End-to-End Network Slice Orchestration and Automation

Nakjung Choi
Head of Mobile Network Systems Department
Network Systems and Security Research
Bell Labs Core Research, Nokia

January 19th, 2024

Agenda

Part I. End-to-End Network Slicing

Part II. Network Slicing Orchestration

Concluding Remarks

Part I. End-to-End Network Slicing

Diversity of Emerging Services

A wide variety of new use cases and business models

Emerging Use Case

Use Case require diverse network characteristics

	Networking Requirements							
	Data rate DL	Data rate UL	Device Density	Latency	Reliability	Mobility	Coverage	Security
Mobile Video Streaming (4K)	15Mbps	500kbps	<10k/km²	<1s	Low	Medium	Country	Low
Video Surveillance (static, 4K)	300kbps	10Mbps	500/km²	<1s	Low	None	City	Medium
Mobile VR/AR Cloud Gaming (4K)	15Mbps	1Mbps	<10k/km²	<10ms	Medium	Medium	Country	Low
V2X Platooning	50Mbps	10Mbps	250/km²	<20ms	High	High	Road/Global	Critical
IIoT Motion Control	1-10Mbps	1-10Mbps	100k/km²	1ms sync.	Very High	Low	Campus	Critical
Tracking and Tracing	100kbps	250kbps	1k/km²	>1s	Low	Medium	Global	Medium

Optimized networks are needed **for specific use case** characteristics!

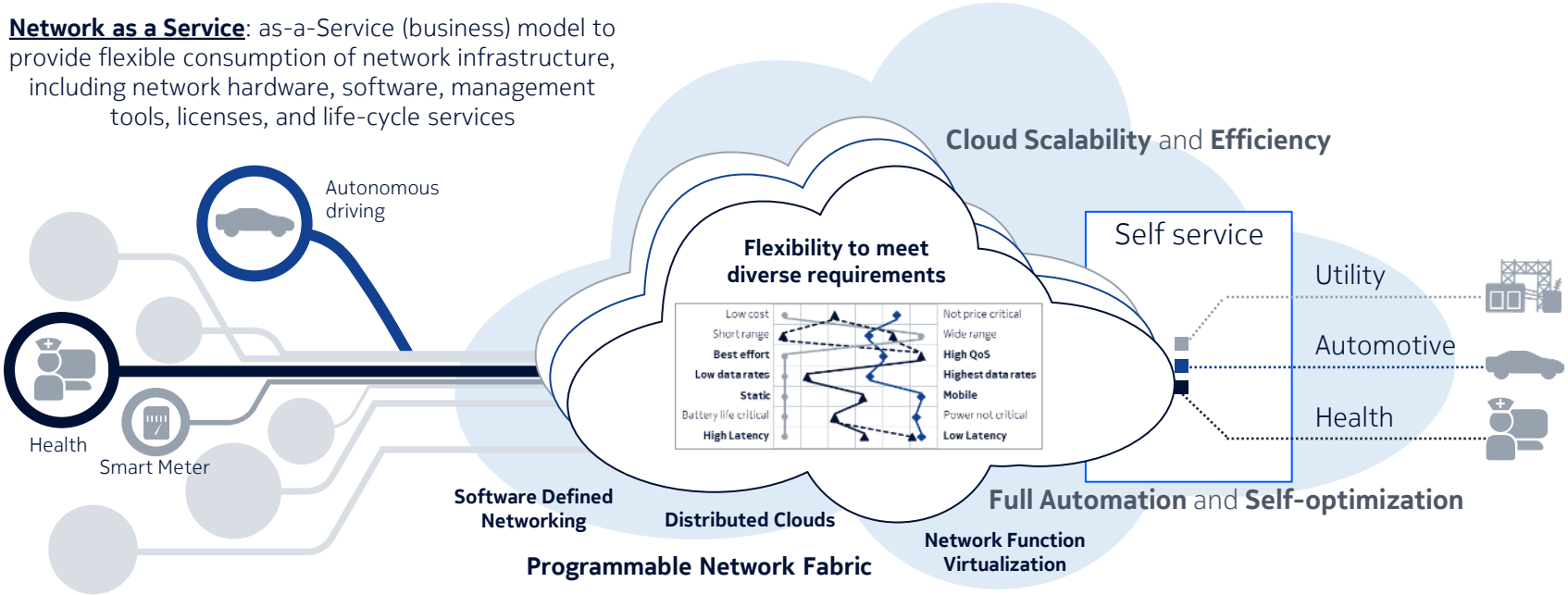
Optimization focus for use case

NOKIA BELL LABS

Concept of Network Slice

Running Multiple Logical Networks on a Common Physical Infrastructure

Network as a Service: as-a-Service (business) model to provide flexible consumption of network infrastructure, including network hardware, software, management tools, licenses, and life-cycle services



- "**Network Slice**" is a **logical network** that **provides specific network capabilities** and **network characteristics**, **supporting various service properties for network slice customers**
- Slicing across **access**, **transport**, **core**, **edge** and **central clouds**

NOKIA BELL LABS

5G Network Slice

Slice Type and Slice Identification

NSSAI (Network Slice Selection Assistance Information): A collection of S-NSSAIs, One UE can signal max 8 S-NSSAI in NSSAI

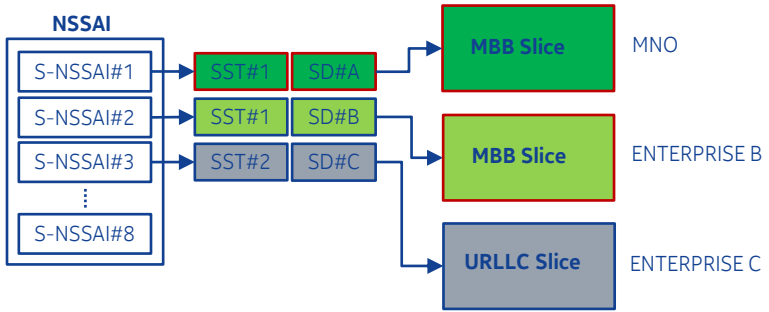
S-NSSAI (Single NSSAI): Identifies single Network Slice, consisting of:

- A **Slice/Service Type (SST: 8 bits):** defines slice characteristics, mandatory
- A **Slice Differentiator (SD: 24 bits):** identifies slices for more than one slice with the same SST, optional

- The **SST** field may have standardized and non-standardized values
- **SD** makes slice as **operator specific**
- **SD** is used to allow operator specific handling of slicing (e.g., different allowed areas, mobile vs FWA service, different changing in 5GC

TS 23.501 Table 5.15.2.2-1 - Standardised SST values

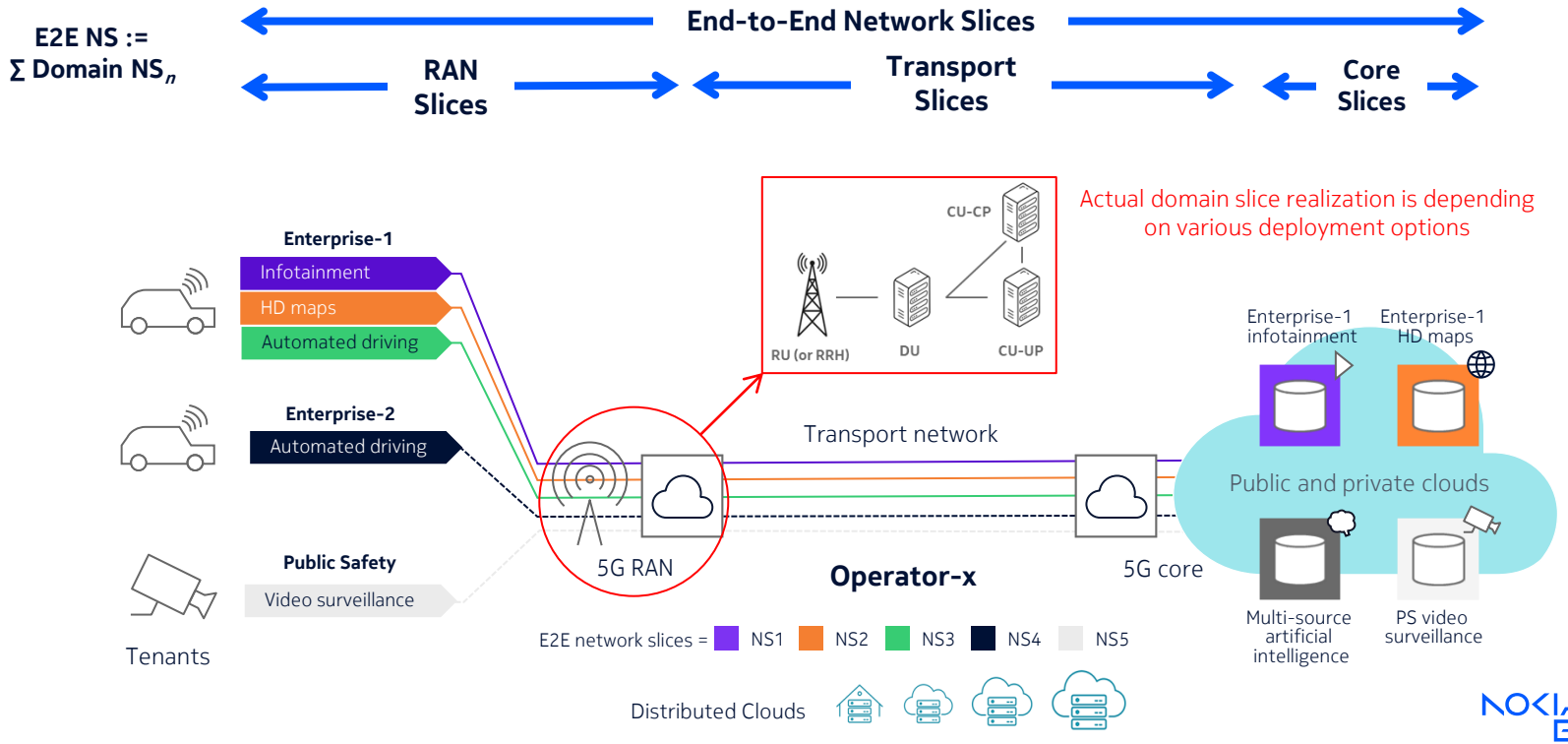
Slice/Service type	SST value	Characteristics.
eMBB	1	Slice suitable for the handling of 5G eMBB
URLLC	2	Slice suitable for the handling of URLLC
MIoT	3	Slice suitable for the handling of massive IoT
V2X	4	Slice suitable for the handling of V2X services
HMTC	5	Slice suitable for the handling of High-Performance Machine-Type Communications
HDLLC	6	Slice suitable for the handling of High Data rate and Low Latency Communications
Standardized	7…127	Standardized SST range (next standardized slices under definition in 3GPP)
Operator	128..255	Operator specific range



End-to-End Network Slice

Reference Architecture

RRH: Remote Radio Head
RU: Radio Unit
DU: Distributed Unit
CU: Central Unit
UP: User Plane
CP: Control Plane

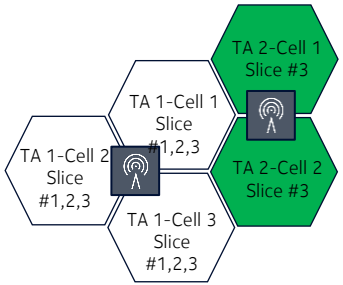


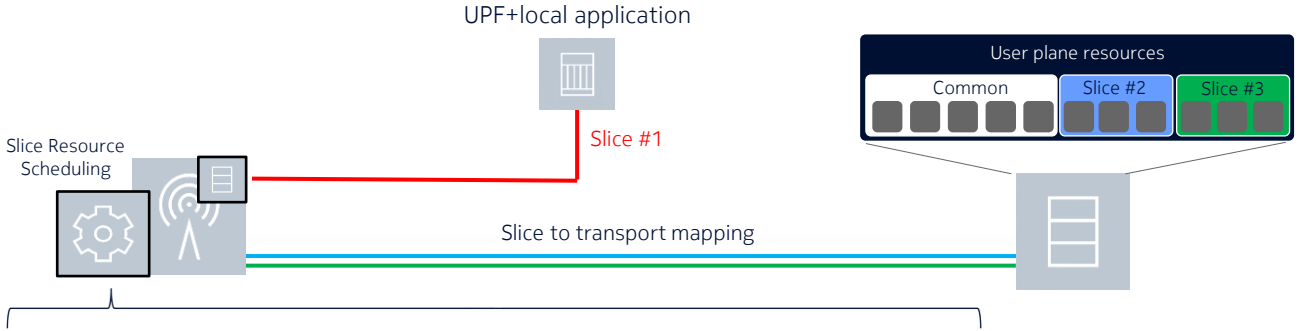
Domain Network Slice

RAN Slice (1/2)

How can RAN Slice be realized?

Geographical	Scheduling	Local Processing & Routing	Cloud Resources
Network Slice is configured per TA (Tracking Area)	Network Slice ID can be used for radio resource scheduling	Network Slice ID can be used for local processing / routing decision	Network Slice ID can be used for allocating dedicated UP resources

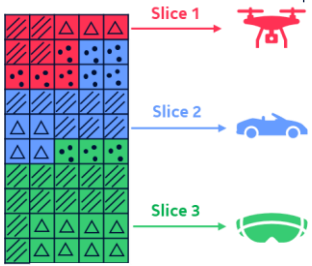




3GPP TS 28.541

The 3GPP Network Resource Model (NRM)

- 3 types of resources: Dedicated Resource (**DR**) △, Prioritized Resource (**PR**) ⁘, Shared Resource (**SR**) ///
- DR exclusively for assigned slice
- Unused PR can be shared across slices
- Sum of three types can be greater than total resources per gNB



Domain Network Slice

RAN Slice (2/2)

Layer	Example: eMBB Slice (Video streaming)	Example: eMTC Slice (Temperature Sensor)	Example: URLLC Slice (Smart Grid)
RRC	State handling optimized for RAN/CN signaling	Handover measurement omitted	Optimized for latency reduction
PDCP	Default	Ciphering omission, header compression	ciphering omission, header compression
RLC	Default	Unacknowledged Mode	Acknowledged Mode
MAC	Default	HARQ optimized for coverage	HARQ optimized for high reliability, RACH priority
PHY	Coding optimized for larger payload	Coding optimized for coverage, energy efficiency	Coding optimized for low latency, short payload

Example RAN configuration for different slices

NOKIA BELL LABS

Domain Network Slice

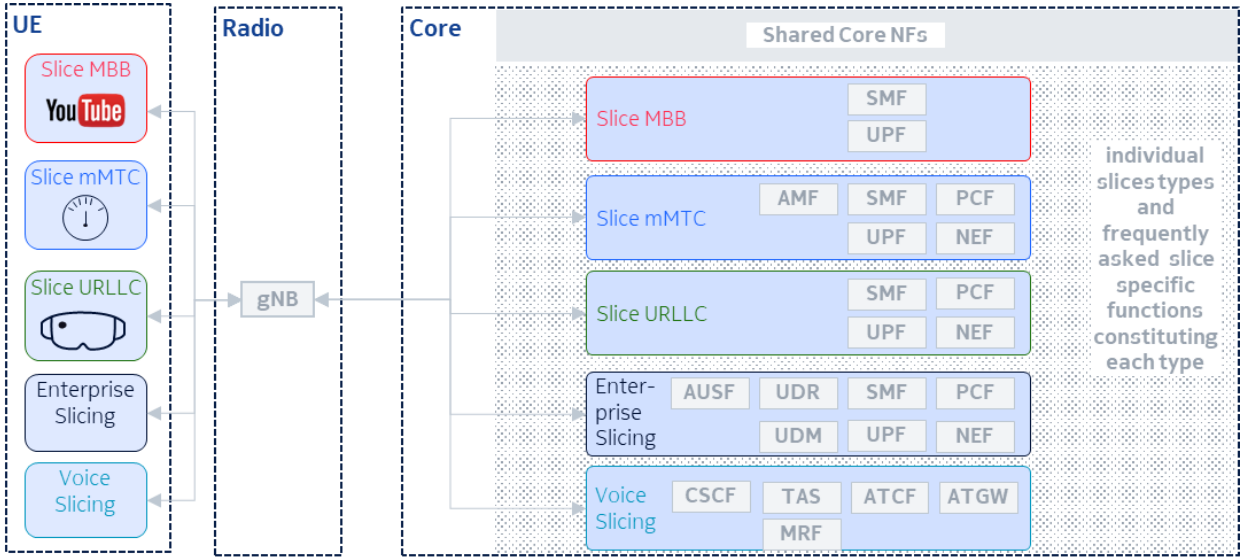
Core Slice

AMF: Access & Mobility Management Function
SMF: Session Management Function
UPF: User Plane Function
N3IWF: Non-3GPP Interworking Function
PCF: Policy Control Function
AUSF: Authentication Server Function
UDM: Unified Data Management
UDR: Unified Data Repository

UDSF: Unstructured Data Storage Function
NRF: Network Repository Function
NSSF: Network Slice Selection Function
AF: Application Function
NEF: Network Exposure Function
LMF: Location Management Function
SEPP: Security Edge Protection Proxy
NWDAF: Network Data Analytics Function

How can Core Slice be realized?

Core Slice s based on cloud installation
Shared/Dedicated Core network is created in cloud for a given slice
For example, 5G Core selection is based on PLMN and S-NSSAI



Diverse Core NFs:

AMF, SMF, UPF, N3IWF, PCF, AUSF, UDM, UDR, UDSF, NRF, NSSF, AF, NEF, LMF, SEPP, NWDAF, …

Private/Public Cloud:

On-premise, far-edge, edge, metro, central, …

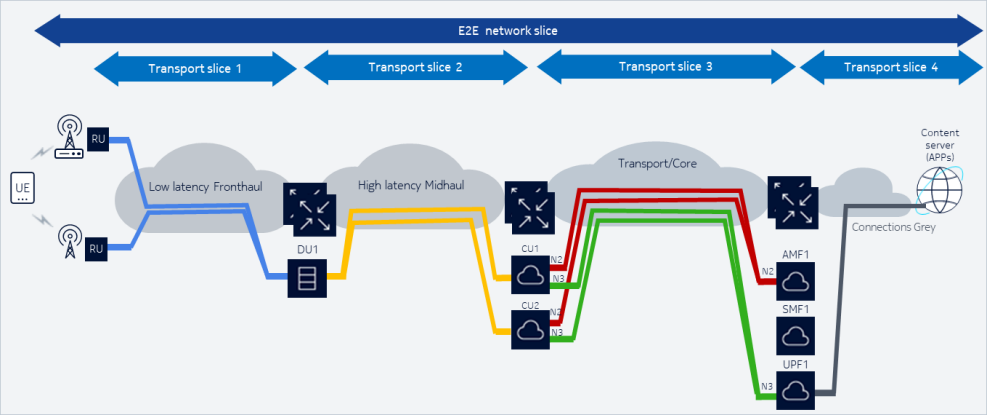
NOKIA BELL LABS

Domain Network Slice

Transport Slice

How can Transport Slice be realized?

A set of Transport Slices	Slice Isolation (soft vs. hard)	SLA, QoS	Standardization
Set of distinct connections Realized by any technology (IP, Optics, Microwave and PON) and at any layer L0/L1/L2/L3	VPN Tunnel Creation for dedicated transport resource provisioning - Any tunnel types (IP, MPLS, SR,…) SDN based solution	Deterministic SLA QoS Level Marking (e.g. DSCP marking, Queue Size, PIR/CIR etc)	Not handled by 3GPP Multiple standardization organization working on it: IETF, BBF, MEF



Multiple transport slices per E2E slice, each having multiple connections

From, 3GPP Slice NRM, ***EP_Transport*** instance includes additional attributes as follow:

- *ipAddress*
- *localLogicalInterfaceInfo*
- *qosProfile*
- *epApplicationRef*
- *externalEndPointRefList*

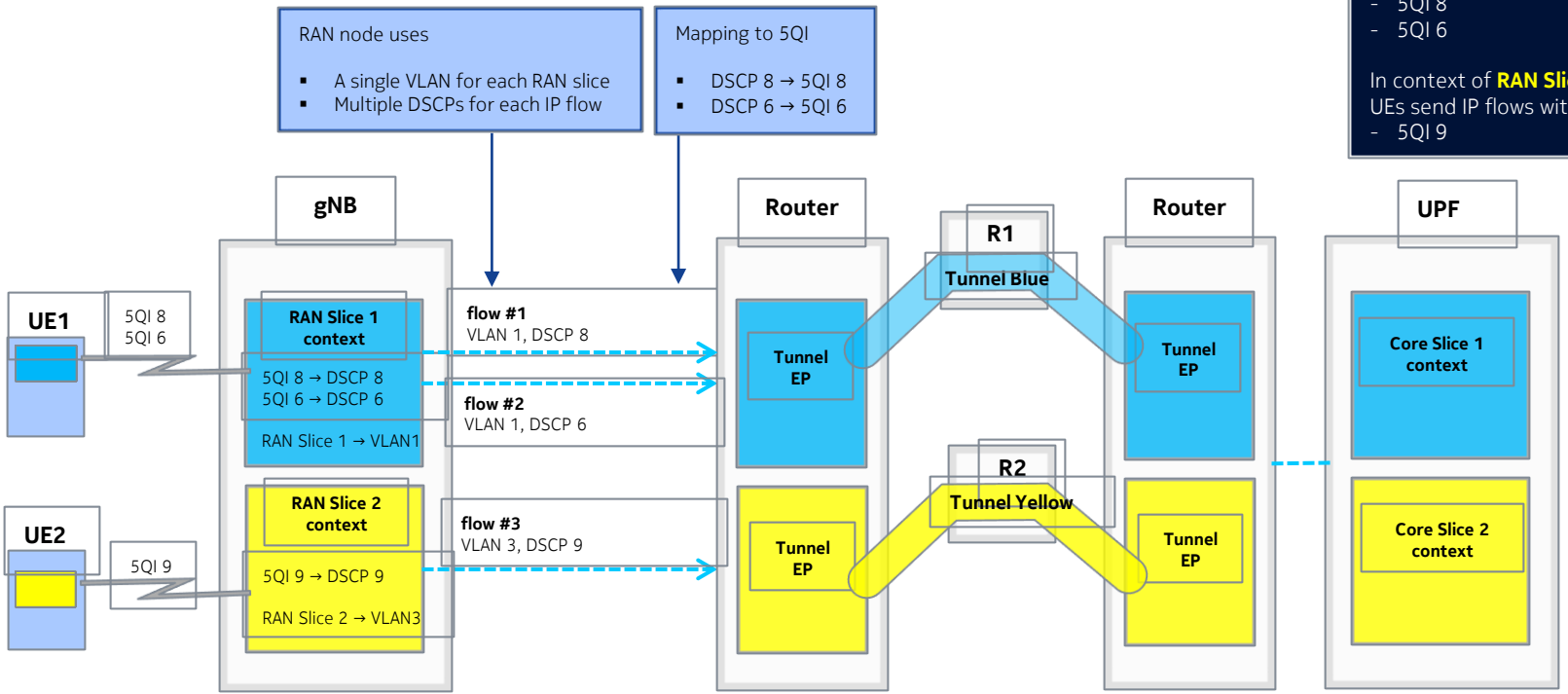
Domain Network Slice Stitching

Example of End-to-End Network Slice Realization

RAN node has context for two RAN slice:
- RAN Slice #1 (i.e. S-NSSAI 1)
- RAN Slice #2 (i.e. S-NSSAI 2)

In context of **RAN Slice #1**, UEs send IP flows with:
- 5QI 8
- 5QI 6

In context of **RAN Slice #2**, UEs send IP flows with:
- 5QI 9



Example of how to stitch Transport Slice with RAN/Core Slices based on VLAN

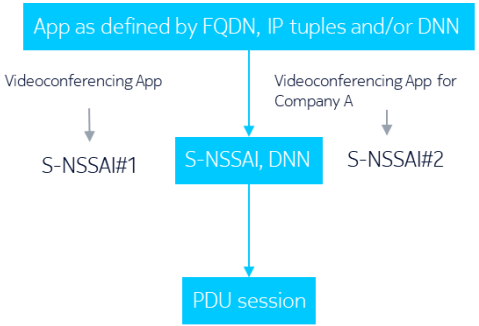
NOKIA BELL LABS

Domain Network Slice

UE Slice

URSP (**UE Route Selection Policy**) is **dynamic slice selection technology** enabling devices to utilize *multiple slices in parallel*

- URSP contains multiple URSP rules providing **a dynamic matching of traffic to PDU sessions**
  - Identification of traffic based on **Traffic Descriptor** (**TD**)
  - Mapping to existing PDU session or creation of a new one based on **Route Selection Descriptor** (**RSD**)
- URSP requires **5G UE**, **5G Core** and **Policy Control** (SA)
- The encoding of URSP is defined in **3GPP TS 24.526**
- UE receives the updated URSP rules and (re-)evaluates their validities
  - URSP is updated by PCF
  - UE moved from EPC to 5GC
  - Change of Allowed NSSAI or Configured NSSAI
  - …



URSP Policy

Rule precedence: 1
Traffic descriptor: Application identifier = OS ID + OS APP ID (Android OS ID + "**ENTERPRISE**")
Route precedence 1: sst: 1, sd: **1A2B3C**, DNN selection "**enterprise**"

Rule precedence: 100
Traffic descriptor: Match-all-type
Route precedence 1: sst:1 sd:**123456**, DNN selection "**internet**"

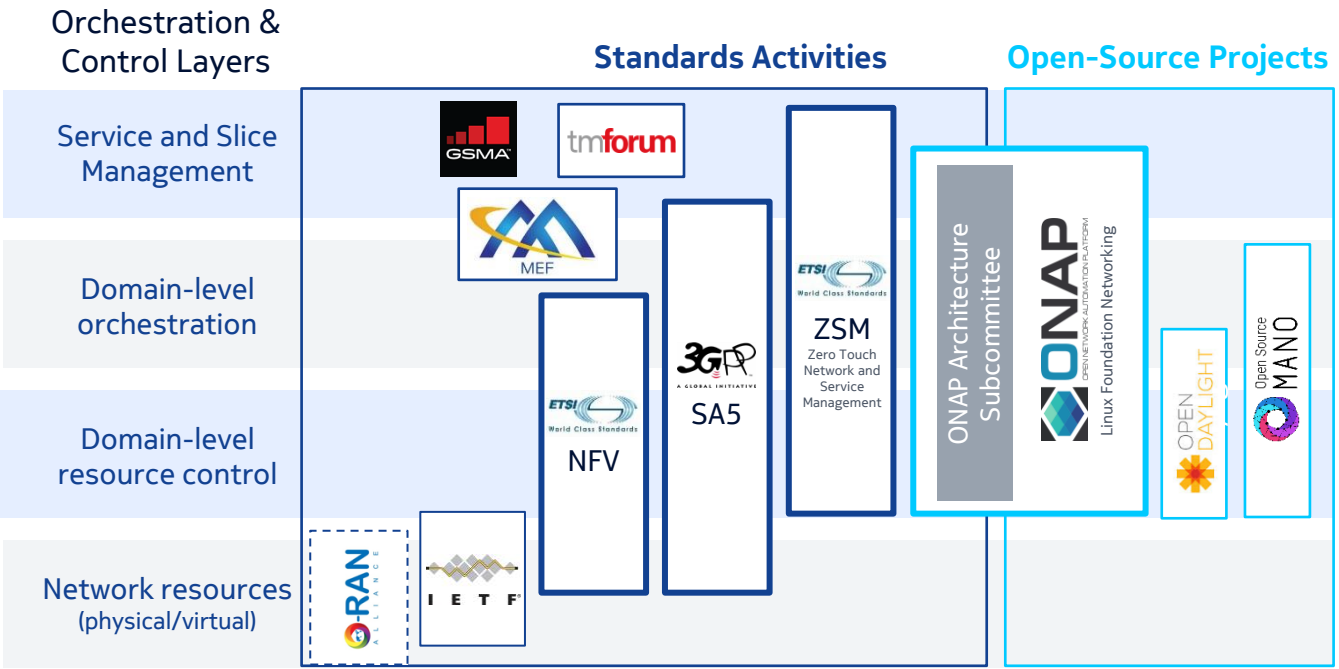
Traffic descriptor parameter name	Description
Application descriptors	It consists of OSId and OSAppId(s)
IP descriptors	Destination IP 3 tuple(s) (IP address or IPv6 network prefix, port number, protocol ID of the protocol above IP)
Domain descriptors	Destination FQDN(s) or a regular expression as a domain name matching criteria
DNN	This is matched against the DNN information provided by the application
Connection Capabilities	This is matched against the information provided by a UE application when it requests a network connection with certain capabilities

Route selection descriptor parameter name	Description
Route selection descriptor precedence	Determines the order in which the route selection descriptors are to be applied
SSC Mode Selection	One single value of SSC mode
Network Slice Selection	Either a single value or a list of values of S-NSSAI(s)
DNN Selection	Either a single value or a list of values of DNN(s)
PDU Session Type Selection	One single value of PDU Session Type
Non-Seamless Offload indication	Indicates if the traffic of the matching application is to be offloaded to non-3GPP access outside of a PDU session
Access Type preference	Indicates the preferred Access Type (3GPP or non-3GPP) when the UE establishes a PDU Session for the matching application

Part II. Network Slicing Orchestration

Network and Service Orchestration

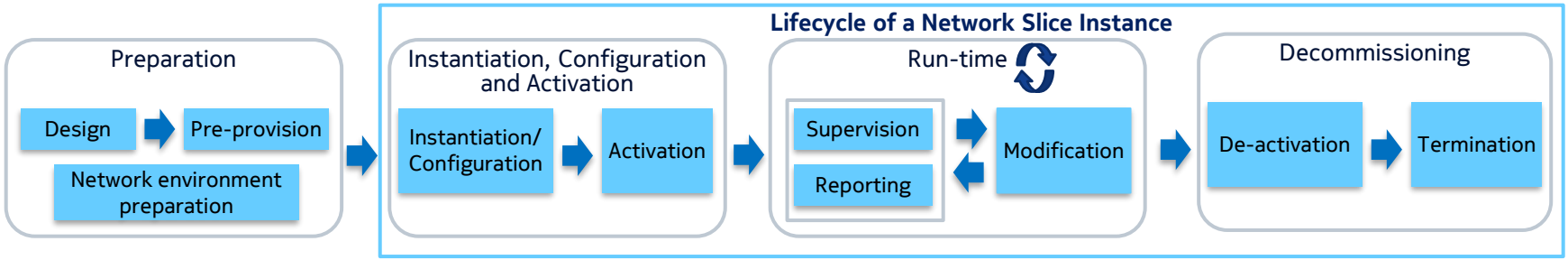
Key Standards and Open-Source Projects



NOKIA BELL LABS

Slice Orchestration and Automation

TS 28.801 – Mgmt & Orch of Network Slicing for Next-Generation Network



Preparation phase

In the preparation phase the NSI does not exist. The preparation phase includes the **creation and verification of network slice template(s)**, the **onboarding of these, preparing the necessary network environment** which are used to support the lifecycle of NSIs and any other preparations that are needed in the network.

Instantiation, Configuration and Activation phase

During instantiation/configuration all resources shared/dedicated to **the NSI have been created and are configured**, i.e. to a state where the NSI is ready for operation. **The activation** step includes any actions that makes the NSI active, e.g. diverting traffic to it, provisioning databases (if dedicated to the network slice, otherwise this takes place in the preparation phase) etc. NS instantiation, configuration and activation can include instantiation, configuration and activation of other shared and/or non-shared NF(s).

Run-time phase

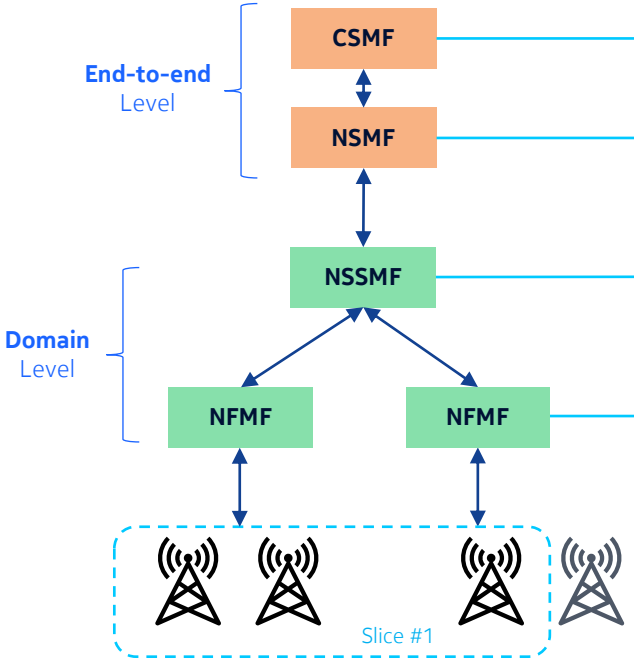
In the run-time phase the NSI is capable of traffic handling to support communication services of certain type(s). The run-time phase includes **supervision/reporting (e.g., for KPI monitoring),** as well as activities related to modification. **Modification** could map to several workflows related to runtime tasks, e.g., **upgrade, reconfiguration, NSI scaling, changes of NSI capacity**, changes of NSI topology, association and disassociation of network functions with NSI.

Decommissioning phase

The decommissioning phase includes deactivation (taking the NSI out of active duty) as well as the reclamation of dedicated resources (e.g., termination or re-use of network functions) and configuration of shared/dependent resources. After decommissioning the NSI does not exist anymore.

Slice Orchestration and Automation

TS 28.801 – Mgmt & Orch of Network Slicing for Next-Generation Network



CSMF translates communication requirements to network slices

Network slice instance (NSI) monitoring, reporting and LCM
- Slice level health monitoring and SLA assurance
- Slice life cycle management and closed-loop operations

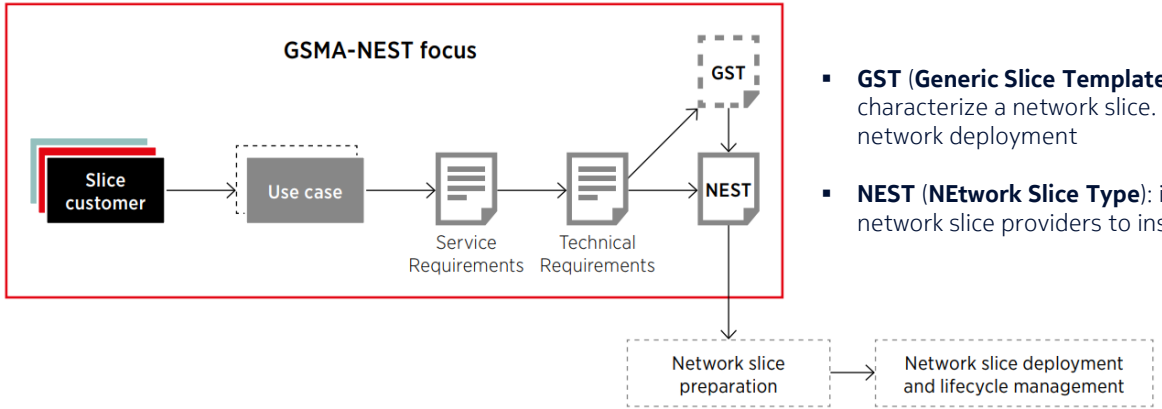
Network slice subnet instance (NSSI) monitoring, reporting and LCM
- Alarm correlation and statistics aggregation on slice subnet level
- Network slice subnet (NSSI) life cycle management and provisioning according to slice profile

Network function (NF) monitoring, reporting and configuration
- Raw slicing alarm/counter collection on NF level
- NF slicing parameter configuration and provisioning

CSMF	Communication Service Management Function
NSMF	Network Slice Management Function
NSSMF	Network Slice Subnet Management Function
NFMF	Network Function Management Function

Slice Orchestration and Automation

GSMA NG. 116: Generic Network Slice Template



- **GST** (**Generic Slice Template**): A set of potential slice attributes that can characterize a network slice. GST is generic and is not tied to any specific network deployment
- **NEST** (**NEtwork Slice Type**): is a GST filled with values and is essential for network slice providers to instantiate equivalent slices

[GST Attributes]

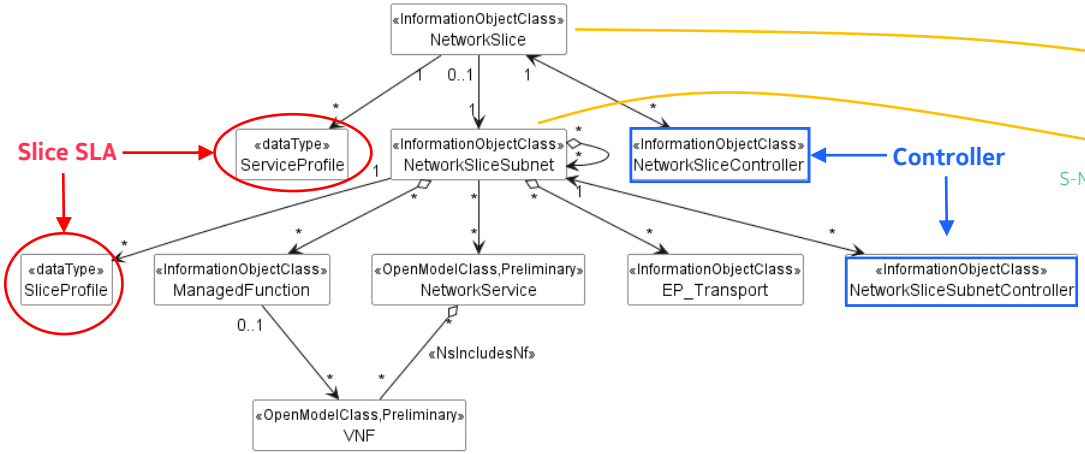
Availability, area of service, **delay tolerance**, deterministic communication, **downlink throughput per network slice**, **downlink maximum throughput per UE**, **energy efficiency**, group communication support, **isolation level**, maximum supported packet size, mission critical support, **maximum number of PDU sessions**, **maximum number of UEs**, performance monitoring, performance prediction, positioning support, radio spectrum, Session and Service Continuity (SSC) support, simultaneous use of the network slice, **slice quality of service**, supported device velocity, synchronicity, UE density, **uplink throughput per network slice**, **uplink maximum throughput per UE**, **packet delay budget**, ...

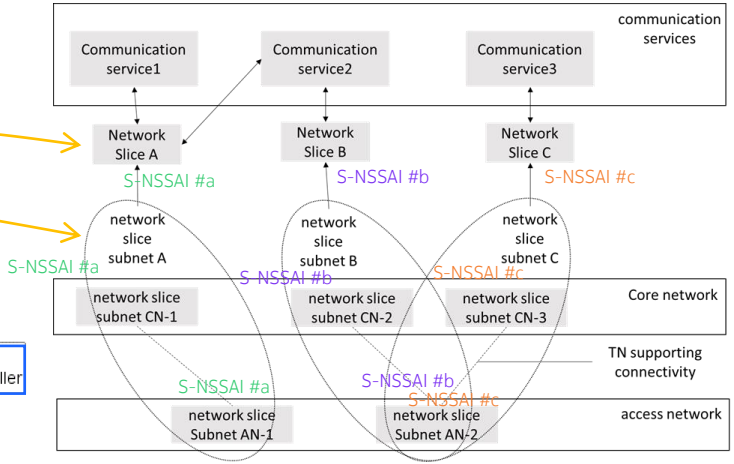
[NEST for URLLC]

Attribute		Value
Availability		99.999
Session and Service Continuity Support		1
Slice quality of service	3GPP 5QI	82
Supported device velocity		2

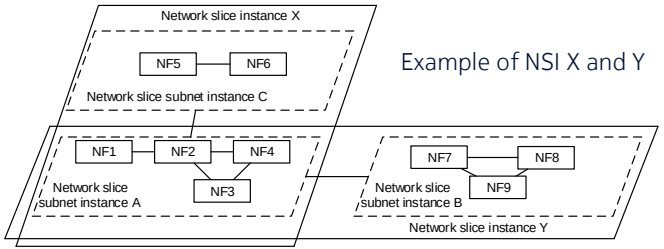
Slice Orchestration and Automation

3GPP TS 28.541 - 5G Network Resource Model (NRM)





A **service instance** is realized by one or more **network slice instance**s (**NSI**s), that in turn consists of **network slice subnet instance**s (**NSSIs**)



Example of NSI X and Y

Slice Orchestration and Automation

3GPP TS 28.541 - 5G Network Resource Model (NRM)

Attribute name	Support Qualifier
serviceProfileId	M
pLMNInfoList	M
maxNumberofUEs	O
coverageArea	O
dLLatency	O
uLLatency	O
uEMobilityLevel	O
networkSliceSharingIndicator	O
sST	M
availability	O
delayTolerance	O
dLDeterministicComm	O
uLDeterministicComm	O
dLThptPerSlice	O
dLThptPerUE	O
uLThptPerSlice	O
uLThptPerUE	O
dLMaxPktSize	O
uLMaxPktSize	O
maxNumberofPDUSessions	O
kPIMonitoring	O
userMgmtOpen	O
v2XCommModels	O
termDensity	O
activityFactor	O
uESpeed	O
jitter	O
survivalTime	O
radioSpectrum	O
dLReliability	O
uLReliability	O
maxDLDataVolume	O
maxULDataVolume	O
nBIoT	O
synchronicity	O

ServiceProfile

Information model definitions for network slice NRM

Attribute name	Support Qualifier
sliceProfileId	M
pLMNInfoList	M
CNSliceSubnetProfile	CM
RANSliceSubnetProfile	CM
TopSliceSubnetProfile	CM

SliceProfile

Slice Orchestration and Automation

3GPP TS 28.541 - 5G Network Resource Model (NRM)

Attribute name	Support Qualifier
dLLatency	O
uLLatency	O
maxNumberofUEs	O
dLThptPerSliceSubnet	O
dLThptPerUE	O
uLThptPerSliceSubnet	O
uLThptPerUE	O
dLMaxPktSize	O
uLMaxPktSize	O
maxNumberOfPDUSessions	O
nROperatingBands	O
sliceSimultaneousUse	O
delayTolerance	O
energyEfficiency	O
termDensity	O
activityFactor	O
coverageAreaTAList	O
resourceSharingLevel	O
uEMobilityLevel	O
uESpeed	O
dLReliability	O
uLReliability	O
dLDeterministicComm	O
uLDeterministicComm	O
survivalTime	O
positioning	O
synchronicity	O
nssaaSupport	O
n6Protection	O
nonIPSupport	O

TopSliceSubnetProfile

Attribute name	Support Qualifier
coverageAreaTAList	O
uEMobilityLevel	O
resourceSharingLevel	O
maxNumberofUEs	O
activityFactor	O
dLThptPerSliceSubnet	O
dLThptPerUE	O
uLThptPerSliceSubnet	O
uLThptPerUE	O
uESpeed	O
dLReliability	O
uLReliability	O
nROperatingBands	O
dLLatency	O
uLLatency	O
delayTolerance	O
sliceSimultaneousUse	O
dLMaxPktSize	O
uLMaxPktSize	O
energyEfficiency	O
termDensity	O
survivalTime	O
dLDeterministicComm	O
uLDeterministicComm	O
positioning	O
synchronicity	O

RANSliceSubnetProfile

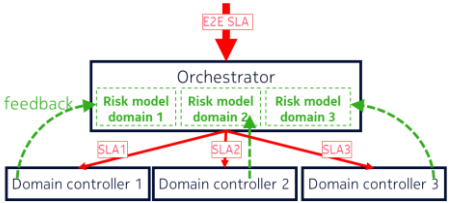
Attribute name	Support Qualifier
maxNumberofUEs	O
dLLatency	O
uLLatency	O
dLThptPerSliceSubnet	O
dLThptPerUE	O
uLThptPerSliceSubnet	O
uLThptPerUE	O
maxNumberOfPDUSessions	O
coverageAreaTAList	O
dLReliability	O
uLReliability	O
resourceSharingLevel	O
dLMaxPktSize	O
uLMaxPktSize	O
sliceSimultaneousUse	O
delayTolerance	O
energyEfficiency	O
dLDeterministicComm	O
uLDeterministicComm	O
survivalTime	O
nssaaSupport	O
n6Protection	O
nonIPSupport	O

CNSliceSubnetProfile

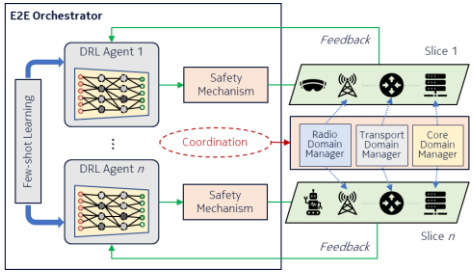
Slice Orchestration and Automation

E2E Slice Deployment & Runtime SLA Management

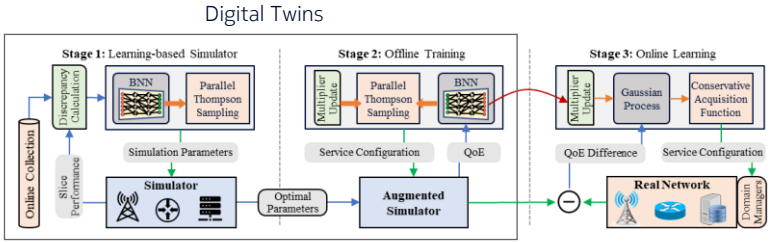
[1] "Applying Machine Learning to End-to-end Slice SLA Decomposition," IEEE NetSoft S4SI Workshop 2020
[2] "Decomposing SLAs for Network Slicing," IEEE Communications Letters 2020
[3] "Constraint-Aware Deep Reinforcement Learning for End-to-End Resource Orchestration in Mobile Networks," IEEE ICNP 2021
[4] "OnSlicing: online end-to-end network slicing with reinforcement learning," ACM CoNEXT 2021
[5] "Atlas: Automate Online Service Configuration in Network Slicing," ACM CoNEXT 2022
[6] "RoNet: Toward Robust Neural Assisted Mobile Network Configuration," IEEE ICC 2023



E2E SLA Decomposition[1,2]

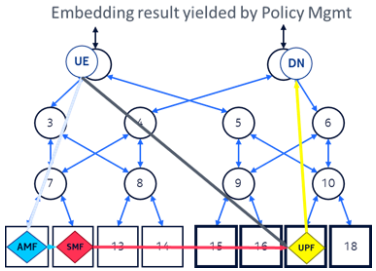


DRL-based Runtime SLA Assurance[3,4]

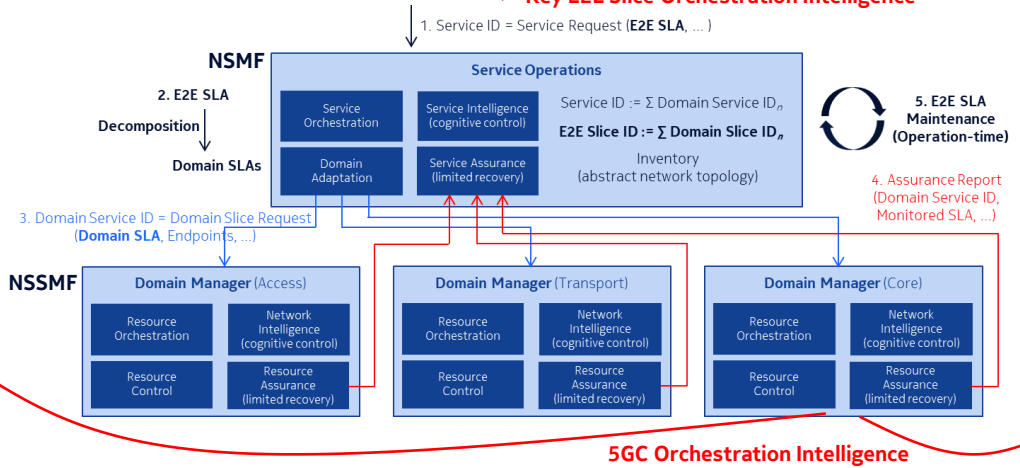


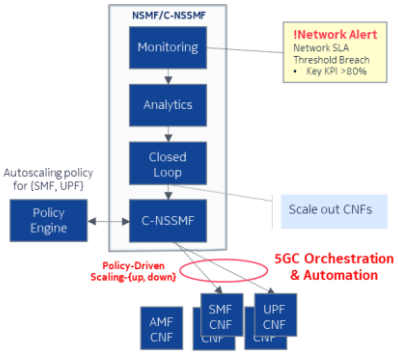
Digital Twins for E2E Orchestration[5,6]

Key E2E Slice Orchestration Intelligence



Secure Slice Placement

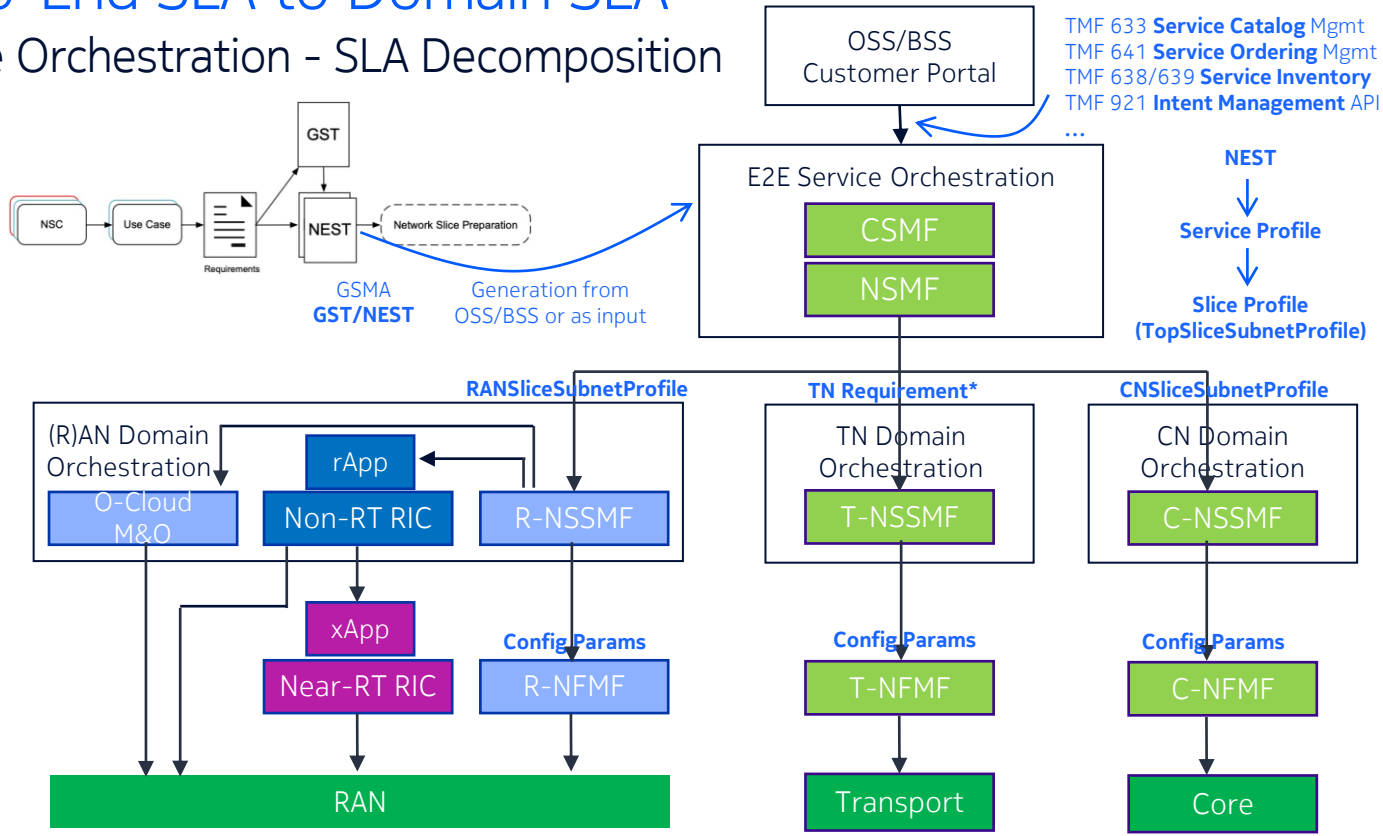




5GC Auto-scaling Close Loop

End-to-End SLA to Domain SLA

E2E Slice Orchestration - SLA Decomposition



*** IETF Yang model**: https://datatracker.ietf.org/doc/draft-ietf-teas-ietf-network-slice-nbi-yang/)

NOKIA BELL LABS

Concluding Remarks

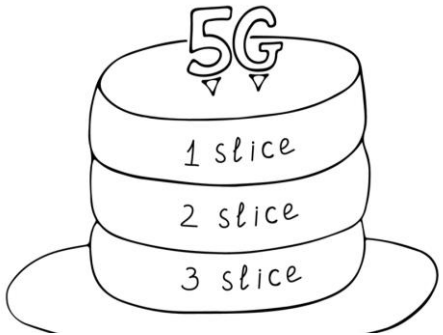
NOKIA BELL LABS

Evolving Network Slicing Capabilities

Further Enhancement

3GPP Rel. 18 Network Slicing enhancements

- Better support for highly customized and localized service
- Dynamic network slices for temporary deployments
- Enhanced operator control for slice quota enforcement
- Enhanced operator control of when UEs can connect and use a particular network slice
- Service continuity in congestion and network maintenance
- Enhanced network slice management support



Considerations for further enhancements

- Huge number of slices, e.g., 2k
- Highly scalable and dynamic slice resource management
- Slice security, e.g., Network Slice Specific Authentication and Authorization (NSSAA)

NOKIA
BELL
LABS