generalized Network Slice as a Service

TODO (Liang): Add generalized NS-aaS

3.1. Virtualized Customer Premise Equipment

3.1.1. Traditional CPEs

A CPE is any service provider equipment located at customer premises.

IT may be either a layer-2 or a layer-3 device (the routing gateway)

performing different network functions depending on the access

technology (DSL modem, PON modem, etc.). Any services provided such

as Internet access, IPTV, VoIP, etc. or network functions for

example, local NAT, local DHCP, IGMP proxy-routing, PPP sessions,

routing, etc. are also part of CPE. The installation of different on

premise devices, entails a high cost for service providers in terms

of both initial installation and operational support, since they are

typically responsible for the end-to-end service.

+-----+ campus

|----| CPEx | -----[ ]

| +-----+

----- Broadband | +-----+ branch

( ) ----------------|--->| CPEy |------[ ]

( CSP ) MPLS | +-----+

(\_\_\_\_) access| +------+ main site

|--->| CPEz |----- [ ]

+------+

Figure 1: Traditional CPE architecture

Traditional CPE deployments are shown in figure Figure 1. These are

service provider network functions installed on customer site to

provide above mentioned functionalities along with remote site

connectivity. CSP is responsible for management and administration

of connections and state with proper policy, bandwidth, security and

QoS requirements.

3.1.2. Trends in CPE infrastructure

A virtualized CPE architecture moves several network functions from

on premise to the service provider network to facilitate provisioning

of new services to customers based on a lean CPE functions on

premises such as minimizing number of network functions on customer

premises, perhaps only layer-2 visibility among them with no need for

routing gateways in the home network is suppressed. Several routing,

NAT, firewall capabilities may be performed in the service provider's

cloud. A customer's site is highly simplified with vCPE solution,

perhaps requiring only access level connectivity on premise and

moving other network functions to ISP's cloud.

|-----------------------|

| +------+ |------------------+-------+ campus

| |--| | | | vCPEx | -----[ ]

| | | | |------------------+-------+

| | | | | <====Broadband ==>

| ----- | | vCPE | | ------------------+-------+ branch

| ( ) |->| | | | vCPEy |------[ ]

| ( CSP )| | | |-------------------+-------+

| (\_\_\_\_\_) | | | |<==== MPLS/4G. ==>

| | | | |-------------------+-------+ main site

| |->| | | | vCPEz |----- [ ]

| +------+ |-------------------+-------+

|-----------------------|

Figure 2: irtualized CPE, with distributed architecture

Figure 2 shows a virtualized architecture in which many functions are

moved to CSP's cloud simplifying CPE on premises tremendously.

Additional details of deployment architecture models are captured in

[I-D.pularikkal-virtual-cpe] where full dissemination of data path

and control plane functions is described. Here only a high-level

relevance of virtualized CPE is shown. The figure shows vCPEx,

vCPEy, vCPEz are virtualized CPEs on multiple sites of a specific

customer, there may be set of different network functions in each x,

y and z CPE. The vCPE instance in CSP cloud is integrated to each

site performing service chains of network functions and resource

allocations specific for ingress and egress path of each site.

3.1.2.1. Challenges

A vCPE is a well-known concept which when combined with WAN

technologies provides end to end visibility and reachability to

remote sites. It has been solved using network function

virtualization (NFV) approaches and via offload of compute intensive

functions to the CSP cloud for ease of management by CSP. However,

there is no standard approach to connectivity or management of

various CPE functions. Furthermore, it is highly desirable for

customers to control and monitor their own network resources at both

remote and local sites. Using network slicing, a greater level of

agility can be achieved, with each customer dynamically managing its

own network with the assistance of network slicing framework.

3.1.3. vCPE as a network slice

The benefit of self-managing a vCPE network slice is the capability

to move network functions on premise of to the cloud. An obvious use

case will be customer initiated gradual migration of network

functions from a site to CSP cloud.

+-------------+ +-------------+

| Global Slice| | Slice |

| Mgr | | Resource Mgr|

+-------------+ +-------------+

| ^

| NS protocol or i/f |

V V

|--------------------------------------------------|

| |

| +-------------+ +-------------+ |

| | vCPE Slice | | CSP | |

| | Mgr/Monitor | | vCPE subnet | |

| +-------------+ +-------------+ |

| |

| +--------+ +--------+ +--------+ +--------+ |

| | vCPEy | | vCPEy | | trans | | vCPEz | |

| | subnet | | subnet | | subnet | | subnet | |

| +--------+ +--------+ +--------+ +--------+ |

| |

|--------------------------------------------------|

| |

| NS transport protocol or i/f |

V V

[Campus] [branch] [Transport] [main site]

Figure 3: vCPE as a Network Slice

In Figure 3, a slice for vCPE is shown. Using slice subnet approach,

each vCPE site instance may be considered as a subnet, along with the

WAN transport as another subnet. The network functions are chained

in a distributed fashion between site vCPEs and CSP vCPE subnet. A

monitoring function interfaces with CSP's global slice manager for

resource management and an interface to physical infrastructure

through network slice transport protocol, realizes these functions on

the infrastructure.

3.1.3.1. Network Operator's View

Having a dedicated sliced network catering to dynamic customization

of network functions with guaranteed resource method, simplifies

network operations. In case of such vCPE type solutions, it is

common for each customer to have its own private IP address space,

therefore, the resource isolation must include address isolations as

well. This may be achieved based on existing label techniques or

through new network slicing data path protocol.

4. Network Slicing in 3GPP Mobile Network

Network Slicing is a core feature of the currently under development

3GPP 5G Phase 1 Mobile System, because it makes it possible for

different vertical applications, such as IoT and broadband

applications, to be deployed over a common infrastructure. More

details can be found in [TS\_3GPP.23.501], [TS\_3GPP.23.502]

[TR\_3GPP.38.801], [TR\_3GPP.33.899], and [TS\_3GPP.28.500].

4.1. Network Slices in 3GPP Systems

In 3GPP systems a Network Slice is a complete logical network which

provides telecommunication services and network capabilities.

Distinct Radio Access Network (RAN) Network Slices and Core Network

Network Slices will interwork with each other to provide mobile

connectivity. A device may access multiple Network Slices

simultaneously through a single RAN.

3GPP defines Slice IDs (named (S-)NSSAI in the standard) composed of

a Slice Service Type (SST) and optionally a Slice Differentiator

(SD). SST refers to an expected network behavior in terms of

features and services (e.g. specialized for broadband or massive

IoT), while SD helps distinguishing among several Network Slice

instances.

Figure 4 describes the general layout of Network Slicing in Mobile

Networks. A Core Network Slice is composed, on the control plane

side, of a Session Management Function (SMF), which manages PDU

sessions, and on the user plane side, of a User Plane Function (UPF)

and possibly other functions. It is interconnected with a RAN Slice

to complete the user plane. Some functions on the Control Plane are

common and shared between multiple RAN and Core Network Slices. A

primary example of such a shared function is the Access and mobility

Management Function (AMF).

Common Functions Core Network Slice Instance

+-----------------+---------------------+

| +--------+ | +--------+ |

| | Control| | | Control| |

+--------+ Plane +----------+ Plane | |

| | | AMF... | | | SMF... | |

+---+--+ | +--------+ | +----+---+ |

|Device| +-----------------+ | |

+---+--+ | +--------+ | +------+-----+ |

| | | | | | User Plane | | +---------------+

+--------+ RAN +--------+ Functions +------+Data Network or|

| | | | | UPF... | | | The Internet |

| +--------+ | +------------+ | +---------------+

+-----------------+---------------------+

RAN Slice Instance

Figure 4: 3GPP Network Slices

4.2. Challenges

A core challenge here is to identify or develop a set of technologies

suitable to implement the infrastructure over which 3GPP Network

Slicing will be built, without requiring major rework of the 3GPP

specifications. Among the specific challenges that an IETF NS

framework will need to address, it will need to support sharing

Network Functions between several slices, building slices recursively

from smaller slices, implementing roaming across different domains,

etc. The following subsections describe creation, management and

operation of 3GPP network slices as currently planned in the

specifications, in order to better understand those challenges.

4.3. 3GPP Network Slices

To create a Network Slice Instance, Mobile Network Operators will

start by describing it by assembling together "Network Slice

Subnets", which are smaller components included in a RAN or Core

Network Slice. Network Slice Subnets include NFs and reserved

network resources, in term of KPIs such as minimum and maximum

throughput, delay, packet loss, etc. Network Slice Subnets can be

shared between several Network Slices. Both Network Slices and their

Subnets are described by the operator through the OSS/BSS management

system. The OSS/BSS translates this input from the operator into

descriptors that are sent to an Orchestrator. The Orchestrator,

through the rest of the NFV-MANO system, configures compute and

network elements to create Network Slice Subnets and compose them as

a Network Slice. Beyond creation, RAN or Core Network Slice

activation is orchestrated as the activation of individual Subnets,

possibly in a given order. Network Slices are isolated from each

other to avoid control plane congestion on one slice (e.g. using one

SMF in slice dedicated for broadband applications) to affect the

control plane of other slices (e.g. to affect potentially critical

IoT applications). Since some common Core Network functions (AMF,

PCF, UDM, etc.) are shared between multiple dedicated Core Network

Slices, the interaction between shared NFs and NFs in dedicated

Network Slices should be isolated from each other as well.

Network Slices creation will support different combinations of "n"

network services, "m" client devices and "p" interconnections with

external (sliced or non-sliced) networks and services. In 3GPP, RAN

and Core Network Slices are typically dedicated to a certain type of

network services such as Broadband or IoT, but may serve one or more

network services of a this type. Additionally, in some Mobile

Networks, parts of the Core Network may not be implemented over a

slice, while others are (e.g. SMF could be in a slice, while common

functions are not). While this can lead to a sub-optimal isolation

between slices, this effect can be partially compensated by

overprovisionning non-sliced sections of the network.

4.4. Managing 3GPP Network Slices

Mobile Network Operators can modify the configuration of a RAN or

Core Network Slices, while it is in use. Example of such operations

include:

o Increase or decrease compute capacity of NFs

o Increase or decrease network capacity

o Update the configuration of NFs

o Add, replace or remove a NFs

o Add, replace or remove a Network Slice Subnet

Some operations affecting a shared slice may not be possible without

affecting other Network Slices, and in this case may be replaced by

other operations: for example, instead of changing the configuration

of a shared AMF to accommodate the needs of a SMF, another Network

Slice Subnet with an AMF may be created, and replace the original

AMF's slice for this SMF.

The management system monitors performance of individual Network

Slice components and coalesce performance data and events for the

whole RAN or Core Network Slice. This includes user and control

traffic load data, QoS/SLA data, e.g. indicating whether services

were provided at expected QoS/SLA level. The management system uses

this information for example to decide to scale up or down NFs.

Performance data and events from a shared Network Slice component

will be attributed by the 3GPP management system to one of the RAN or

Core Network Slices that contain or interact with this shared

component.

To support roaming, Mobile Network Operators will need to configure

the interconnection between Network Slices on the Home Network and

Network Slices on the Visited Network. On the visited side, the

Operator ensures that the proper Network Slice is selected for a

roaming device. User traffic will flow through the visited network

slice either directly to an external data network, or through the

interconnected home network slice (both cases will need to be

supported). From the end user perspective only the performance of

the whole (visited + home) network slice is important.

Mobile Network Operators may expose limited 3GPP Network Slice

management to third party Communication Service Providers (CSP), who

may in turn consume this service or provide it to their own

customers, as a form of "Slice as a Service" described in Section 3.

Using this interface a CSP can request the creation of a Network

Slice using specifications of NFs, isolation, security, performance

requirements (such as traffic demand requirements for the coverage

areas, QoS for service). When an Operator exposes management data

(e.g. fault management data, performance data) about a Network Slice

shared by multiple customers of a CSP, exposed management data of

each customer can be isolated from each other.

+--------+

Limited NS | |

Limited NS Instance |Customer|

+-------+ Instance +-------------+ Management | |

|Mobile | Management |Communication+<-----------+--------+

|Network+<------------+Service |

| | |Provider +<-----------+--------+

+-------+ +-------------+ Limited NS | |

Instance |Customer|

Management | |

+--------+

Figure 5: 3GPP Limited Network Slice Management Exposure

4.5. Operating 3GPP Network Slices

TODO

5. Services with Resource Assurance

5.1. Enhanced Broadband

Today, video consumes the largest amount of bandwidth over the

Internet. As the higher resolution formats enter mainstream, even

more bandwidth will be needed to stream 4K/8K/360 degree formats.

The scenario in this section are discussed in regards to need for

demands beyond best-effort network delivery, in particular

requirements due to growth in data rate capacity, connection density

and interactive media. These are equally applicable to both fixed

and mobile networks.

5.1.1. Media delivery networks

+-----+

|=>| DASH|

| +-----+

+------------+ +-------------+ ----- +-----+ | +-----+

| Content |<==>| Transcoding |<=> ( ) ==>| CDN |=|=>| HDS |

| Aquisition | | Function | ( ISP ) +-----+ | +-----+

+------------+ +-------------+ (\_\_\_\_) | +-----+

|=>| HLS |

+-----+

Media delivery formats

Figure 6: Traditional Streaming Media Infrastructure

5.1.2. Enhanced Media Streaming Description

Today the video output format is HD with 1080p resolution with few

services delivering up to 4K. Both Video-on-demand and live-linear

channels (streaming live event feed) can be supported. Most often

media services are delivered using streaming platforms Figure 6.

5.1.2.1. Factors Influencing Enhanced Broadband Use Cases

Media delivery comprises of different functional components, as shown

in Figure 6 above and often an overlay or OTT infrastructure is used.

The deployment requires acquiring content, transcoders and CDN

servers and decoders to support different delivery formats All these

may be considered specialized service functions in media streaming

infrastructure. The entire operation is (a) not flexible in terms of

resources placement (on premise vs cloud vs proximity to destination)

(b) is built on best-effort of available resources, (c) Is reactive

when the congestion occurs leading to client-server based end to end

stream optimization derived from network conditions.

5.1.2.2. Traditional Media Streaming Service Verticals

There are 3 categories of media or content distribution

a Video on Demand (VOD)

b Live streaming/Linear channels

c Video conferencing

While a and b are one way content consumption, Video conferencing

requires 2-way or multi-way connection. It may consist of either

person-person or person-group video communication.

5.1.2.3. New Verticals - Virtual Reality (VR)/Augmented Reality (AR)

Virtual Reality(VR)/Augmented Reality(AR) is the future use case of

eMBB services. A 360-degree video is mostly low resolution,

requiring ~25 Mbps network bandwidth for streaming. For a network

based AR/VR bandwidth required will be in the order of Gbps and

latency less than 10 milliseconds for a fully immersive experience

such as cloud-based VR gaming, fully-interactive media experience.

However, media processing for AR/VR will still be identical to in-

network processing functions as shown in figure 1 and corresponding

latencies could lead to downgrade of user experience. Therefore,

upon request for an AR/VR stream a special infrastructure is required

that differs from best-effort network.

5.1.3. eMBB Type Slices

A purpose-built network slice for eMBB streaming shall ensure to

minimize processing overheads, it may be done by placement of network

functions closer to subscribers.

o Resource scaling: eMBB resources should be allocated dynamically

because bandwidth is expensive and requirements are high, such

vertical service operators may not want to pay for unutilized

bandwidth. Therefore, slices should adjust in negotiated chunks

of scale both bandwidth and service functions. For example, if a

stream is viewed by 8 people initially, the resource for 20 users

is allocated. It will subsequently grow or shrink in chunks of

resource for 20 subscribers.

o The transport aspects of eMBB has multiple aspects

\* Fan-out network: user to access network resource slicing

constrain specification

\* Point to Point or virtual network: resource specification for

content acquisition to distribution network.

o Latency Guarantee varies for live streaming, on-demand streaming

and connected AR/VR streaming

o Slice priority: eMBB slices must also be consideration that the

resources allocated does not compromise availability of emergency

services.

+----------------------------------+

| Provider Slice Orchestrator |

| |

| +------------------+ |

| | eMBB Resource | |

| +--> | Spec Guard |---+ |

| | +------------------+ | |

| | | |

| | +----------+-------+ | |

| +--->| Resource Monitor|<--+ |

| +---------+--------+ |

| ^ | |

|-----------+-------------+--------+

| |

| Real time feed|back

| |

eMBB | |

Network | v dynamic resource adjustment

+------------+------------+-------------+

| +----------+-------+ +-----------+ |

| | Acquired Content|<-->| eMBB slice| |

| | subnet | | Customizer| |

| +---------+--------+ +-----------+ |

| | | | +-+

| | | =======> | |

| +--------+ +-------+ | | +-+ handheld

| | CDN1 | | CDN2 | | | +---+

| | subnet | | subnet| ========>| |

| +--------+ +-------+ | | +---+ PC

| | | | |

| +-----------------+ | | +---------+

| | Encoders subnet |================+=+====>| |

| +-----------------+ | +---------+ TV

+----+----------+---------+-------+-----+

Figure 7: Reference eMBB slice

See Figure 7 above for a reference slice.

5.1.4. Network Operator's View

A typical eMBB slice flow from a network operator is as follows

o There is an eMBB slice offering template/form. A service vertical

provider requests

1. Regional network locations of CDN and location of acquired

content.

2. Describes transport requirements for its own distribution

network comprising of connectivity between content acquisition

and Fan-out points.

3. A granularity of transport resource chunk.

4. It may request access to subscriber database from multiple

access network types (mobile, fixed) creating value add for

both service provider.

5. For each access type resource requirement is specified.

o Registers self with access rights to resource monitoring and

negotiation loop. Slice operator has an abstracted view of its

own slice instance topology.

o Network operator has end to end (acquired content to cached

content to user) visibility across different domain segments and

corresponding transport resources. A well-coordinated network

slice protocol enables resource allocation across different

segments.

Note in addition to eMBB, traditional CDN use cases can be deployed

in a slice as well, see examples in [RFC6770].

+-------------------------------------------------------------+

| +-----------------------+ |

| +-------->| Provider Slice Manager|<----+ |

| | +-----------------------+ | |

| | | |

| +------+-----------+ +-----------+-------+ |

| | Global | | eMBB Slice | |

| | Resource Manager |<------------> | Resource Allocator| |

| +------------------+ +-------------------+ |

| |

+-------------------------------------------------------------+

| |

------- NS control -------------- NS control--

| |

------------------ -----------------

| -------------- | | -------------- |

| | eMBB Manager | | | | eMBB Manager ||

| |-------------- | | -------------- |

| | | |

| | | |

| -------------- | | -------------- |

| | eMBB Network | | | | eMBB Network ||

| |-------------- | | -------------- |

-------------------- -----------------

| | | |

V V V V

------------------NS transport ----------------

| | |

V V V

---------------- ---------------- -----------

| Infrastructure | |Infrastructure | | DC |

| Domain A | | Domain B | | Domain C |

---------------- ---------------- -----------

Figure 8: Transport provider network operator view. See Figure for

reference

5.2. Massive machine to machine communication

5.2.1. Wireless Sensor Networks

Sensor networks are widely deployed in industries such as

agriculture, environmental monitoring and manufacturing. The general

workflow of wireless sensor network is provided in Fig. Figure 9.

6.Decided Behavior

+-------------------+

| |

+----v------+ |

| Sensor | |

|(1. Data | |

|Collection)| |

+----+------+ |

|2.Collected Data | 3.Aggregated +---------------------+

+------------->+----------+ Data | Data Center |

| Sink Node/ |----------> (4. Data Analysis |

| Base Station| | & |

+---------->+--------------+--<------| Behavior Decision) |

|2.Collected Data | 5. Decided +---------------------+

+----+------+ | Behavior

| Sensor | |

|(1. Data | |

|Collection)| |

+----^------+ |

| |

+-------------------+

6.Decided Behavior

Figure 9

Figure: Workflow of wireless sensor network

As figure Figure 9 shows, sensors mainly collect data & behavior;

rarely communicate with each other in traditional wireless sensor

network. While in the scenarios discussed in this section, sensors

or embedded devices will be more intelligent and carry out more

frequent interactions that raises more challenges for mobile

networks.

5.2.2. Massive Internet of Things Description

Machine-to-machine type communication will dominate communication

paradigm in various industries such as healthcare, manufacturing,

transportation, etc. In order to support the massive internet of

things, traditional mobile networks have to be redefined -- by

creating the connectivity fabric for everything and bringing new

levels of on-device intelligence.

5.2.2.1. Factors Influencing Massive Internet of Things Use Cases

There are three main challenges raised by Massive Internet of Things

use cases:

o Scalable connectivity: there will be billions of smart devices

connect to mobile networks worldwide by 2020;

o Wide area coverage: sensor could be embedded into various

household equipments, medical instruments, vehicles, or even

public facilities;

o Frequent small amount data transmission: due to limited power,

most of the embedded sensors work intermittently rather than

continuously.

5.2.2.2. New mMTC Verticals

A few examples of new types of scenarios that require unique

infrastructure are mentioned below.

5.2.2.2.1. Smart City

Smart city networks is an integration of several public

infrastructures together through M2M communications. For example

o Automatic metering for gas, energy, water, etc;

o Environment monitoring for pollution, temperature, humidity, etc;

o Light management inside buildings or even the whole city;

o Traffic signal control;

o Public safety alerting for natural disaster.

Building a smart city requires a variety of IoT networks to inter-

operate together; these IoT networks are run by different departments

with different access privileges for administration and access

control. A smart-city network should be isolated from the public

Internet.

5.2.2.2.2. E-Health

E-health refers to the application that remote monitor the physical

conditions (e.g., heart rate, pulse, blood pressure etc.), and

accordingly take necessary medical measures remotely. Being a life-

critical service, e-health communication network must be reliable and

fast but small-size of data exchange. In addition, the privacy and

security of user's data must be guaranteed.

5.2.3. mMTC Type Slices

mMTC involves potentially a large number of small and power-

constrained devices, therefore, resource allocation at scale is of

particular importance in mMTC type slices. Furthermore, different

kind of IoT devices may exhibit quite different traffic patterns

e.g., continuous (heart rate monitors) & periodic delay tolerant

(temperature sensors), delay sensitive (e.g., weather forecast &

disaster alerting), mobility mode, security awareness etc. The mMTC

type slices should be conscious of various requirements of scale,

data pattern, reliability, security and energy efficient

communications.

5.2.4. Network Operator's View

Different from eMBB and uRLLC type services, mMTC service does not

have so much strict requirements on bandwidth and latency. Massive

and ubiquitous connectivity support would become the biggest

challenge of mMTC service. That is, for an network operator, mMTC is

mainly concentrated in the access network side and most of the

information flow should not pass through the transmission or core

network, both for security and communication efficiency. The

mobility management IoT gateway functions could be placed closer to

terminals (e.g., base-stations, edge clouds, etc.). Consequently, an

mMTC type slice should consist of plentiful access network resource,

as well as normal yet reliable transmission network and core network

resources in general.

5.3. Ultra-reliable low latency communication

5.3.1. Brief introduction

Not only, mission critical communication services but industrial

manufacturing, production processes, remote medical surgery, and

transportation safety (high mobility cases), etc scenarios require

ultra-reliable communications with no packet loss.

5.3.2. Challenges

In uRLLC scenarios, both data and control planes may require

significant enhancements to transmission or information distribution

protoocols. [TR\_3GPP\_38.913] specifies generic KPIs for access

network userplane latency as 1ms and reliability factor of 99.999%

for transmission of a packet of size 32 bytes. Although KPIs vary

for different scenarios such as V2X(3-10ms, 99.999%), eMBB (4ms UL/DL

each), In order to meet these, latency and reliability of the

transport in mobile networks should also be considered.

5.3.2.1. New service verticals

In the following sections three new uRLLC scenarios are described.

5.3.2.1.1. Industrial Operation and Inspection

Operations in remote industry sites usually need the support of

mobile transport network. Accurately operating machinery (low

latency and jitter) from remote locations requires high-quality

communication links between the control site. Factors to consider \*

low latency and low jtter in communication path \* Short time interval

between an operator sending control signal tp equipment response.

In an industrial closed control loop (Sensor -Controller - Actuator)

as shown in figure Figure 10, a typical control cycle time where

network is involved should be below 10ms [White-paper-5GAA].

++++++++++ +++++++++++++++

+ Sensor +-->+ Transmitter +---+

++++++++++ +++++++++++++++ |

| ++++++++++++ ++++++++++++++

+-->+ Base +---->+ Controller +

+---+ Station +<----+ +

| ++++++++++++ ++++++++++++++

++++++++++ +++++++++++++++ |

+Actuator+<--+ Receiver +<--+

++++++++++ +++++++++++++++

Figure 10: Industrial closed control loop

5.3.2.2. Remote Surgery

Remote surgery which enables surgeons to perform critical specialized

medical procedures remotely, allowing their vital expertise to be

applied globally. Providing accurate control and feedback for the

surgeon entails very strict requirements in terms of latency,

reliability and security.

5.3.2.3. Vehicle-to-everything (V2X)

Vehicle-to-everthing (V2X) network uses precise knowledge of the

traffic situation across the entire road/highway network to optimize

traffic flows, reduce congestion, and minimize accidents. For uRLLC

scenario,

o V2X in access network uses Vehicular Ad Hoc Network (VANET) type

protocols for vehicle-to-vehicle and an access medium

communication (either ITS-band or commercial-cellular). The

topologies are dynamic and mobility is high. In order to support

fully autonomous reliable driving, a highly reliabile

communication channel is required.

o Often, V2X may involve a part transport and core networks for

functions such as subscriber/vehicle admission and intensive

computational resource for aggregating information from multiple

traffic zones.

5.3.3. Network Operator's View

A uRLLC network slice only accepts service specifc traffic and

discards any other type of traffic to avoid negative impact on uRLLC

service operation. Even within the same vertical different kind of

services should be isolated. For example, in the V2X vertical, the

network slice used for autonomous driving should not be used for in-

vehicle infotainment. Capabilities required by uRLLC service

provider include

o Locations of the access nodes for terminals (devices, vehicles) to

the transport network and locations of the controller to construct

its own network topology within the network slice. In high

mobility scenario such as automotive verticals, the dynamic

topology adjustments are required without loss of data.

o Each service vertical has different performance requirements in

terms of latency, reliability and data rate etc., therefore, the

uRLLC network slice should allow customization for these

parameters.

\* A uRLLC service provider should be able to registers self with

access rights to resource monitoring and negotiation loop.

From a network operator provides a uRLLC Slice with following

considerations

o Should support/provide specific data and control planes protocols

with significant enhancements for deterministic latency and

reliability (e.g. DetNet in data plane).

\* Allow uRLLC provider to access user admission and

authentication to its network slice in advance.

\* The network coverage for a uRLLC service provisioning may be

limited to a confined area, either indoor or outdoor, network

operator needs to be able to coordinate resource allocation

across different access types and network segments. The fig

Figure 11, shows provider and operator view of the network.

The monitoring of resources is done in the context of

performance. A performance degradation would require resource

adjustment. As shown below, in one possible sliced model will

have its own customizer that uses internal performance

observing logic with in its slice by coordinating with

different subnets/domains using southbound NS transport

protocol and transfers this information to operator via a

northbound NS protocol for resource adjustment. It is implied

that domains maybe different access technolgoies and need for a

common performance metric propagation and resource allocation

is important for a uRLLC slice to function properly.

+-----------------------------+

| Operator Slice |

| Orchestrator |

| +---------+ +-----+ | uRLLC service +---------+

| | Resource| | Perf| <-|---------------| uRLLC |

| +--- | view | | Spec| | template | service |

| | +---------+ +-----+ | +---------+

| | +----------+--------+ |

| +--->|Performance Monitor| |

| +---------+------^--+ |

| | | |

|------------------------|-+--+

| | resource adjustment

| |

performance metrics| |

| |

uRLLC slice | v

+---------+-------------+-------------+

| +--------+--+ +-----------+ |

| | Subs |<-->|uRLLC slice| |

| | Mgmt | |Customizer | |

| +-------+---+ +---------^-+ |

| +-------+------------| |

| | | +---v-----+ +

| +--------+ +-------+ | micro | |

| | GC-1 | | GC-2 | | resource| |

| | subnet | | subnet| | mgr | |

| +--------+ +-------+ +---------+ |

| | | |

+----+----------+---------+-------+---+

| | | |

V V V V

------------NS transport --------------

| | |

V V V

+--------------+ +------------+ +----------+

| Domain A | | Domain B | | Domain C |

+--------------+ +------------+ +----------+

Figure 11: Reference for uRLLC Network Slice.

5.4. Critical Communications

Critical communications are used during emergency situations. Often

referred to as mission critical, the communication has to be reliable

and non-disruptive. Different scenarios of critical communications

relate to public safety responders, military, utility or commercial

applications, mainly using reliable voice or short data messaging

over wireless communication systems. First responders such as

firefighters, paramedics and other responders, for their daily and

emergency communications needs to be able to communicate without

disruption.

5.4.1. Public Safety Infrastructure

5.4.1.1. Current Improvements over traditional services

Traditional technologies for emergency communications are narrow band

radio networks such as Land Mobile Radio (LMR) systems. They are

terrestrially-based professional push to talk wireless communications

systems commonly used for critical communications by public safety

organizations such as police, firefighters, and other emergency

response organization. LMR and related systems such as TETRA or P25

have dedicated frequencies and channels assigned to individual groups

of users for instant connection through a simple interface. Next-

generation public safety communications are planned to be built with

enhanced broadband voice, data and video communications services

beyond narrowband LMR with broadband LTE networks for high speed data

(ref 22.179 and FirstNet).

5.4.1.2. Challenges for Enhanced Critical communication

3GPP defines, on-network critical communication can be established

with the help of a network infrastructure to manage the call. It can

also be off-network, where the UEs communicate directly to each

other. The scope does not discuss point to point off-network

communication as it is not relevant to the topic. Most important

challenges for on-network communications include

o Expensive to deploy a separate broadband network: The coverage of

a separate network at the scale of area, state or nationwide that

is interoperable is not cost effective, especially as new

communication technologies emerge, public safety systems should be

able to adapt easily to state of the art.

o Lack of flexibility: in terms of adding new value added services

or ability to take advantage of commercial services.

o Ability to reliable support of basic mission critical services

such as voice: loss of information in voice communication is no

acceptable in emergency services, if common infrastructure is to

be used, it must assure no loss of information.

5.4.2. Enhanced Critical Service Type Slices

The traditional critical communications use dedicated separate

infrastructures in order to be reliable and non-disruptive. In

contrast, LTE based mechanisms acquire different bearer QoS Class

Identifier (QCI) for different type of barriers (data, voice, video).

The eMC (enhanced mission critical) network slices benefit from the

following:

o Insertion and authorization of subscribers in a group

communication: In a critical infrastructure, the subscriber

authentication may be done earlier at the entry point

automatically through slice selection functional entity.

o Pre-allocated QCIs: Generally, QCIs are requested on per session

basis which could slow down overall call control setup and is

undesirable for emergency services. When operating in a slice,

these resources maybe reserved ahead of time in a coarse-grained

manner instead of per session.

MC Network slices are relatively straight forward as it only concerns

with guaranteed bit rate (GBR) on per media basis and management of

groups. The MC network slice need an ability to request transport

services based on GBR for reliable communication. A reference

network slice below shows an MC organization providing service

agreement that would be a network slice template with resource

specification. The eMC slice sets up different subnetworks of

different subscriber groups and manages its membership. These

subnets are realized into the infrastructure across different domains

through a network slice transport mechanism. The MC network slice

must be capable of active resource monitoring to prevent congestions

to ever occur as well as request additional transport resources in

case of emergency event occurence.

+----------------------------------+

| Provider Slice Orchestrator |

| |

| +------------------+ | service +------------------+

| | eMBB Resource | |<-----------| Mission Critical |

| +--> | Spec Guard |---+ | agreement | Organization |

| | +------------------+ | | +------------------+

| | | |

| | +----------+-------+ | |

| +--->| Resource Monitor|<--+ |

| +---------+--------+ |

| ^ | |

|-----------+-------------+--------+

| |

| Resource request

| | prioritized resource adjustment

MC Network|Slice v dynamic group management

+------------+------------+-------------+

| +----------+-------+ +-----------+ |

| | Group Subs Mgmt |<-->| MC slice | |

| | | | Customizer| |

| +---------+--------+ +-----------+ |

| | | | +-+

| | | +---------+ + +-->| |

| +--------+ +-------+ | GRP | | +-+ MC-UE

| | GC-1 | | GC-2 | | selector| | +-+

| | subnet | | subnet| +---------+ | --->| |

| +--------+ +-------+ | +-+ MC-UE

| | | |

+----+----------+---------+-------+-----+

| | | |

V V V V

------------NS transport ----------------

| | |

V V V

---------------- ---------------- -----------

| Infrastructure | |Infrastructure | | MC server|

| Domain A | | Domain B | | Domain C |

---------------- ---------------- -----------

Figure 12: Reference for Mission Critical Network Slice.

6. Network Infrastructure for new technologies

6.1. ICN as a Network Slice

ICN as in Information-Centric Networking is a culmination of multiple

future Internet research efforts in various parts of the world, now

being pursued under IRTF's research task group called [ICNRG].

6.1.1. Information Centric Networks Description

Information-Centric Networking (ICN) addresses Internet's network

architectural design gaps based on evolving applications requirements

and end user behaviour which is significantly different from what IP

was designed for, which was optimized for host-to-host communication

paradigm. ICN is a non-IP paradigm based on name-based routing and

offers many desirable networking features to applications such as,

caching, mobility, multicasting and computing in a manner different

from traditional host-centric communication model. With respect to

5G and network slicing, ICN paradigm is in line with the move towards

service-centric architectures enabled through frameworks like SDN,

NFV, and Edge Computing. At a high level, ICN offers a name-based

abstraction to application that doesn't require further translation

(as in domain names to IP mapping in current IP networking), making

it suitable to several communication modalities such as multi-point-

to-multi-point, D2D and Adhoc communication.

6.1.1.1. New Verticals - ICN based service delivery

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

1) In ICN, applications, services and content are addressed using

names, hence end host resolution services like DNS can be avoided,

this achieves name resolution to edge content or services without

incurring additional RTT delays; 2) Service flows will be offered

mobility and multicasting support, as the networking is session-less

and optimized towards efficient movement of named data or networking

named services and host level communication; 3) Services can be

deployed at the very edges with ease as ICN routers are compute

friendly, this is because states in the forwarding table can be that

of either content or service resources; 4) Further saving bandwidth

in the upstream link through opportunistic caching is an inherent

feature of ICN, this also leads to energy efficient networking.

6.1.1.2. Considerations for Information Centric Network Applications

When offered as a programmable and customizable logical network

slice, ICN based services can be offered through a network slice in

parallel with traditional IP based services. ICN can be realized as

a slice [\_5GICN] based on the choice of data plane resource offered

by the operators in different segments of the network such as the

access, core network or central data centers. While the same

resources can be used to support services over IP, proper resource

isolation shall allow it to co-exist with ICN slices as well. ICN

though initially was aimed to server CDN application, it is equally

adept to server real-time applications such as audio/video

conferencing [ICN-AV] or AR/VR applciations. TODO (Ravi): different

kind of services that ICN can offer - e.g. IoT, multimedia ICN

assumes that the network slicing framework is built upon a

programmable pool of software and/or hardware based data plane

resources. The pool of resources comprises of o Hardware decoupled

network functions, that may be containers or VMs. o Deeply

programmable hardware resources include GPU, FPGAs [ClickNP], Smart

NIC [Netronome] operated using P4 abstractions, that are supported

over x-86 platform. Programmable hardware may also include

commercial chips supported using P4 or POF allowing one to realize

high performing novel data planes, e.g. [Barefoot]

6.1.2. ICN Type Slices Asks

In ICN, applications use Interest/Data abstractions over named

resources resolved by ICN's routing plane. An ICN slice shall be a

programmable ICN-domain, in which content learning and distribution

will be done using existing or new ICN aware routing and data plane

protocols. As a result, it should be possible to deploy network

functions such as ICN routers and content producers and distributors

that serve and speak ICN protocols. Just as multiple service

instances can be part of a slice, an ICN slices can multiplex

heterogeneous services; on the other hand an ICN slice can be as

granular as a service instance too. The latter approach has

implications with respect to consumer privacy, access control of name

data objects, and granularity of mobility handling.

6.1.3. Network Operator's View

A basic ICN slice can be manifested as a resource isolated logical

network while sharing resources with other connectivity or service

slices. An ICN slice relies on programmability and virtualization

framework to manage the service slices, to allow maximum flexibility

through logically centralized control plane for services. Through a

network slice template -

o ICN service providing entity could specify specific locations

(edge of network domains) to deploy ICN-routers or other ICN-NFs

(ICN aware network functions). Its service definition varies with

the type of service, for e.g. in case of a VoD service, it can

include the demand with respect user demand for a particular set

of content, distributed cache or storage resource, and compute

resources to execute video-centric service functions.

o An ability to establish connectivity between ICN network elements

in all segments and create an ICN based topology, this can be done

using specific service control plane based on application events

arriving in a dynamic manner

o Mechanism to carry ICN user traffic over the infrastructure, ICN

slice can be made aware to the RAN explicitly or implicitly using

traffic classification function at the edge or can be enabled in

an overlay manner.

o In addition, bandwidth and other network resources may be

requested.

How multiple services will be deployed within an ICN aware slice may

or may not be exposed to the network operator, depending on if the

ICN slices are natively managed by it or a by other service providers

6.2. Network Slices in a Communication Endpoint

In this section connected endpoint use case are described to

highlight significance of slicing in an end point.

6.2.1. Connected Vehicle

Connected vehicles are example of scenarios where a communication end

point is split into 3 different type of services that vary in in

terms of topology, bandwidth, latency, mobility and security.

a V2I in short-range: requires adhoc routing protocol, reliable data

plane and higher layer security and authentication;

b Traditional broadband for Infotainment: requires high speed

connection bandwidth.

c In network assistance for localized services: low speed, reliable

connection for a short period of time. This service need to be

highly secure and isolated because it connects vehicle to

manufacturers who can alter component settings.

6.2.2. Sliced User Equipment

a user equipment, if authorized may be allocated dedicated resource

for mission critical services and best-effort slice for normal

connectivity.

6.2.3. Network operator view

A network operator that registers a subscriber is required to know

how a UE is used and which services, offered as a slice it is part

of. An highly secure 3-way authentication between operator, service

provider and user is required to enable a slice on a device.

7. Overall Use case Analysis

The discussion in above usecases can be summarized as following in

terms of the requirements for network slicing framework.

7.1. Network Slicing Requirements

Four key requiremnts [NS Gaps] are identified from the above use case analysis:

* Network Slining Uniform Specification
* E2E Network Slicing
* Network Slicing Domain-Abstraction
* Network Slicing OAM Operations

We observed the following functional requirements of a lower

level granularity:

1. With or without network slice: Network slicing doesnot change the

functionality of a scenario; It only facilitates creation of an

isolated, an independently run infrastructure for that use case

over a common network. It may be relevant to define the concept

of a default slice, which corresponds to best-effort, traditional

network. Whereas, other type of instances of network slices may

be purpose-built for a service.

2. Network Slice is a network that is build on a infrastructure

composed of connectivity, storage and computing (connectivity

is only 1/3 part of the infrastructure).

3. Each network slice may have its own operator that see this slice

as a complete network (i.e router instances, programmability, using

any appropriate communication protocol, caches, provide dynamic

placement of virtual network functions according to traffic patterns,

to use its own controller, finally it can manage its network as its own).

4. Network slicing introduces an additional layer of abstraction by the

creation of logically or physically isolated groups of network resources

and network function/virtual network functions configurations separating

its behavior from the underlying physical network.

5. Subnet Concept or Abstraction: Each usecase has functionality

that can be logically split into subnets. Each subnet supports

only a part of functionality or interconnection. Therefore, a

proper abstract or logical representation of these subnets is a

required. A provider transport with assured network resources

wll be required to inter-connect these subnets.

6. Multi-domain coordination: In all scenarios mentioned, require

multi-domain coordination to connect and administer own subnets.

7. Network slicing would need to be selfmanaged with automated

deployment in order to cope with scalability.

8. Meet low latency or bandwidth demands: All scenarios require

agile resource adjustments. it may not be possible to achieve

this using centralize or API approach. It can also be difficult

to coordinate across different domains. Therefore, a network

slice transport protocol that standardizes resource propagation

in different subnets is needed. It is important for protocol (or

interface) to be lightweight and distributed.

9. Role of Network Function Virtualization: Although not discussed

in details of use cases, NFV plays an important role in terms of

dynamic placement of services, particioning of network resource

and configuring the network functions/virtual network functions

10. Ability to run own control and data plane: Many services need a

network slice for two important reasons, (1) it can not provide

optimal experience on a best-effort network, (2) it is

inefficient and expensive to build a separate infrastructure.

The separation from tradtional network, should allow new services

to use specific network and transport protocol for that service.

11. Similar to virtualization approaches, slice instances sould be

carried end to end with per domain translations performed.

12. Independent per slice management plane: Since a sliced network is

pupose-built, the intelligence to run, control, manage, operate

and administer a slice is with the provider of service in a

slice.

13. Mult-access knowledge: Many services are scoped within an access

domain that could be wireless or different cellular spectrum.

14. Network slicing supports dynamic multi-services, multi-tenancy and

the means for backing vertical market players

7.2. Considertations

These observatons impose several challenges on all network segments

(access, core, edge and trasport networks).

1. Each domain or network segment has different charateristics, for

example, it may be an L2, L3 fixed network or even mobile/

cellular network. Dissemination of resource characteristics

should be done uniformly across all networks to simplify slice

deployment.

2. Within its own domain different traffic engineering techniques

may be deployed for example, FlexE, MPLS, RSVP-TE or SDN based

TE.

3. No two network infrastructures are alike, technologies such as,

edge computing, NFV, SDN, cloud are partially deployed today.

There is no uniformity about whether a service is available as a

physical node or a virtual node. A network slice framework need

to be able to cater to both cases.

4. Optimal placement of resources on-demand is only possible when

infrastructure supports it. A capability exposure of a domain

could facilitate such functions.

5. At a massive scale, it is extremely complex to centralize global

view of resources and be able to distribute on-demand.

Considerations maybe made to incorporate domain-to-domain

communication about data and control for a specific network

slice.

6. Network operators would exploit network slicing for

• Reducing significantly operations expenditures, allowing also programmability necessary to enrich the offered tailored services.

• Providing the means for network programmability

• Additional business offerings to OTT and other vertical market players without changing the physical infrastructure (i.e. Health Vertical Sector, Energy Versical Sector, Automotive Vertical Sector, Media and Entertailment Versical Sector, Factory-of-the-Future Vertical Sector, Smart Home Vertical Sector, Smart City Vertical Sector, Additional Specialized Services Vertical Sector)

8. Conclusion

The goal of Network slices is not only to support various service

verticals with very different network capability and performance

demands but also simplify traditional service delivery models. The

use cases above are not an exhaustive list of all possible scenarios

but broadly network slicing scope shall cover the following types -

o An operator offering Network slice as a service.

o Network slicing for traditional service verticals.

o Network slicing for new resource-demanding service verticals.

\* scale intensive IoT type

\* latency and reliability sensitive

\* bandwidth sensitive

o Network slicing for services with own control and dataplane

transport.

o Network slicing in a communication endpoint

The document consider a network operator's view on how to offer

network slice as a service. There is a lot of emphasis on use

wireless or cellular technologies as one of the communication

segments. With in network slicing framework it becomes important to

generalize resource specific protocols that maybe used in both

wireline and wireless transports.

9. Security Considerations

The security considerations apply to each slice. In addition general

security considerations of underlying infrastructure whether isolated

communication with in a slice apply for links using wireless

technologies.

10. IANA Considerations

There are no IANA actions requested at this time.

11. Acknowledgements

Thanks to Stewart Bryant for providing details for several use cases.

Dong Jie for helping with review of the document.

12. References

12.1. Normative References

[I-D.gdmb-netslices-intro-and-ps]

Galis, A., Dong, J., kiran.makhijani@huawei.com, k.,

Bryant, S., Boucadair, M., and P. Martinez-Julia, "Network

Slicing - Introductory Document and Revised Problem

Statement", draft-gdmb-netslices-intro-and-ps-02 (work in

progress), February 2017.

[I-D.pularikkal-virtual-cpe]

Pularikkal, B., Fu, Q., Hui, D., Sundaram, G., and S.

Gundavelli, "Virtual CPE Deployment Considerations",

draft-pularikkal-virtual-cpe-02 (work in progress),

February 2017.

[RFC6770] Bertrand, G., Ed., Stephan, E., Burbridge, T., Eardley,

P., Ma, K., and G. Watson, "Use Cases for Content Delivery

Network Interconnection", RFC 6770, DOI 10.17487/RFC6770,

November 2012, <http://www.rfc-editor.org/info/rfc6770>.

[RFC7665] Halpern, J., Ed. and C. Pignataro, Ed., "Service Function

Chaining (SFC) Architecture", RFC 7665,

DOI 10.17487/RFC7665, October 2015,

<http://www.rfc-editor.org/info/rfc7665>.

12.2. Informative References

[\_5GICN] IEEE Communication, "Delivering ICN Services in 5G using

Network Slicing. 'Asit Chakraborti, Syed Obaid Amin,

Aytac Azgin, Ravi Ravindran, G.Q.Wang'", May 2017,

<https://arxiv.org/abs/1610.01182>.

[Barefoot]

Barefoot, "Barefoot", 2017,

<https://barefootnetworks.com>.

[ClickNP] ACM SIGCOMM, "ClickNP: Highly Flexible and High

Performance Network Processing with Reconfigurable

Hardware. 'B. Li, et al'", 2017,

<https://www.microsoft.com/en-us/research/wp-

content/uploads/2016/07/main-4.pdf>.

[FG-IMT2020-Gaps]

"FG IMT-2020: Report on Standards Gap Analysis",

<http://www.itu.int/en/ITU-T/focusgroups/imt-2020>.

[ICN-AV] IEEE Transaction on Emerging Network Architecture (under

submission),, "SRMCA: Scalable and Realiable Multimedia

Communication Architecture. 'Asit Chakraborti, Syed Obaid

Amin, Aytac Azgin, Ravi Ravindran, G.Q.Wang.'", 2017,

<https://arxiv.org/abs/1703.03070>.

[ICNRG] IRTF, "ICN Routing Group", November 2016,

<https://irtf.org/icnrg>.

[Netronome]

Netronome, "Netronome", 2017,

<https://www.netronome.com/products/agilio-cx/>.

[Network-Slicing-Concept]

Next Generation Mobile Network Alliance, "Description of

Network Slicing Concept", January 2016,

<https://www.ngmn.org/uploads/

media/160113\_Network\_Slicing\_v1\_0.pdf>.

[TR23.799]

International Telecommunications Union, "Study on

Architecture for Next Generation System", SA2 Study 3GPP-

TR23.799, December 2017,

<http://www.3gpp.org/ftp/Specs/archive/23\_series/23.799/>.

[TR\_3GPP.33.899]

3GPP, "Study on the security aspects of the next

generation system", 3GPP TR 33.899 0.6.0, November 2016,

<http://www.3gpp.org/ftp/Specs/html-info/33899.htm>.

[TR\_3GPP.38.801]

3GPP, "Study on new radio access technology Radio access

architecture and interfaces", 3GPP TR 38.801 1.0.0, March

2017, <http://www.3gpp.org/ftp/Specs/html-info/38801.htm>.

[TR\_3GPP\_38.913]

3GPP, "Study on scenarios and requirements for next

generation access technologies", 3GPP TR 38.913 14.2.0,

March 2017,

<http://www.3gpp.org/ftp/Specs/archive/38\_series/38.913>.

[TS\_3GPP.23.501]

3GPP, "System Architecture for the 5G System", 3GPP

TS 23.501 0.2.0, February 2017,

<http://www.3gpp.org/ftp/Specs/html-info/23501.htm>.

[TS\_3GPP.23.502]

3GPP, "Procedures for the 5G System", 3GPP TS 23.502

0.2.0, February 2017,

<http://www.3gpp.org/ftp/Specs/html-info/23502.htm>.

[TS\_3GPP.28.500]

3GPP, "Telecommunication management; Management concept,

architecture and requirements for mobile networks that

include virtualized network functions", 3GPP TS 28.500

1.3.0, 11 2016,

<http://www.3gpp.org/ftp/Specs/html-info/28500.htm>.

[White-paper-5GAA]

5G Automotive Association, "The Case for Cellular V2X for

Safety and Cooperative Driving", November 2016,

<http://www.5gaa.org/

pdfs/5GAA-whitepaper-23-Nov-2016.pdf>.

[NS Gaps]

"Gap Analysis for Network Slicing ", November 2016,

< draft-qiang-netslices-gap-analysis-00>.

Authors' Addresses

Kiran Makhijani

Huawei Technologies

2890 Central Expressway

Santa Clara CA 95050

Email: kiran.makhijani@huawei.com

Jun Qin

Huawei Technologies

Huawei Campus, No. 156 Beiqing Rd.

Beijing 100095

Email: qinjun4@huawei.com

Ravi Ravindran

Huawei Technologies

2890 Central Expressway

Santa Clara CA 95050

Email: ravi.ravindran@huawei.com

Liang Geng

China Mobile

Beijing 100095

Email: gengliang@chinamobile.com

Li Qiang

Huawei Technologies

Huawei Campus, No. 156 Beiqing Rd.

Beijing 100095

Email: qiangli3@huawei.com

Shuping Peng

Huawei Technologies

Huawei Campus, No. 156 Beiqing Rd.

Beijing 100095

Email: pengshuping@huawei.com

Xavier de Foy

InterDigital Inc.

1000 Sherbrooke West

Montreal

Canada

Email: Xavier.Defoy@InterDigital.com

Akbar Rahman

InterDigital Inc.

1000 Sherbrooke West

Montreal

Canada

Email: Akbar.Rahman@InterDigital.com

Alex Galis

University College London

Department of Electronic and Electrical Engineering

Torrington Place

London WC1E 7JE

United Kingdom

Email: a.galis@ucl.ac.uk