T01: RAN Slicing Challenges, Technologies, and Tools

Navid Nikaein

Assistant Professor - Eurecom

Coordiantor of Mosaic-5G Initiative

Tutorial IEEE/IFIP Network Operations and Management Symposium

23-27 April 2018. Taipei, Taiwan

Provide a comprehensive guide on RAN slicing and orchestration

(1) highlight the importance and timeliness of **softwarization, virtualization, and disaggregation** of RAN to enable **multiservice multi-tenant** RAN toward So-RAN architecture

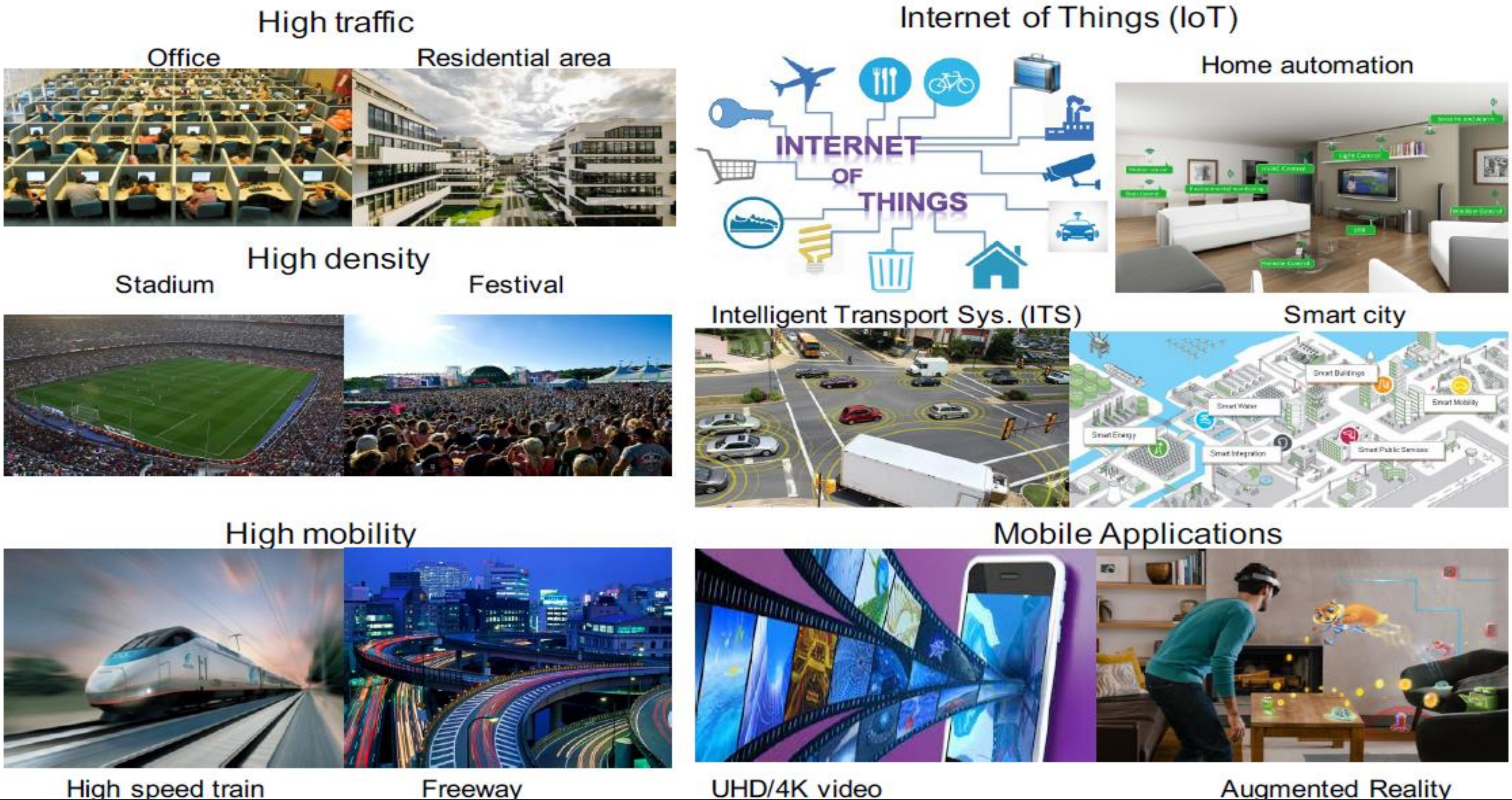
(2) Discuss the interconnection of a RAN slice with core Network Slices to enable an **E2E slice in 5G**

(3) Cover a **well-balanced research and development** topics including challenges, key technologies, and proof-of-concept prototyping

Tutorial Objectives


Connected, Controlled, and Flexible
Digital Society
Value Creation
Consistent experience
Sustainable business model

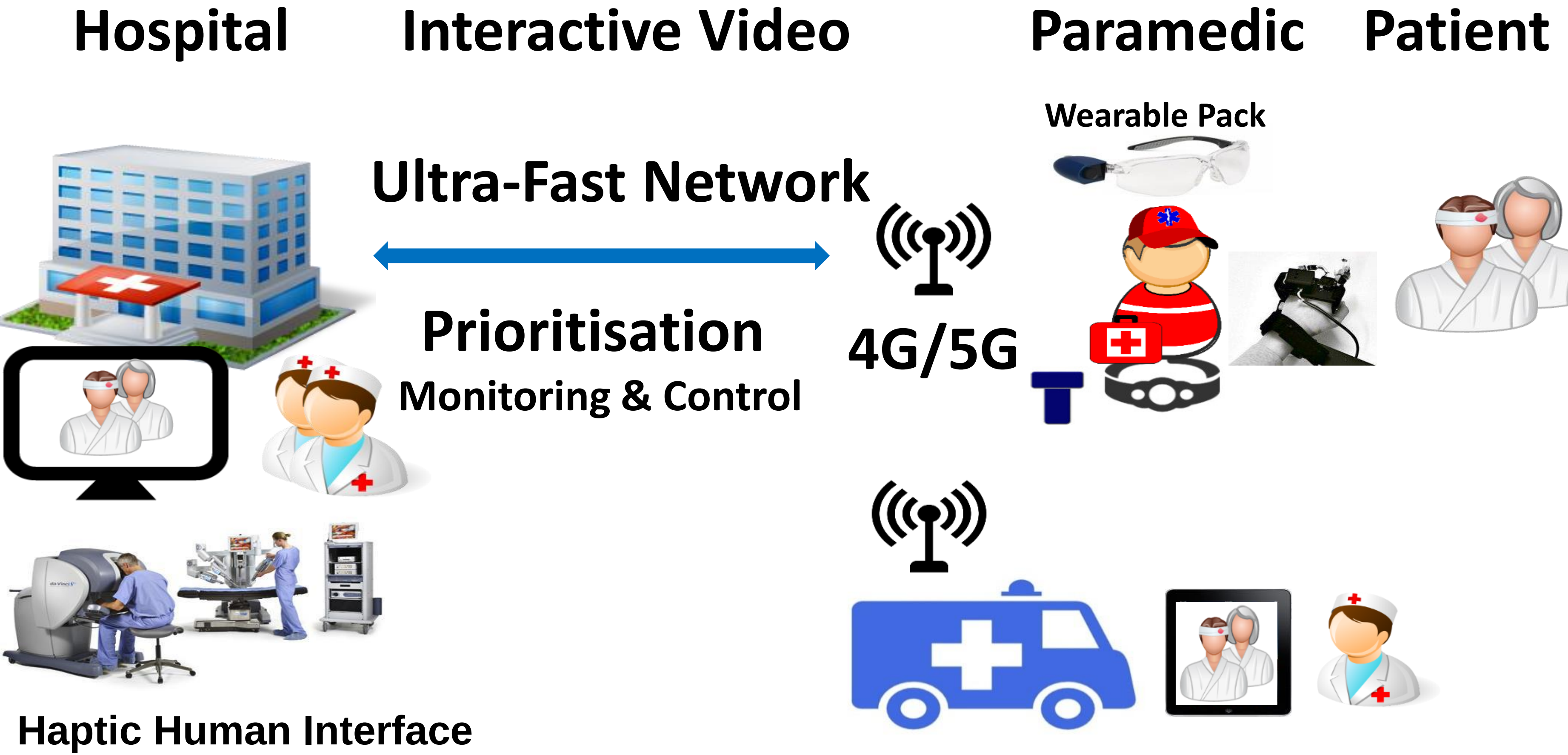
What is 5G?



Many 5G Use-cases

Source: Dohler

Internet of Skills

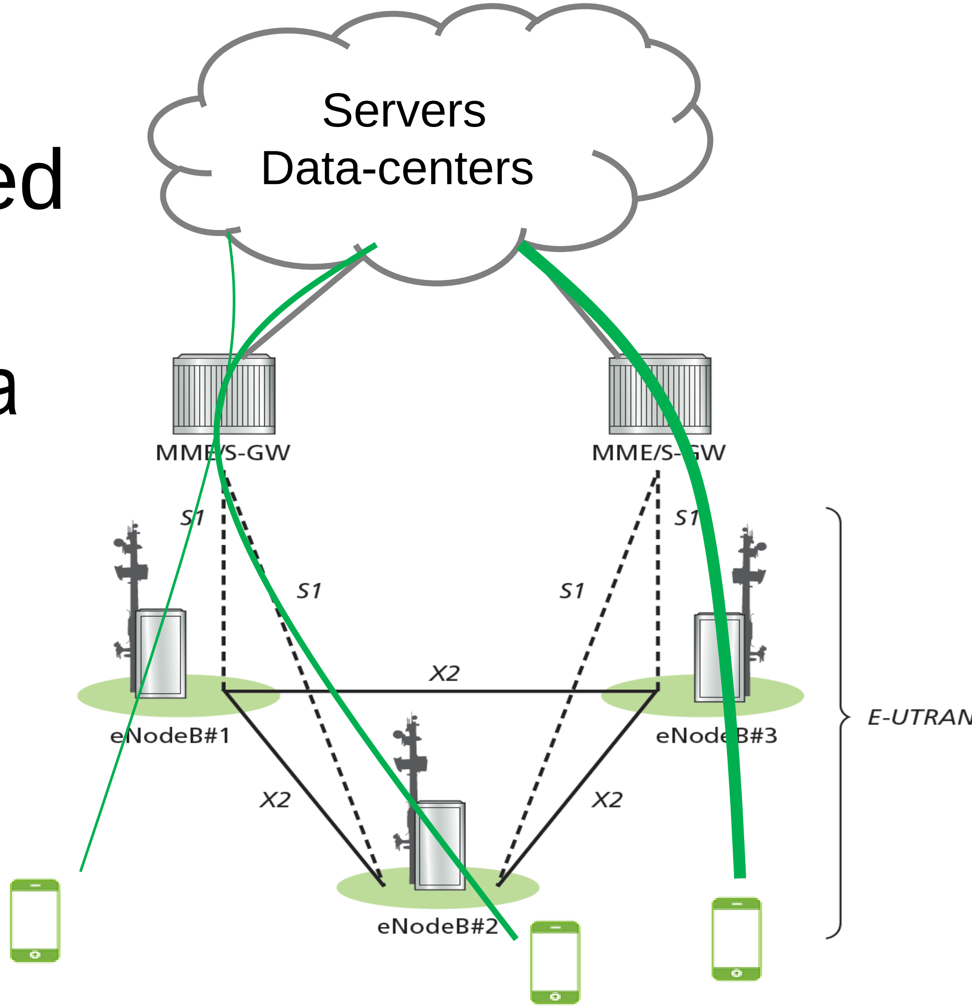


E-Health: When Robotics meets 5G

Communication-oriented

Today's 4G is designed for a limited number of UCs

- Throughput-optimized
- Fixed
- Rigid

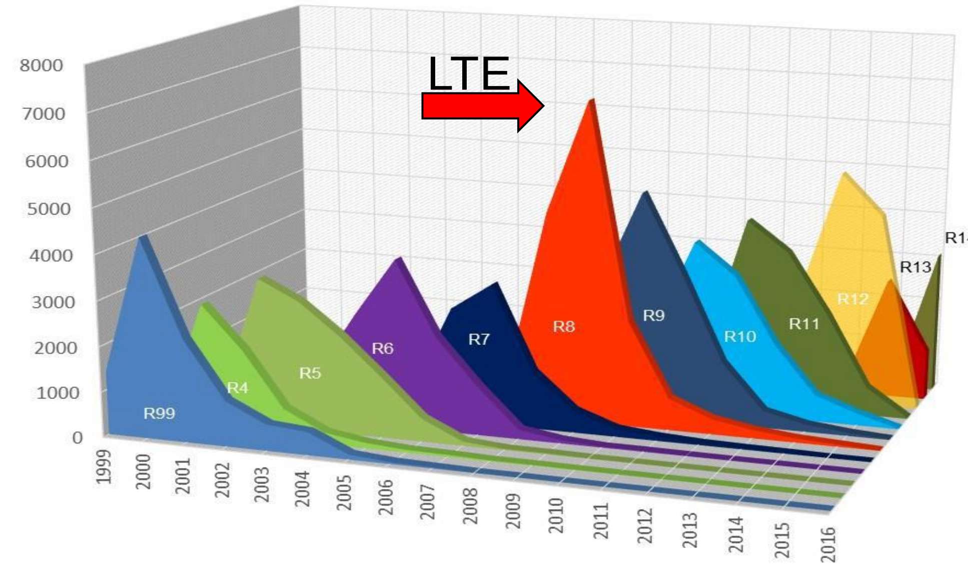


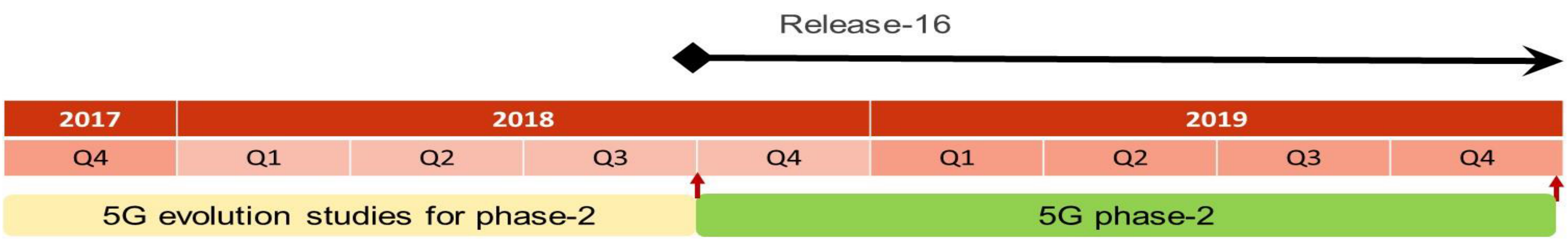
Is 4G enough?

Mindful about

3GPPP facts and figures

514 Companies from 45 Countries
50,000 delegate days per year
40,000 meeting documents per year
1,200 specifications per Release
10,000 change requests per year





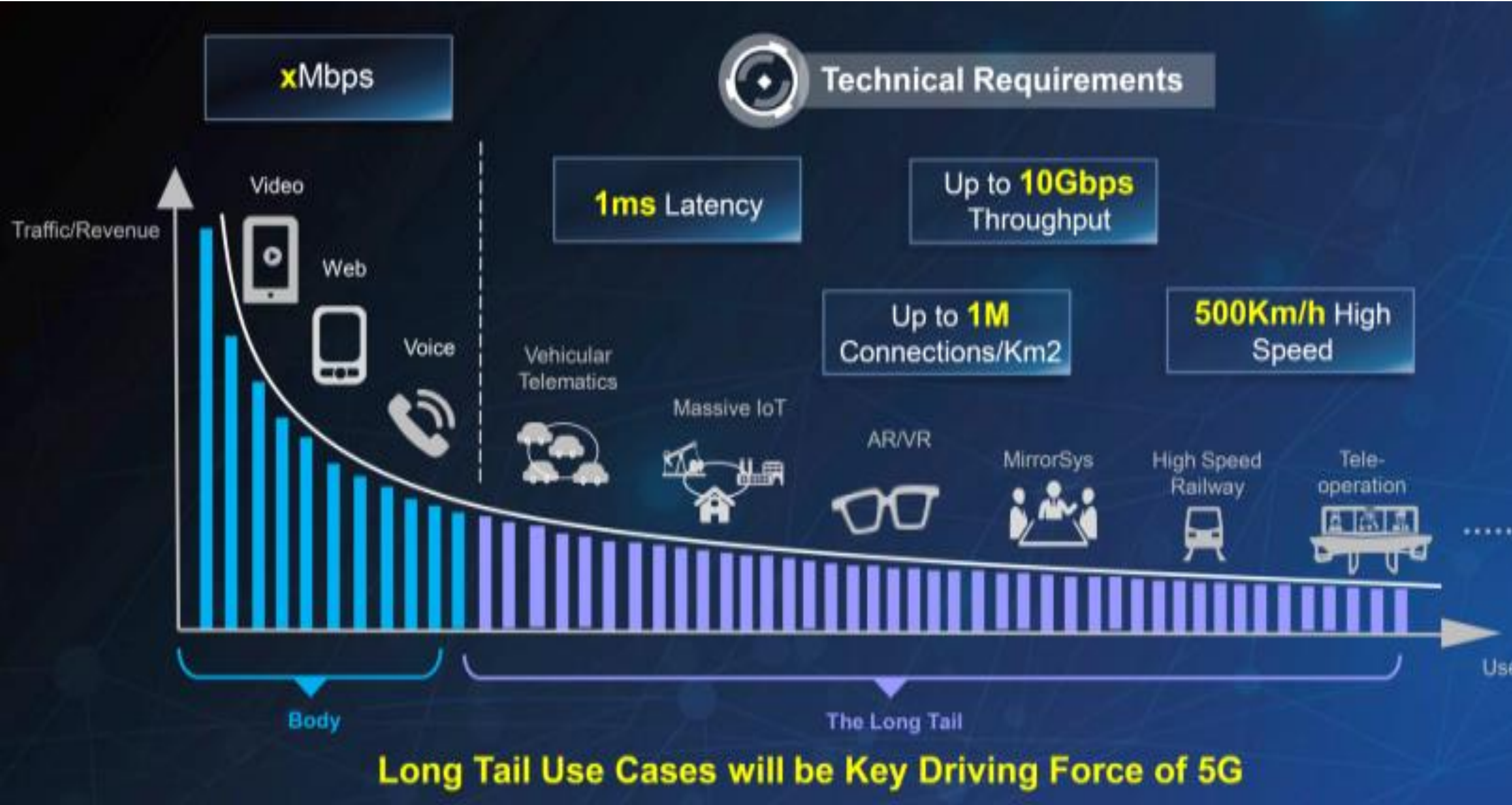
Communication-oriented 4G

Future mobile network will look fundamentally different

There will be no “one-size-fits-all” architecture

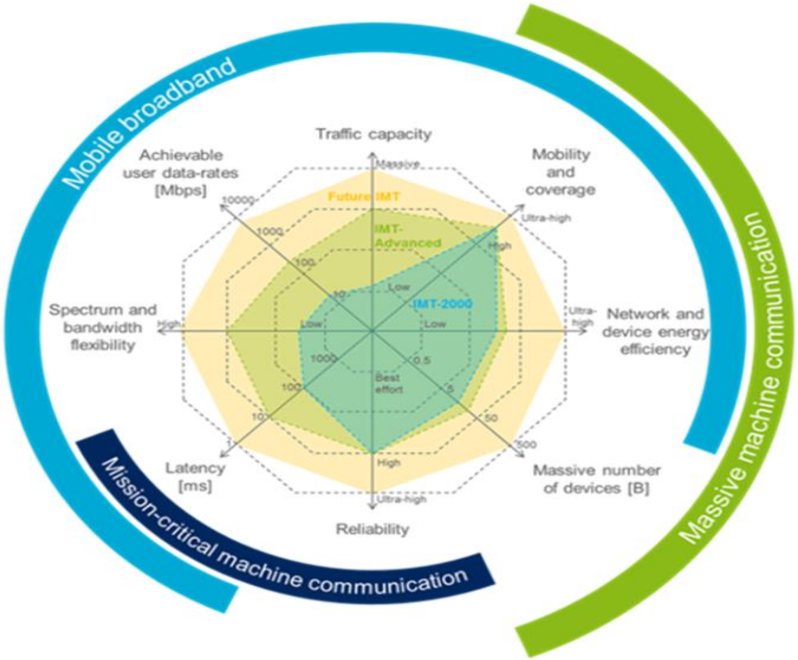
Like it or not

Service-oriented 5G



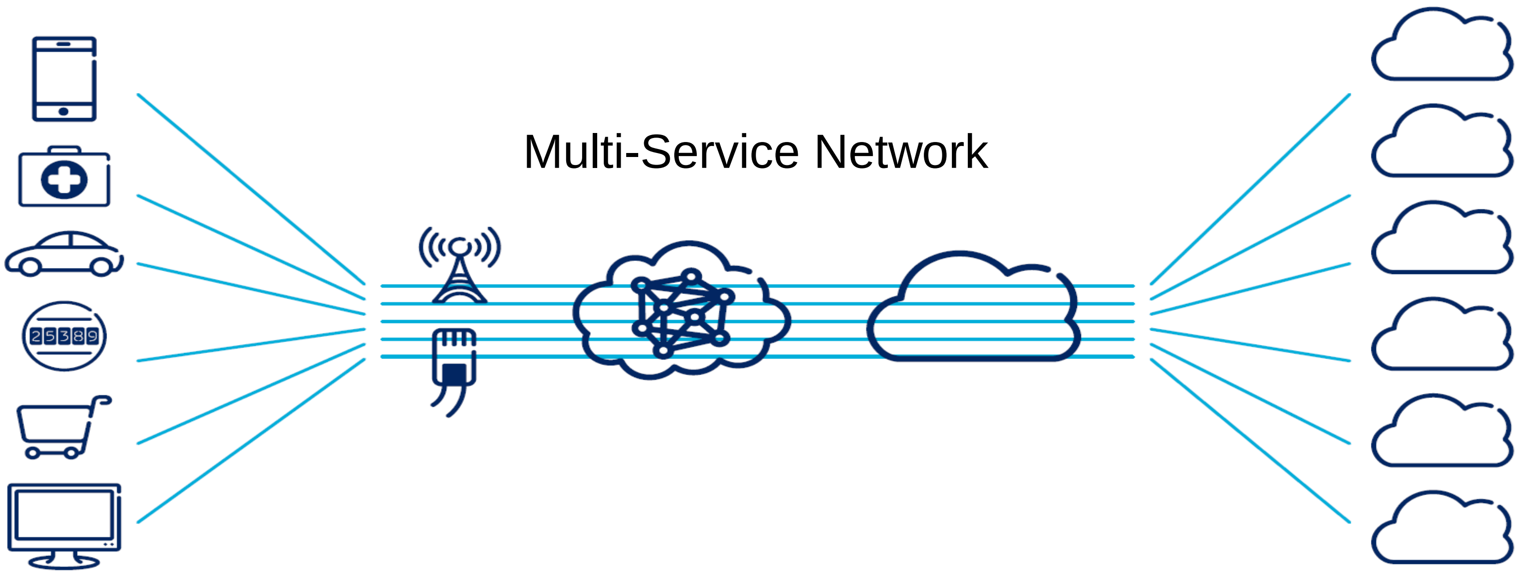
Not a one-size fits all

Building a different network is also not a viable option



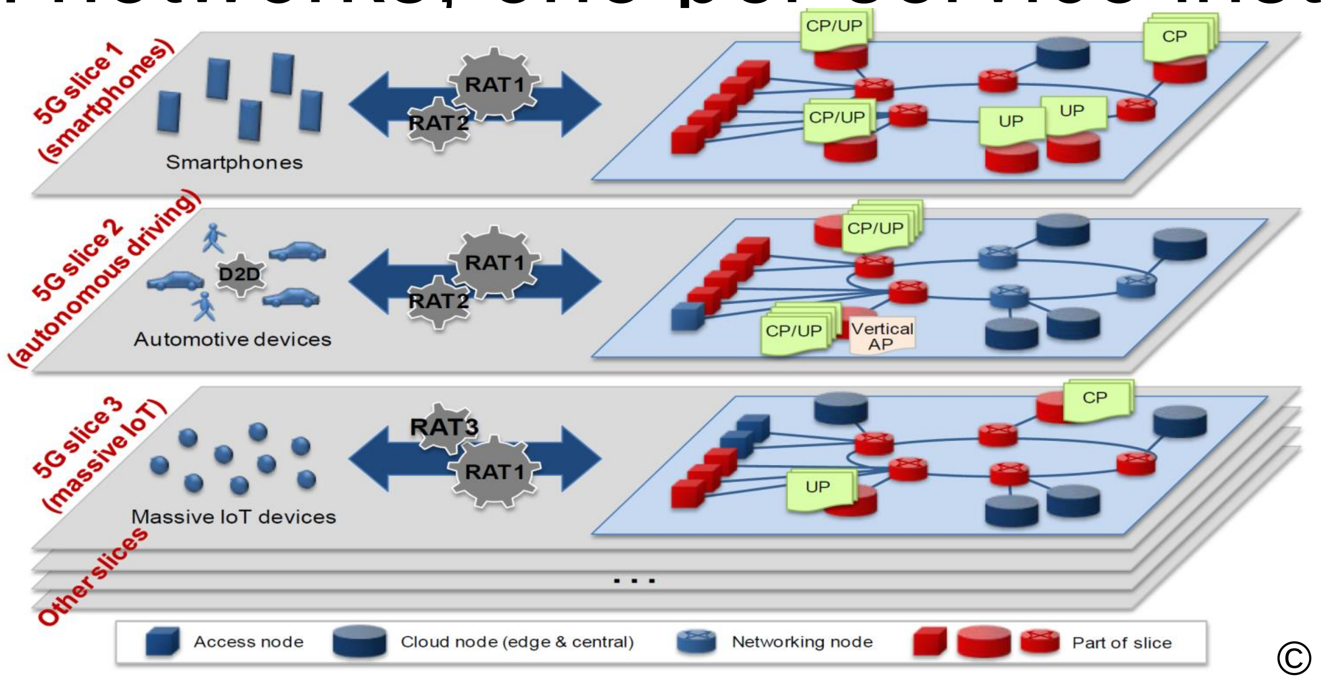
Not a dedicated-network

Turn physical infrastructure into multiple logical networks, one per service instance

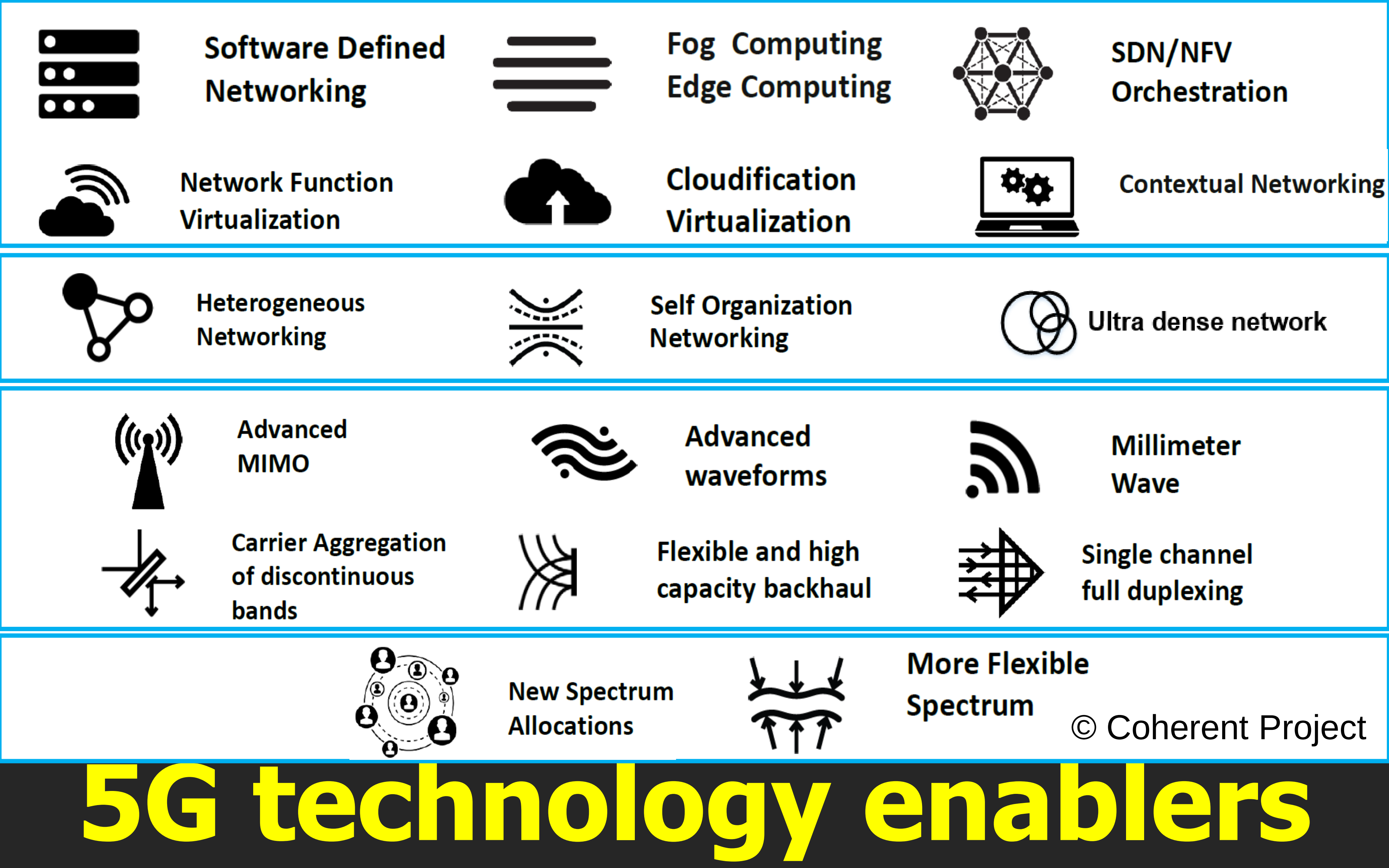


One-Network, Many-Service

Turn physical infrastructure into multiple logical networks, one per service instance

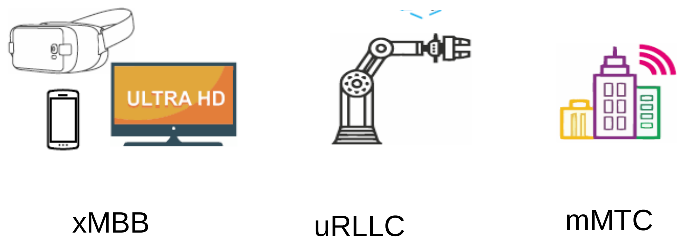


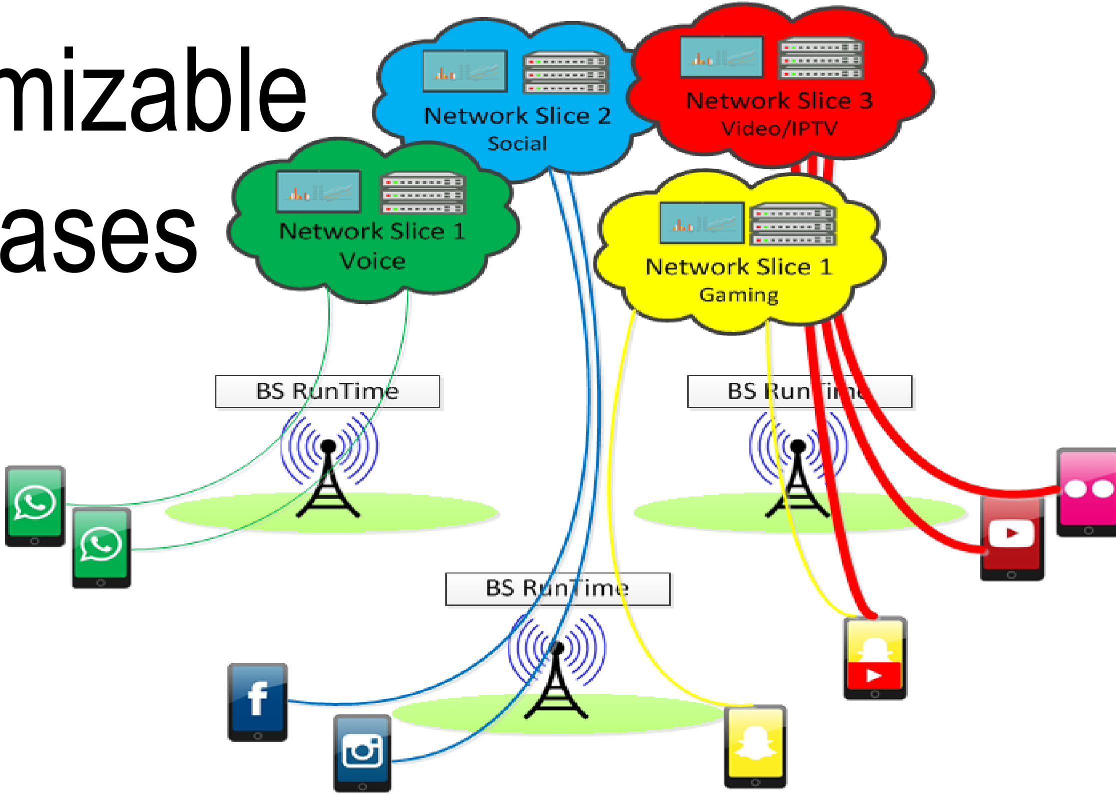
One-Network, Many-Service


Software Defined Networking
Fog Computing
Edge Computing
SDN/NFV Orchestration
Network Function Virtualization
Cloudification Virtualization
Contextual Networking
Heterogeneous Networking
Self Organization Networking
Ultra dense network
Advanced MIMO
Advanced waveforms
Millimeter Wave
Carrier Aggregation of discontinuous bands
Flexible and high capacity backhaul
Single channel full duplexing
New Spectrum Allocations
More Flexible Spectrum
© Coherent Project
5G technology enablers

Network Slicing

Flexible & Customizable for each use-cases





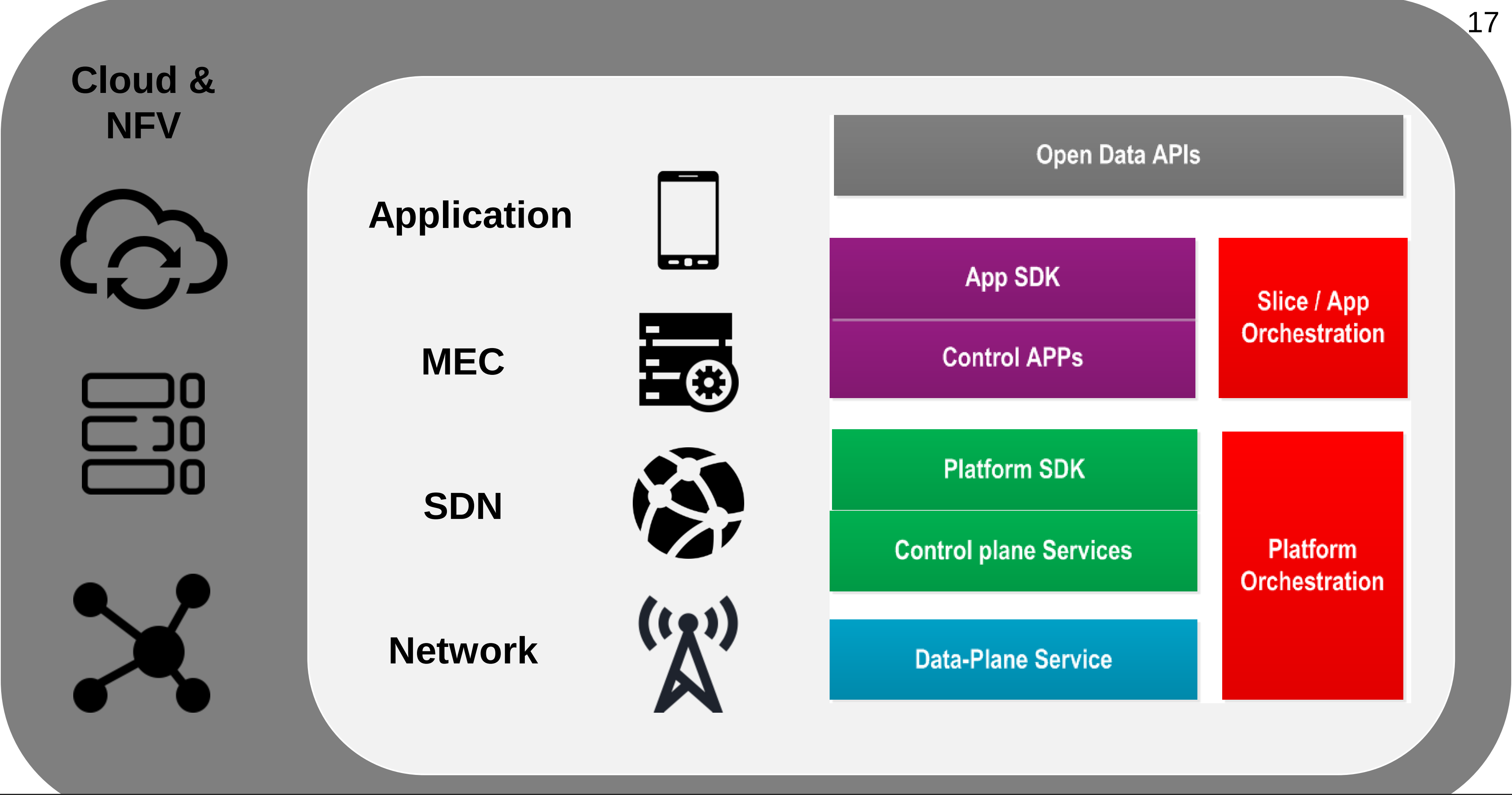
Service-oriented 5G

Slicing Technology Enablers:

- Softwarization
- Virtualization
- Disaggregation

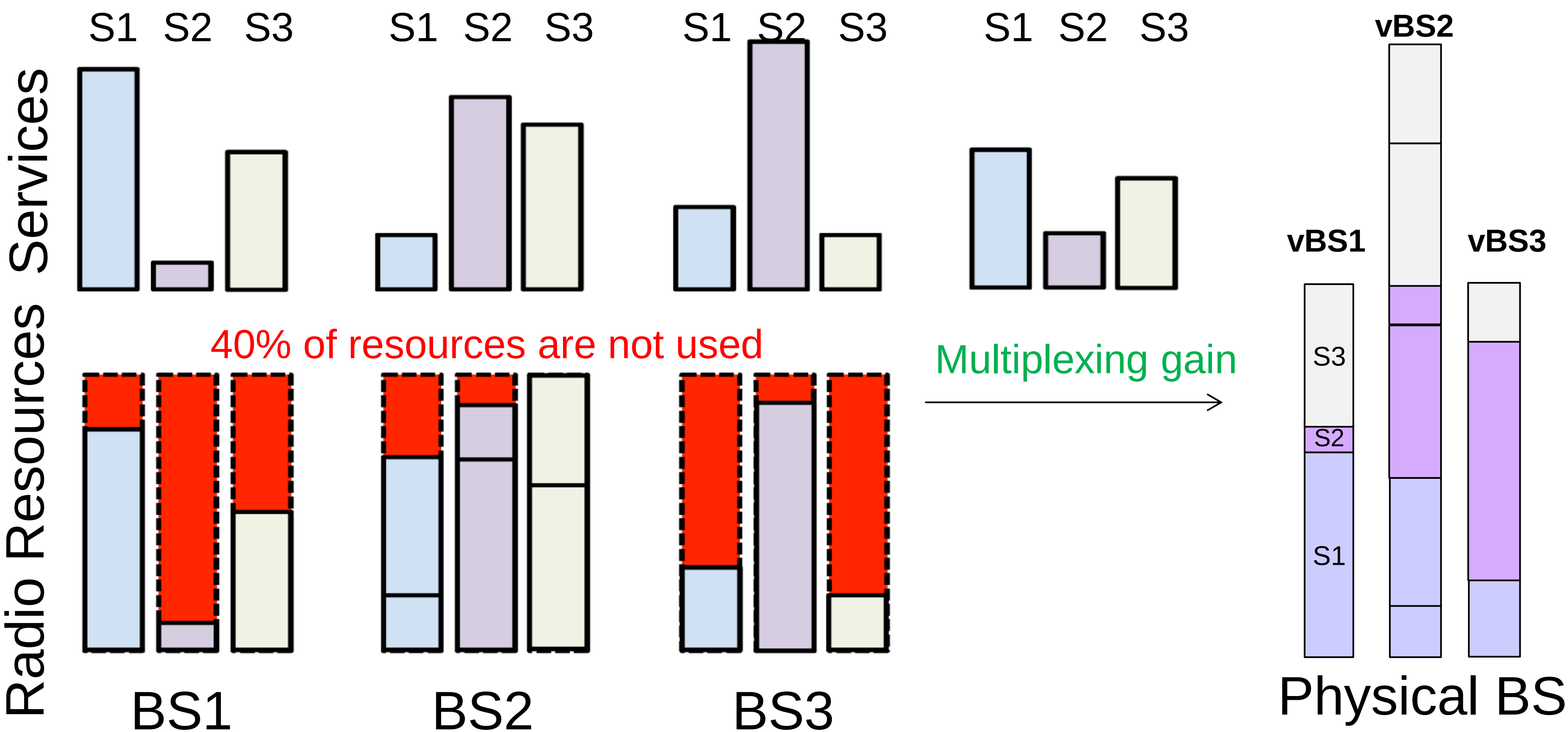
Multi-service multi-tenant network

Service-oriented 5G



Slicing Technology Enablers

Disaggregation

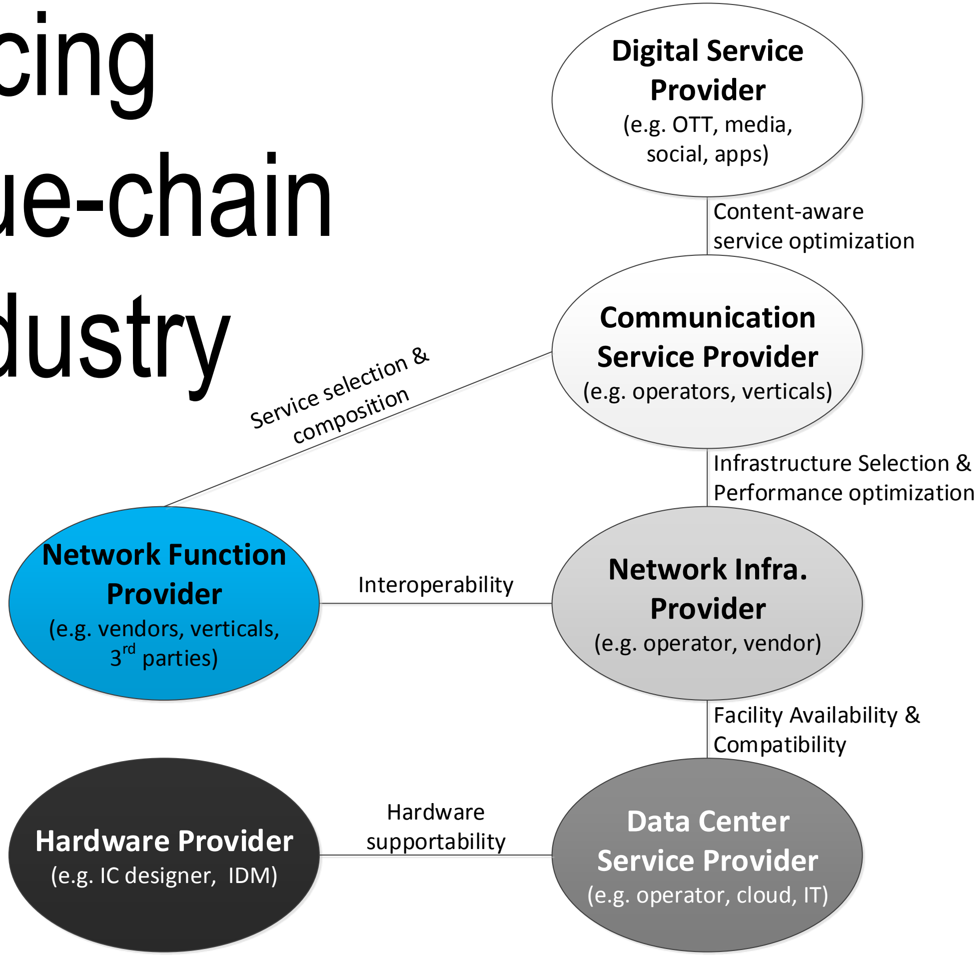


Slicing Technology Enablers

Why will it happen?

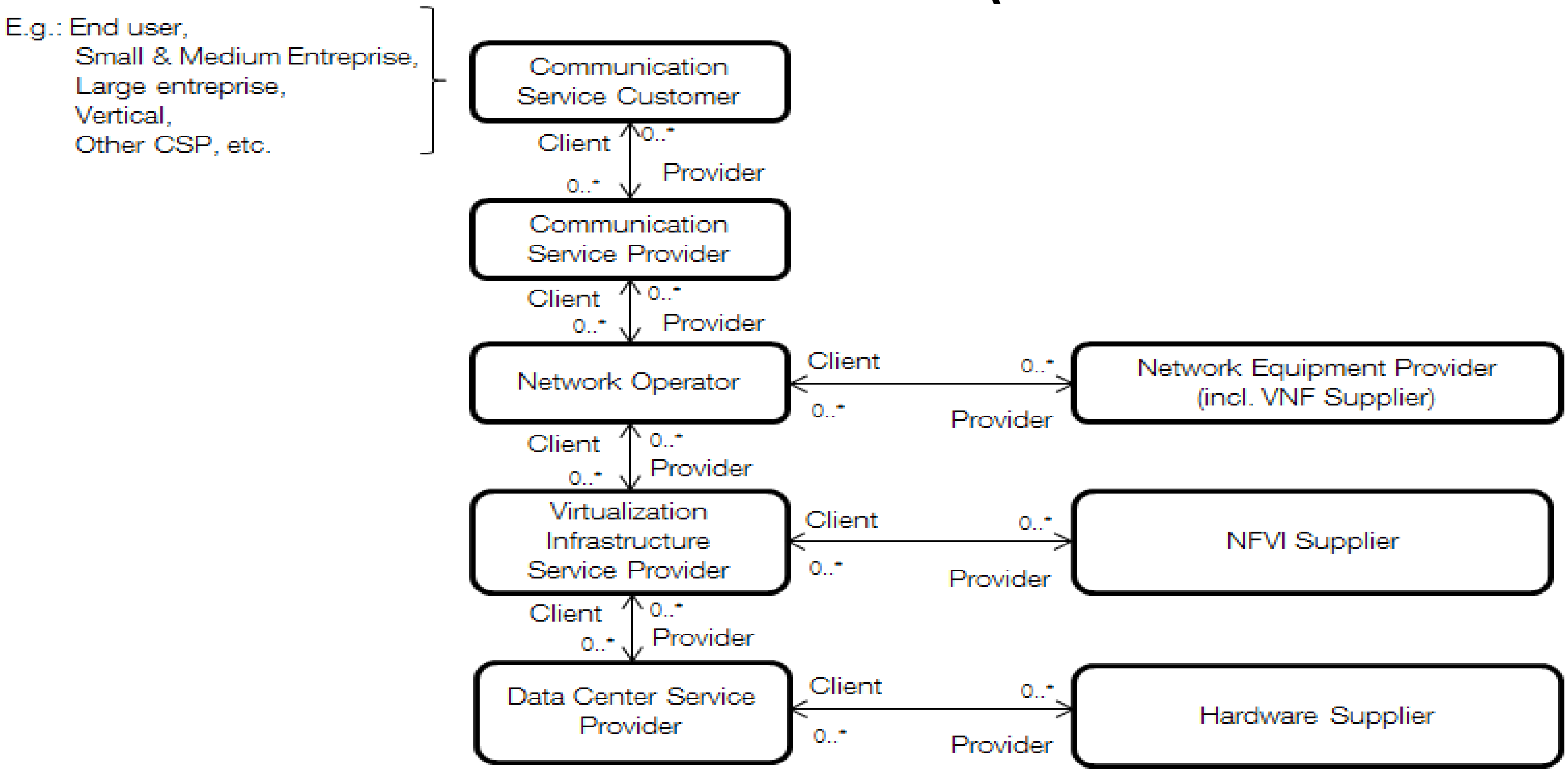
Extreme network flexibility and modularity

Network Slicing Evolves the value-chain of telecom industry



Service-oriented 5G

3GPP Role Model (3GPPP 28.801)



Service-oriented 5G

3GPP re-architects mobile networks

	3G	4G	5G
Downlink waveform	CDMA	OFDM	OFDM, SCFDMA
Uplink waveform	CDMA	SCFDMA	OFDMA, SCFDMA
Channel coding	Turbo	Turbo	LDPC (data) / Polar (L1 contr.)
Beamforming	No	Only data	Full support
Spectrum	0.8 – 2.1 GHz	0.4 – 6 GHz	0.4 – 90 GHz
Bandwidth	5 MHz	1.4 – 20 MHz	Up to 100 MHz (400MHz for >6GHz)
Network slicing	No	No	Yes
QoS	Bearer based	Bearer based	Flow based
Small packet support	No	No	Connectionless
In-built cloud support	No	No	Yes

Service-oriented 5G

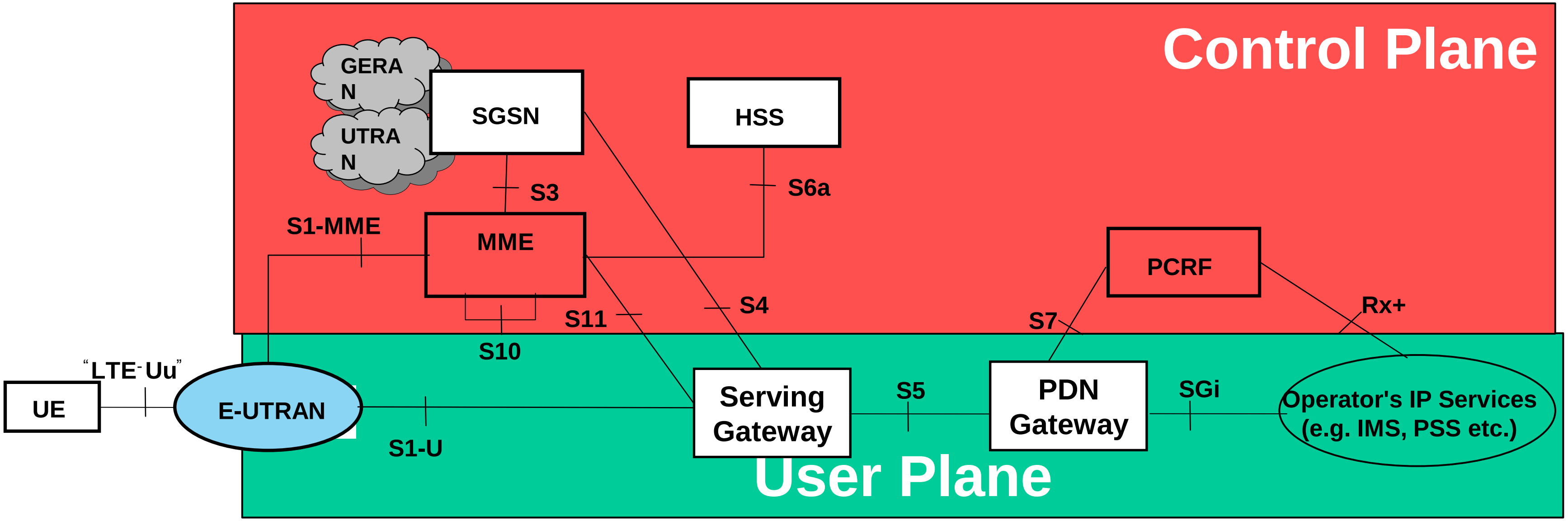
Monolithic BS

Stateful network entities

Transactional communication mode

Certain level of CP and UP separation

Common entity for user mobility and session management



Communication-oriented 4G

3GPP Network Sharing Models

Multi-operator RAN(MORAN)

- Shared RAN nodes, dedicated spectrum, but separated CN per operator

Multi-operator CN (MOCN)

- Shared RAN nodes and spectrum, but separated CN per operator with proprietary services

Gateway CN (GWCN)

- shared RAN and part of core networks

Dedicated core (DECOR)

- Deploy multiple dedicated CNs (DCNs) within a single operator network
- One or multiple MMEs and SGWs/PGWs, each element

Evolved DECOR (eDECOR)

- UE assisted DCN selection
- Network Node Selection Function (NNSF) at RAN to select directly the proper DCN towards which the NAS signaling needs to be forwarded
- Congestion control and load balancing among multiple DCN with shared MME

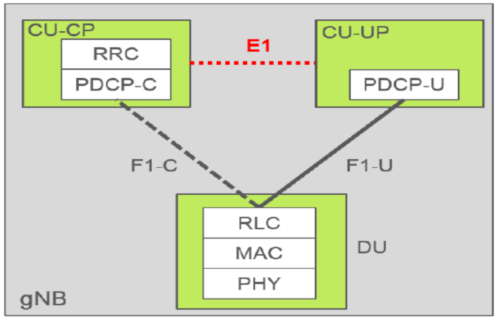
3GPP re-architects mobile networks

3 Tier RAN Node

CU0 → DU[0-n] → RRU[0-m]

Functions Split

CP UP split

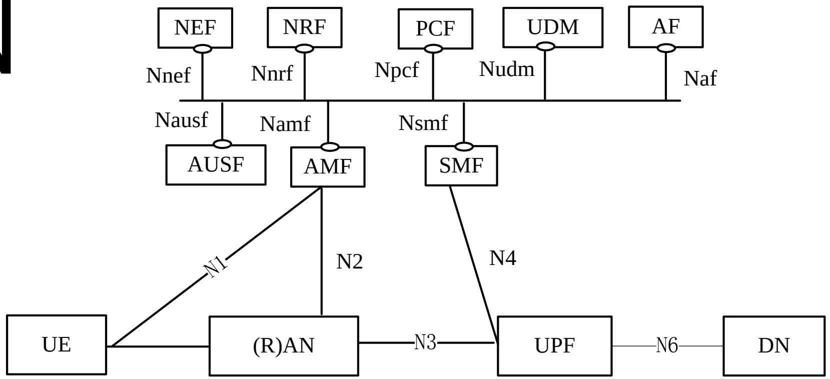


Service-oriented CN

service catalog and discovery

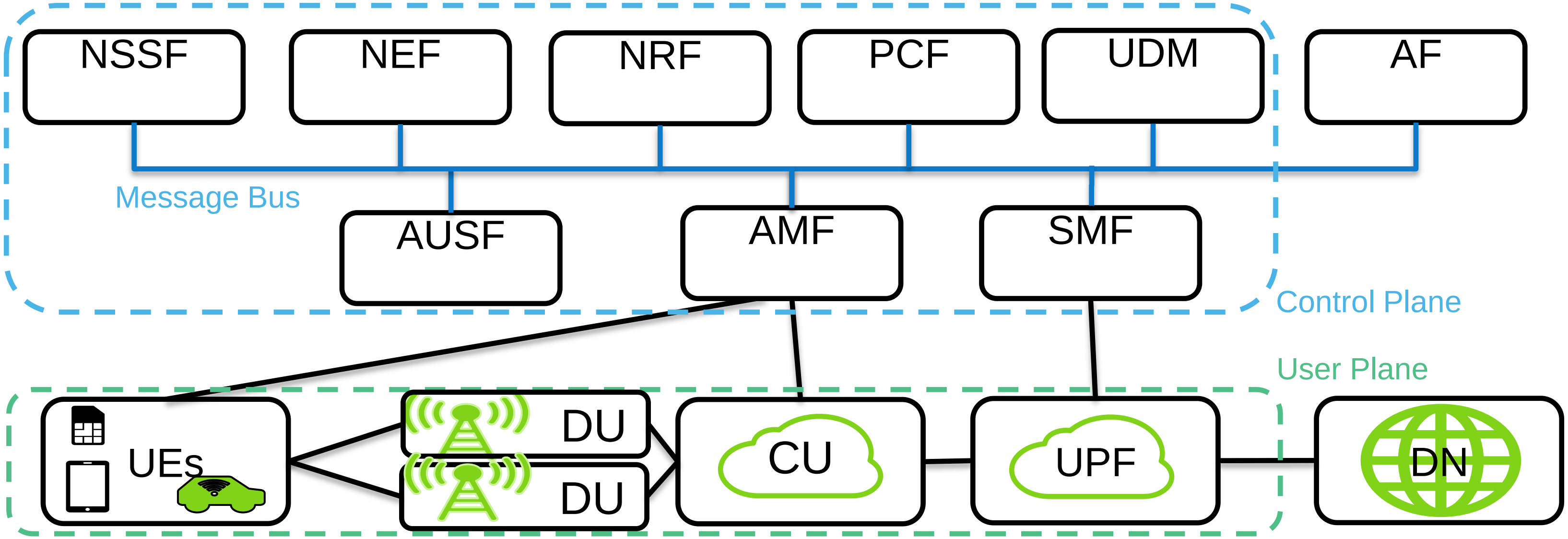
Slice selection function

CP and UP split



3GPP 5G RAN and CN

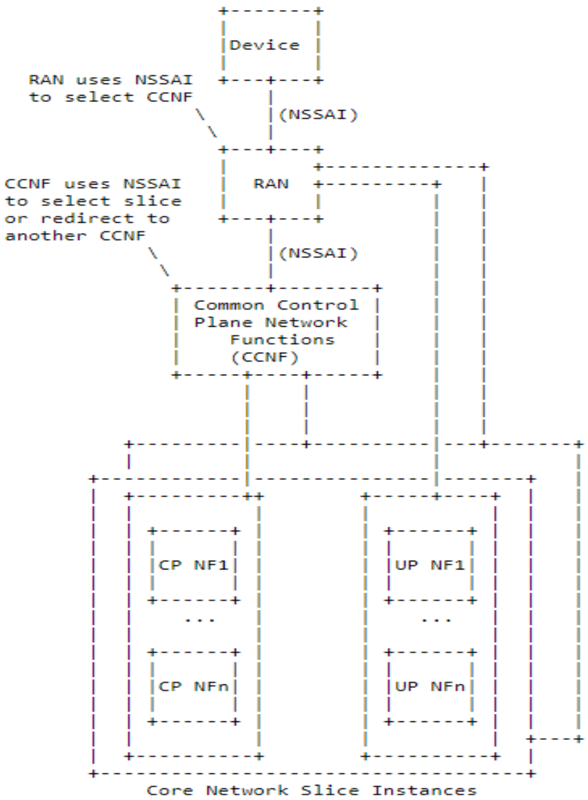
3GPP re-architects mobile networks

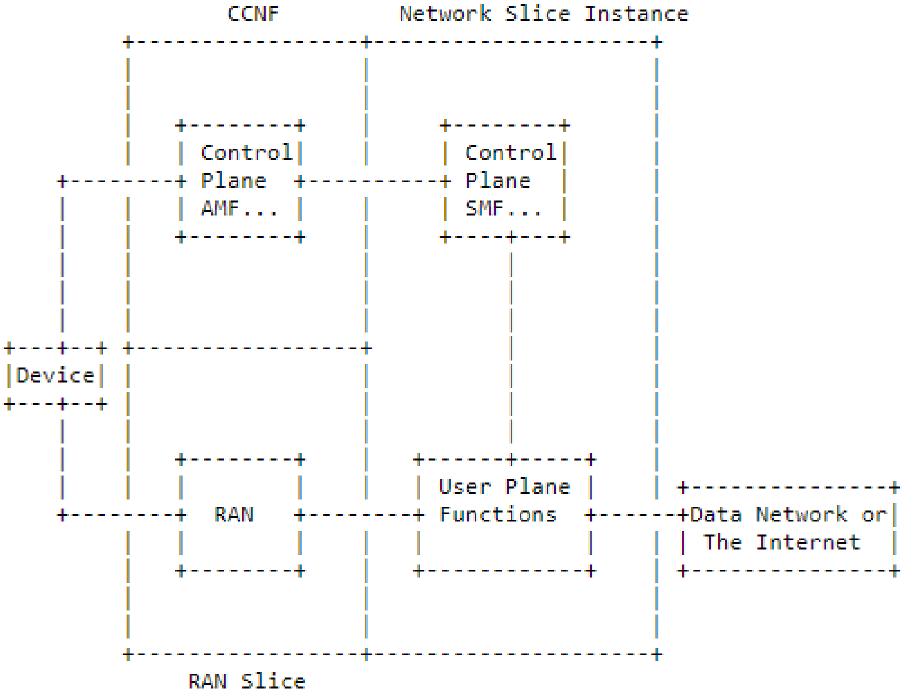


AMF	Access & Mobility Management Function	SMF	Session Management Function
AUSF	Authentication Server Function	UPF	User Plane Function
NRF	Network Repository Function	AF	Application Function
UDM	Unified Data Management	PCF	Policy Control Function
NSSF	Network slice selection function	NEF	Network Exposure Function

Service-oriented 5G

3GPP re-architects mobile networks





3GPP network slicing

Select the set of network slice instances serving the UE

Determine the allowed Network Slice Selection Assistance Information (NSSAI) and the mapping to the subscribed S-NSSAIs

Determine the configured NSSAI and the mapping to the subscribed S-NSSAIs

Determine the AMF set to be used to serve the UE or a list of candidate AMFs by querying the NRF

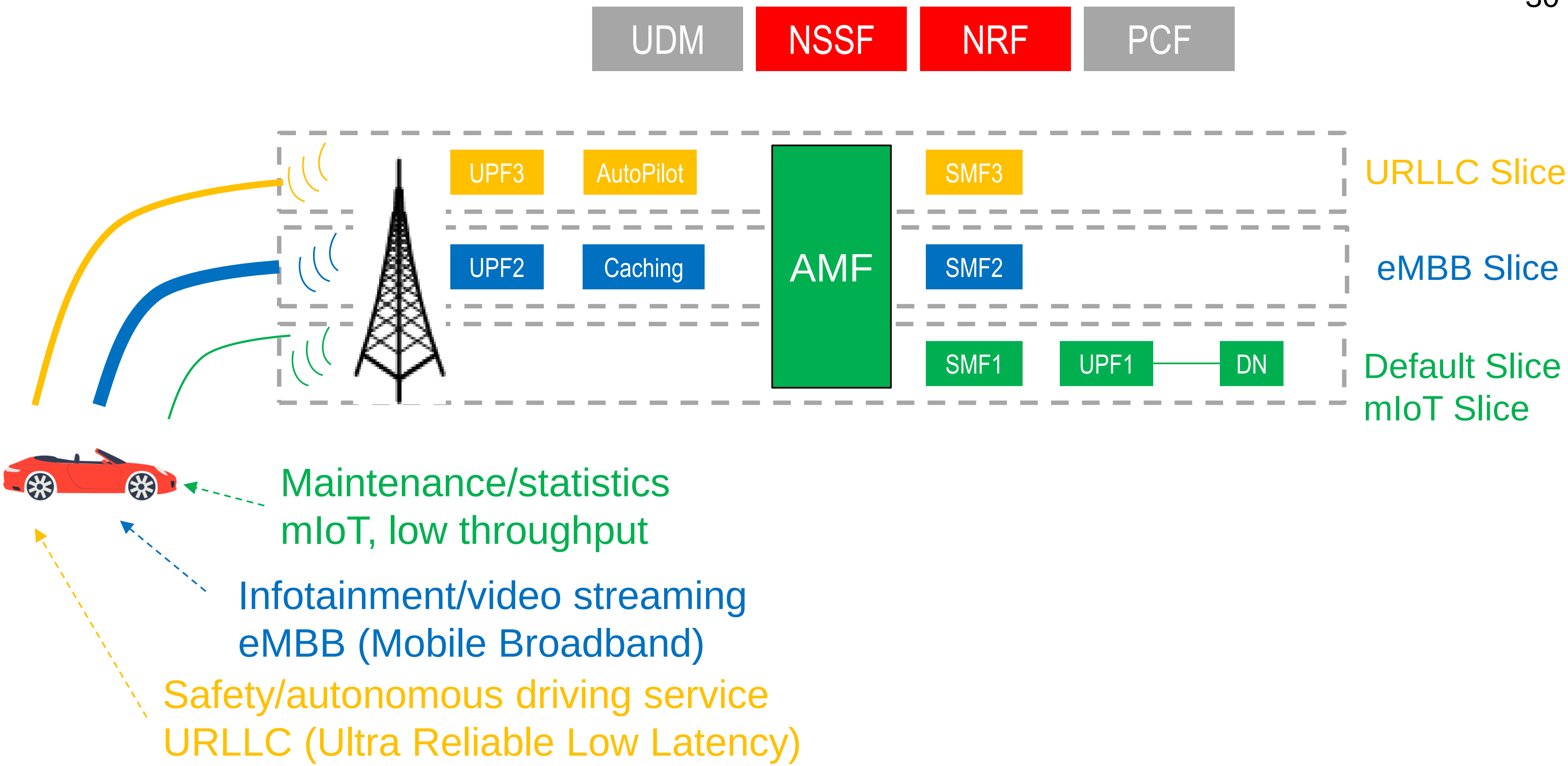
NSSF: Network slice selection function

Provides information on the discovered NF instances upon discovery requests

Maintains the NF profile of available NF instances and their supported services

NF Profile: instance ID, type, PLMN ID, Network Slice identifiers, IP address of NF, NF capacity information, NF specific service authorization information, names of supported services, endpoint addresses of supported services, identification of stored data information

NRF: network repository function

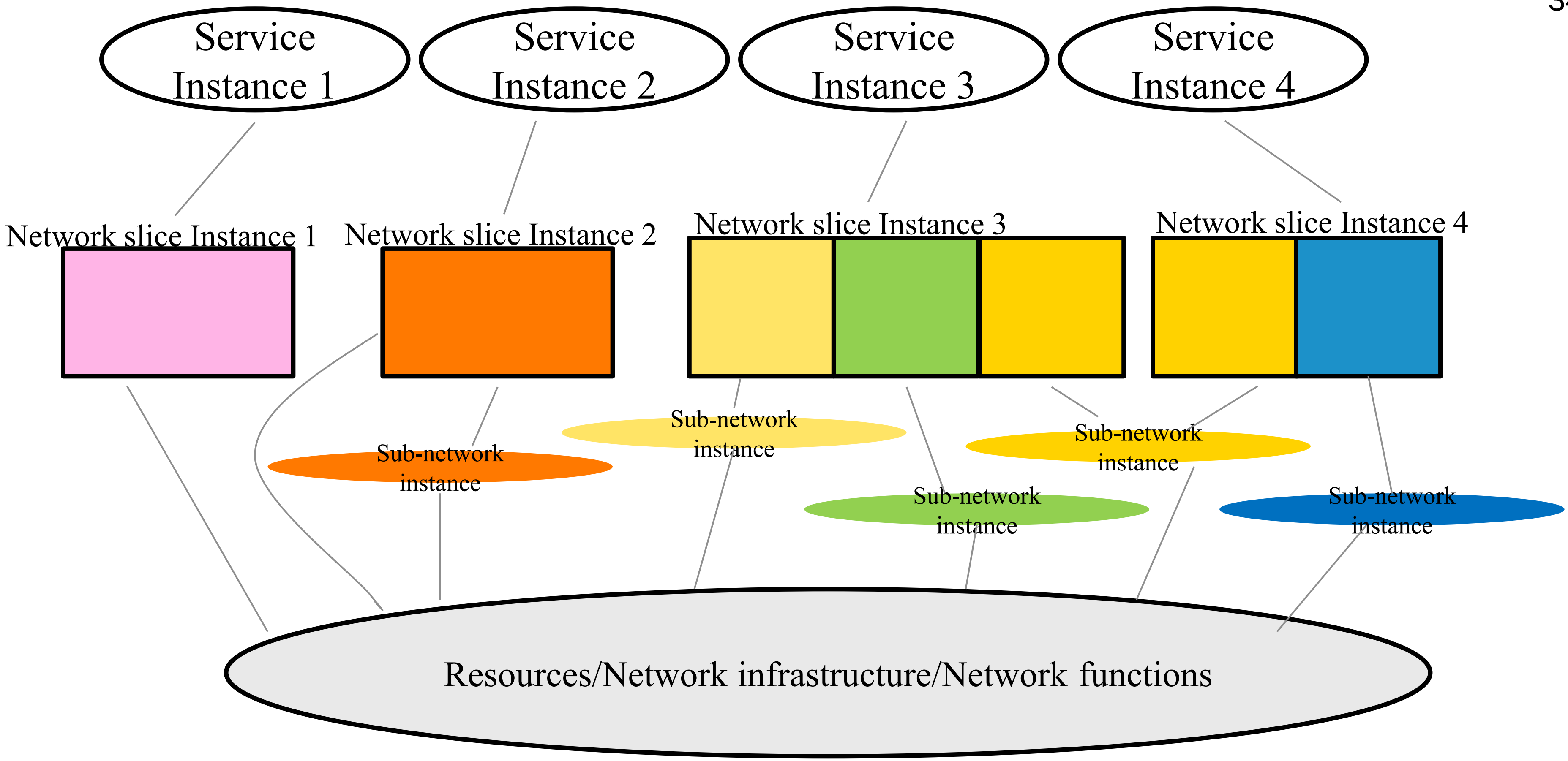


Dedicated or Shared Functions?

Dedicated or Shared Resources?

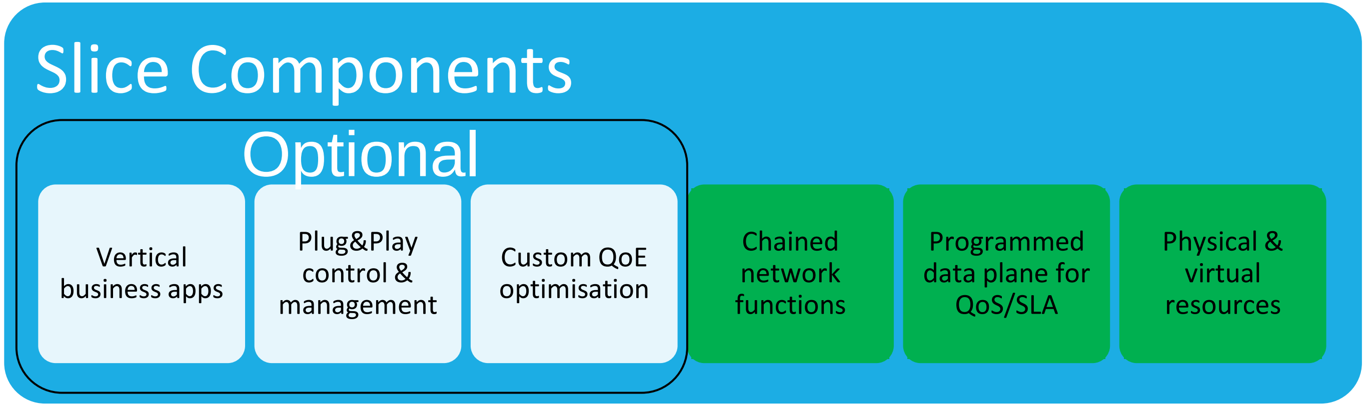
Dedicated or Shared Resources?

RAN Slicing



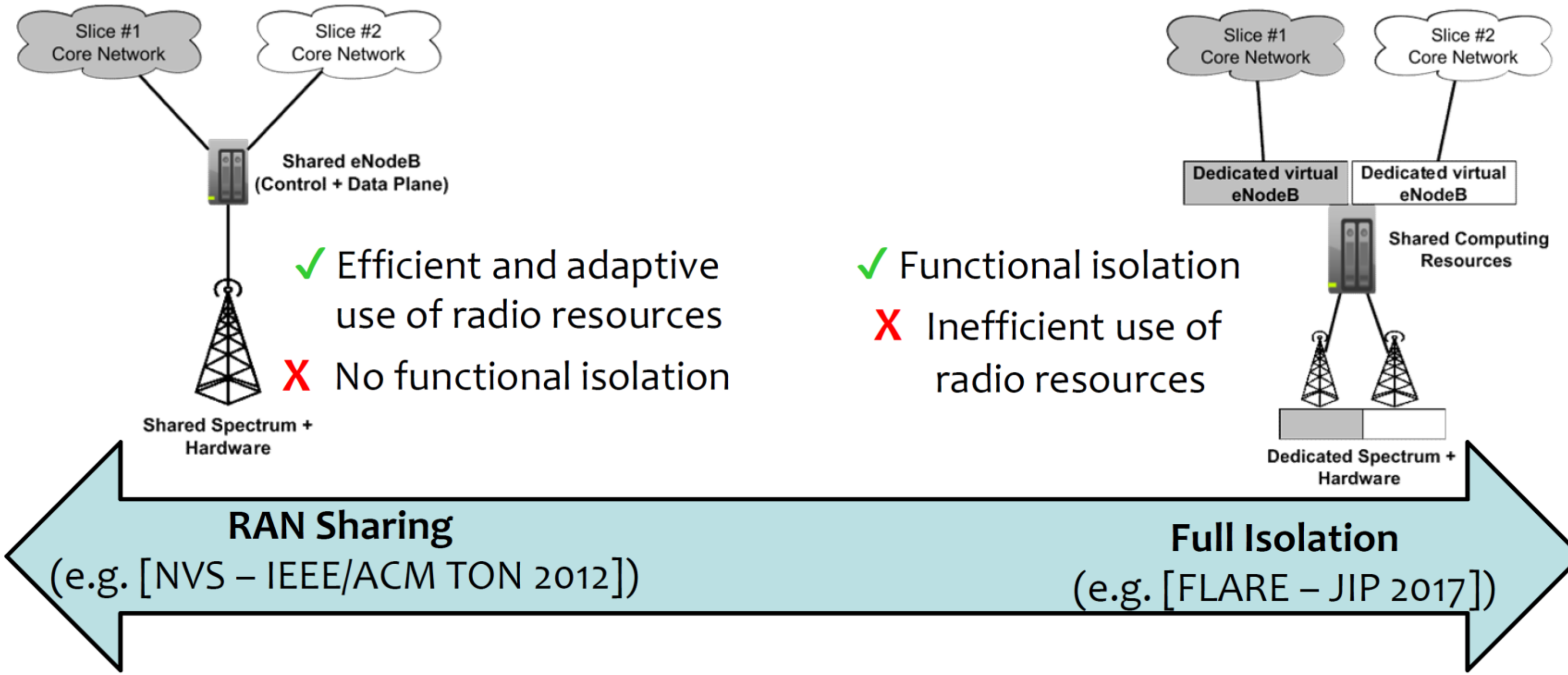
Network Slicing Concept

Composition and deployment of multiple E2E logical networks tailors to a service over a shared infrastructure, and their delivery as a slice



What is a slice?

RAN Slicing



Dedicated and Shared

FlexRAN : a SD-RAN platform enabling RAN sharing (Foukas et al., 2016)

Fully isolation platform with vBSs as different slices (Nakao et al., 2017)

Separated radio resources for intra/inter-slice scheduler (Rost et al., 2017)

RRM is enforced using a resource visor per slice (Ksentini et al., 2017)

ORION: BS hypervisor isolate slice-specific control logics and share the virtualized radio resources (Foukas et al., 2017)

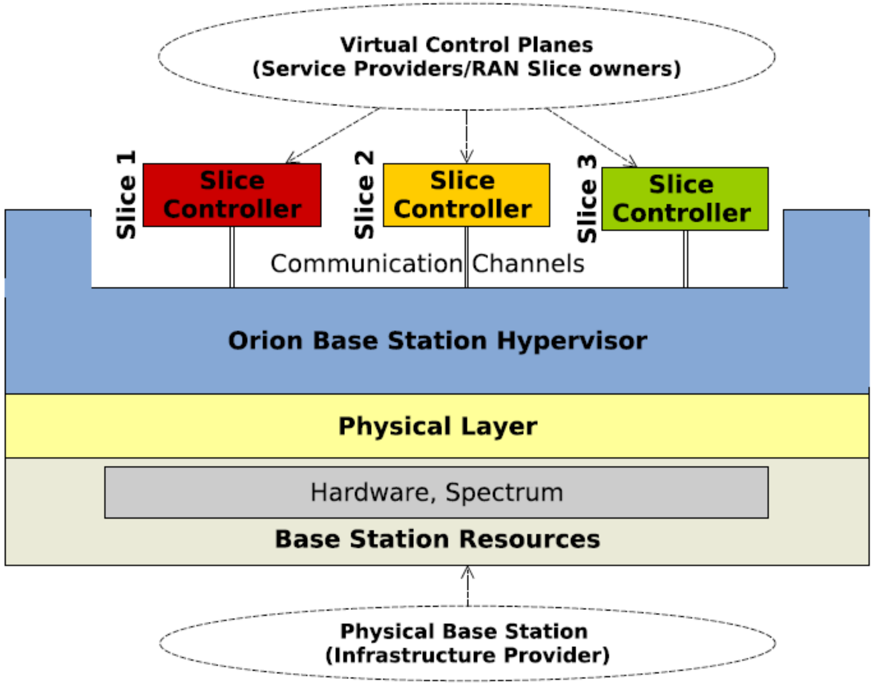
RAN runtime targets customization and multiplexing in several aspects over disaggregated RAN (Chang et al., 2017)

State of the Art

RAN slicing system

(1) Isolate slice-specific control logics while keeping common CP/UP functions

(2) Share radio resources in virtualized or physical form



ORION

Components

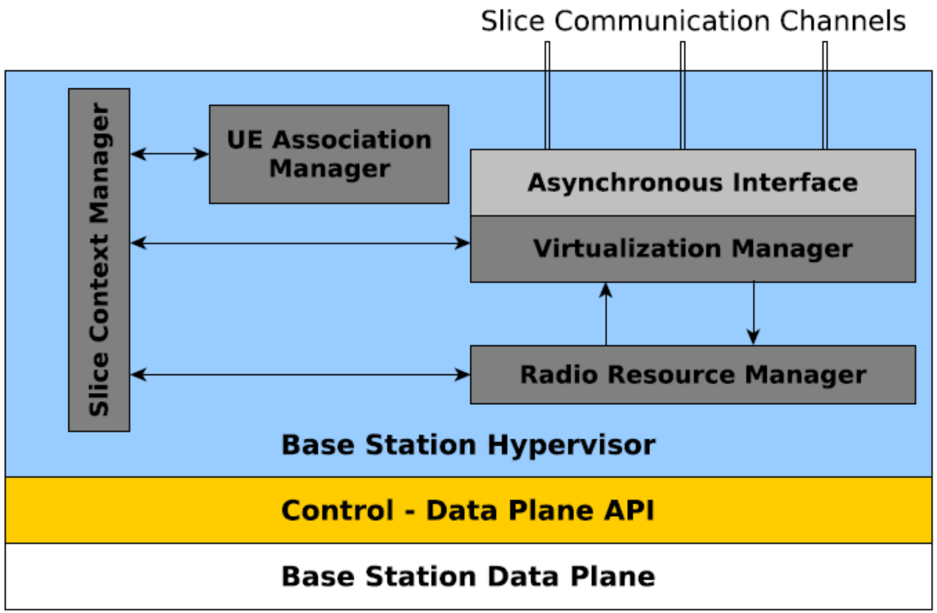
(1) **Slice context manager** performs lifecycle management of each slice (SLA, active UEs, admission control)

(2) **Virtualization manager**

- provides a generic form of abstraction for virtualizing radio resources and data plane state
- presents a virtual/isolated view to each slice virtual control plane

(3) **Radio resource manager** allocates physical resources among slices

(4) **UE association manager** handles slice discovery by UEs and maps UEs to slices



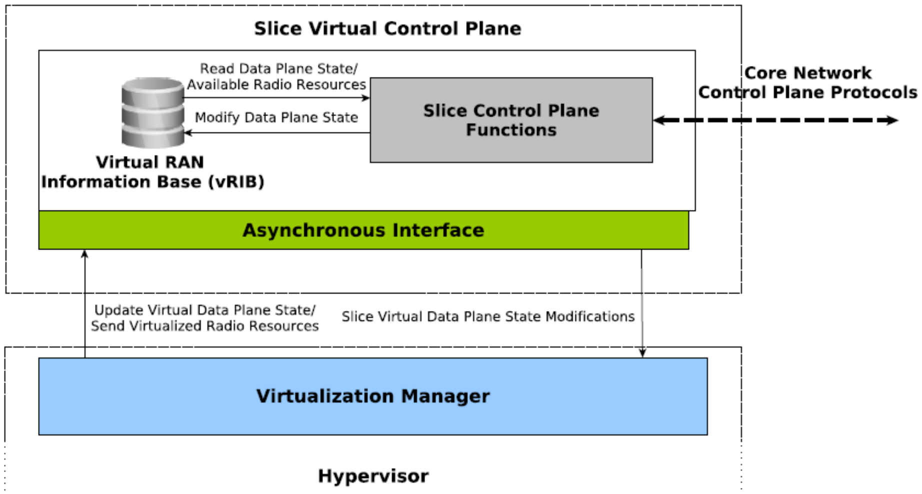
Virtual Control Plane

(1) Interacts with the underlying infrastructure via the virtualization Manager of the Hypervisor

- translated into control-data APIs

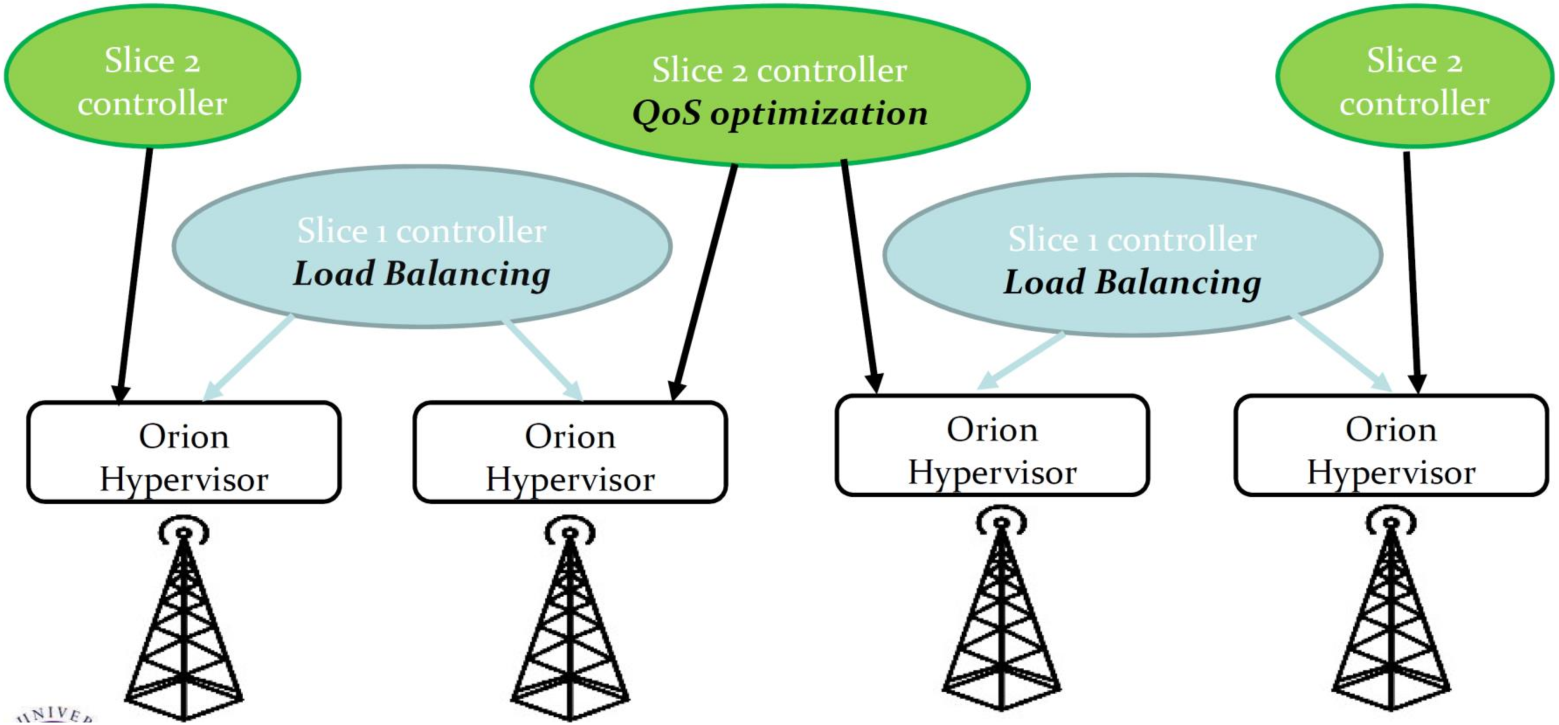
(2) Operates over vRIB, the locally maintained state of virtual radio resources and data plane

- Slice network view and state



ORION

ORION RAN Slicing System

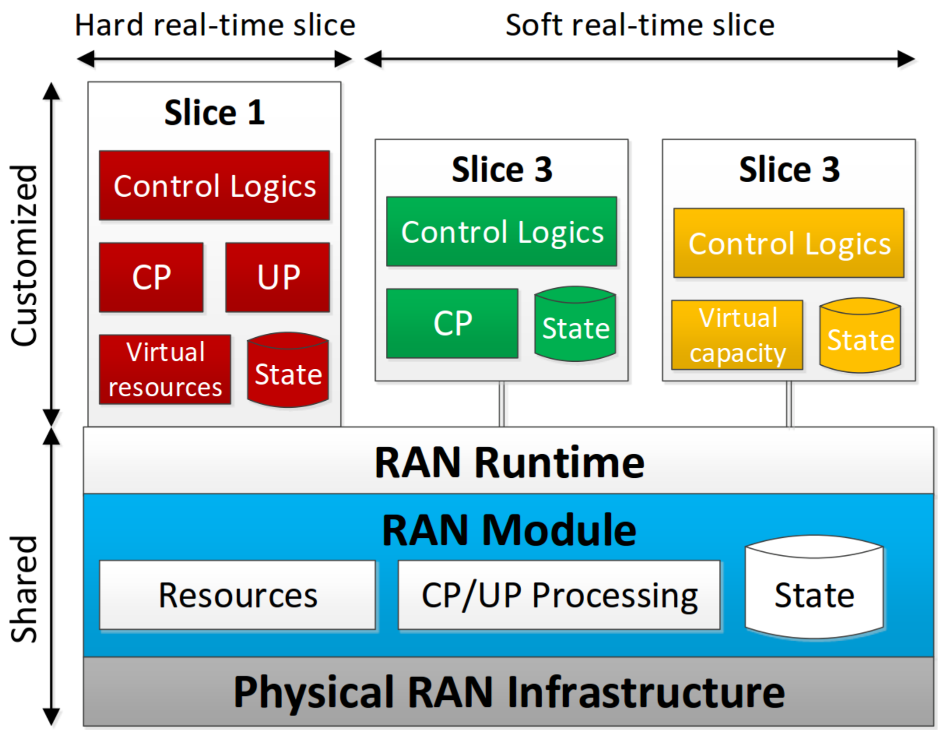


RAN Slicing Execution Env.

(1) run multiple virtualized RAN module instances with different level of isolation and sharing

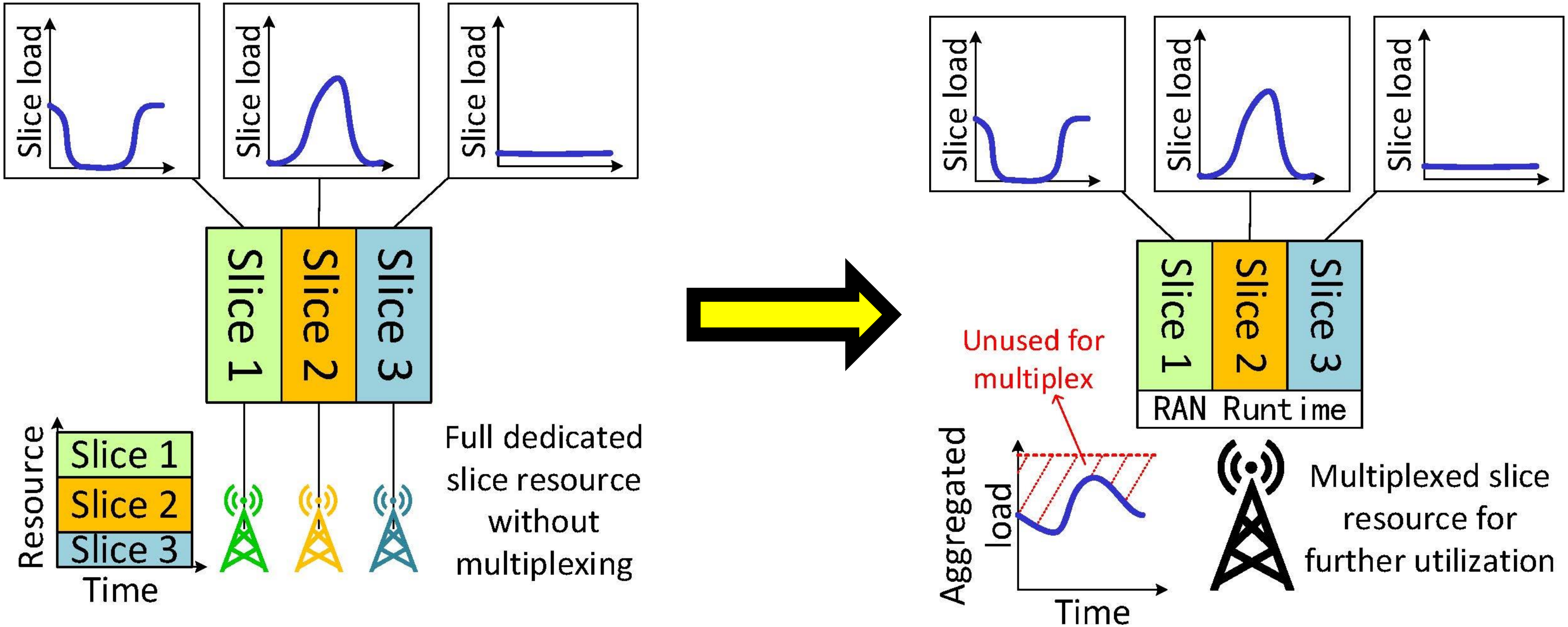
(2) Pipeline RAN functions to either via multiplexed or customized CP/UP functions

(3) Share radio resources in virtualized or physical form



RAN Runtime

Multiplexing Gain



RAN Runtime

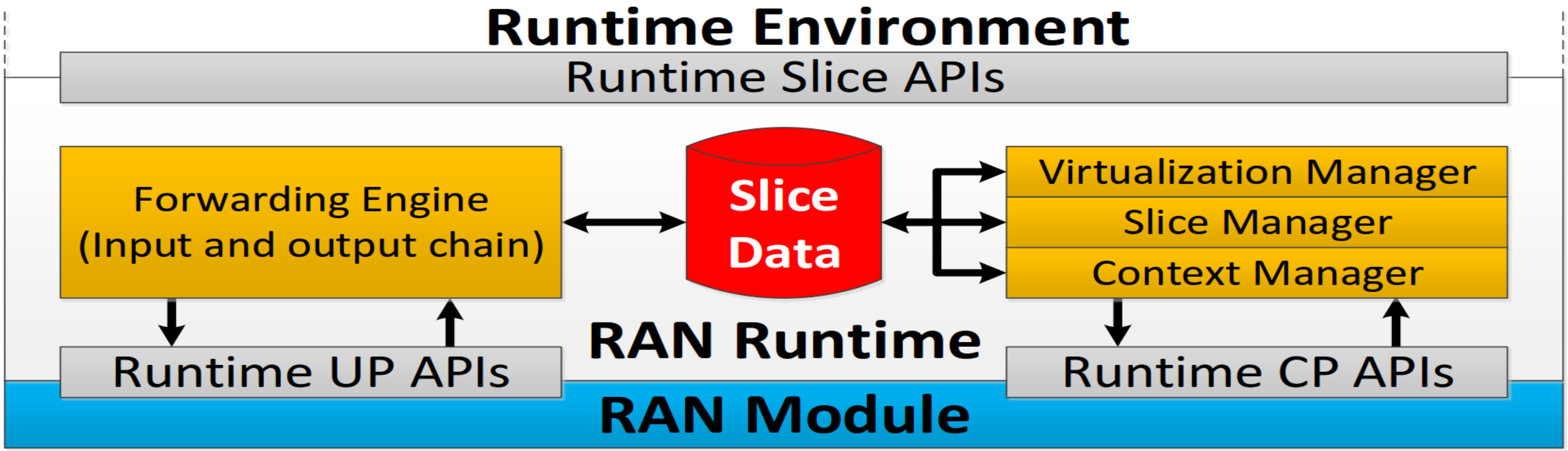
(1) Slice data: Slice context and RAN module context

(2) Context manager: Manage slice data and perform CRUD operation

(3) Slice manager: slice life-cycle, program forwarding engine, conflict resolution

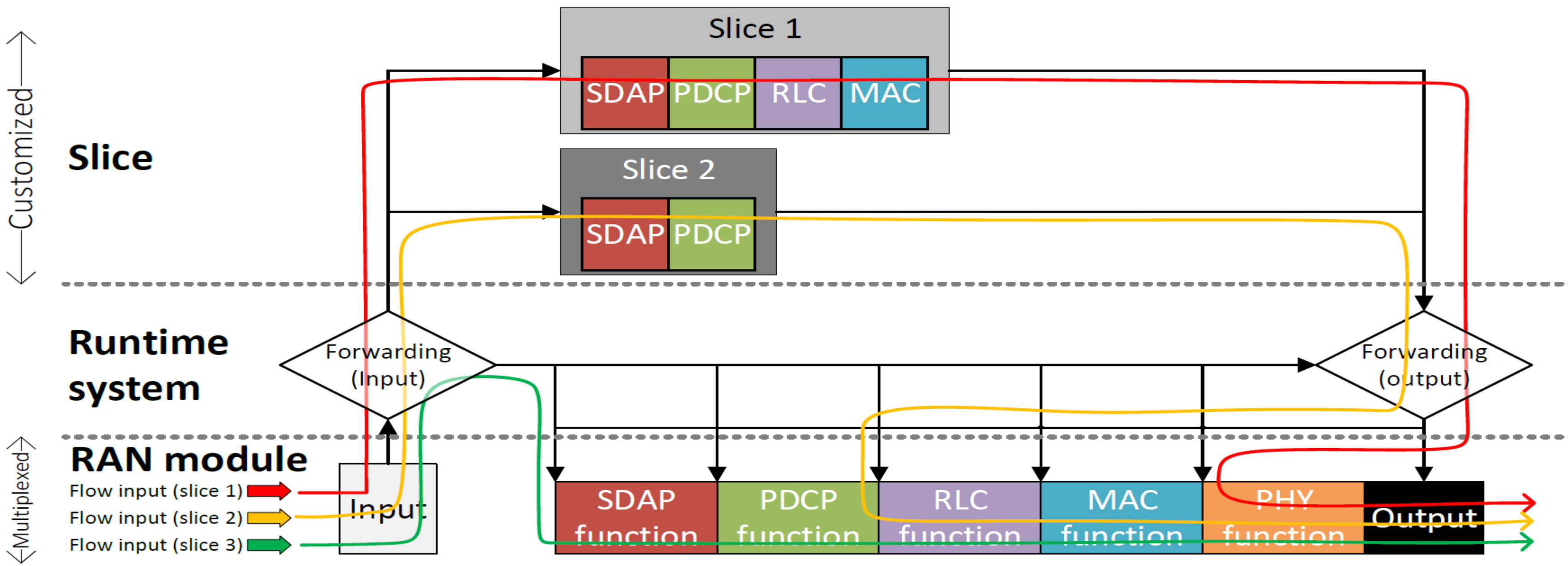
(4) Virtualization manager: resource abstraction, partitioning, and accommodation

(5) Forwarding engine: establish slice-specific UP path



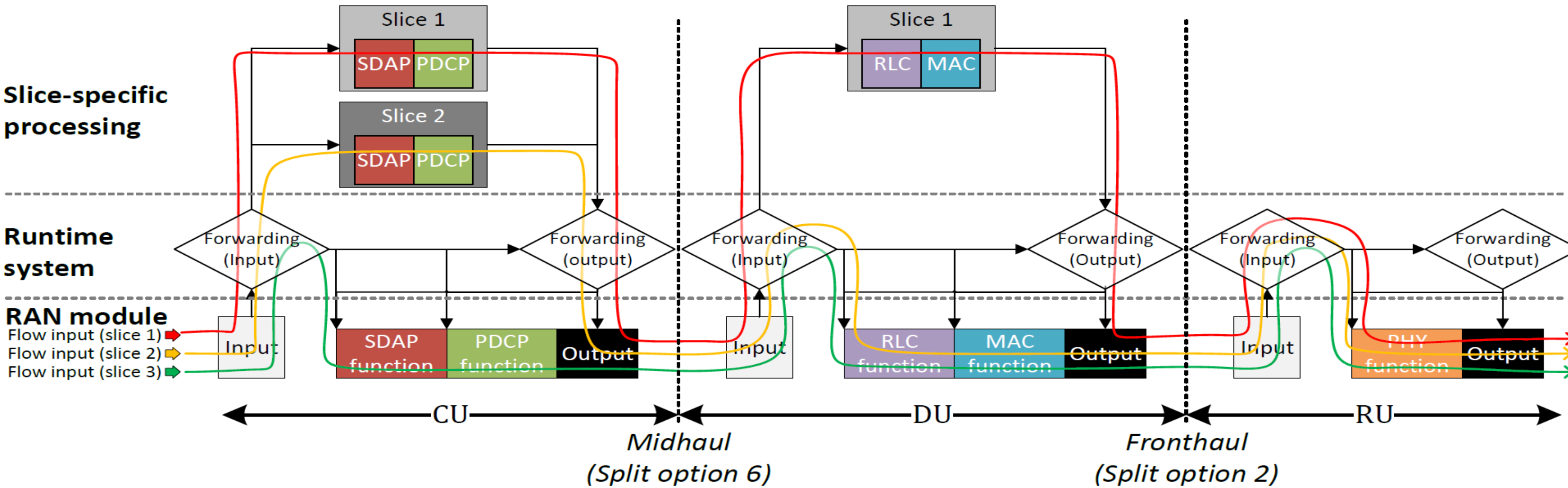
RAN Runtime

Function customization in Monolithic BS



RAN Runtime

Disaggregated BS



RAN runtime

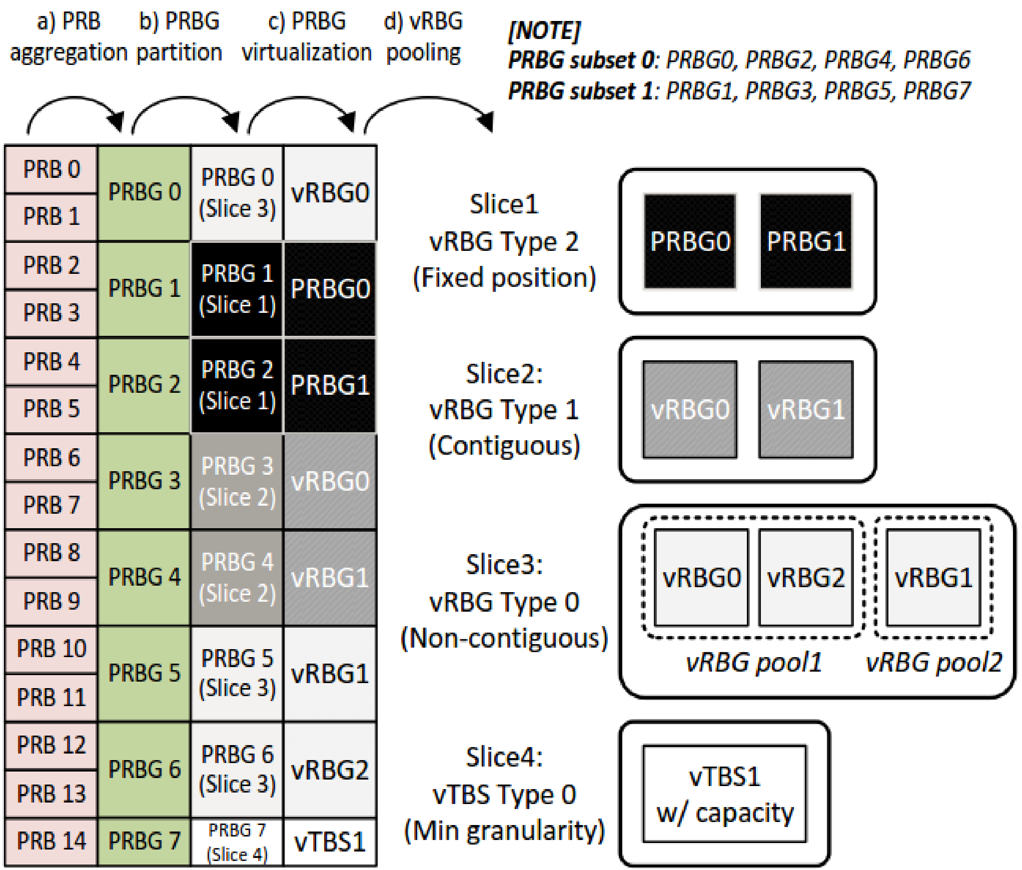
Resource Abstraction

Requested resources	Abstraction types (Resource granularity)	DL resource allocation type	UL resource allocation type
Resource Block	vRBG Type 0 (Non-contiguous)	Type 0, Type 1, Type 2 distributed	Type 1
	vRBG Type 1 (Contiguous)	Type 0, Type 2 localized	Type 0
	vRBG Type 2 (Fixed position)	Type 2 localized	Type 0
Capacity	vTBS Type 0 (Min RBG granularity)	All Types	All Types

RAN Runtime

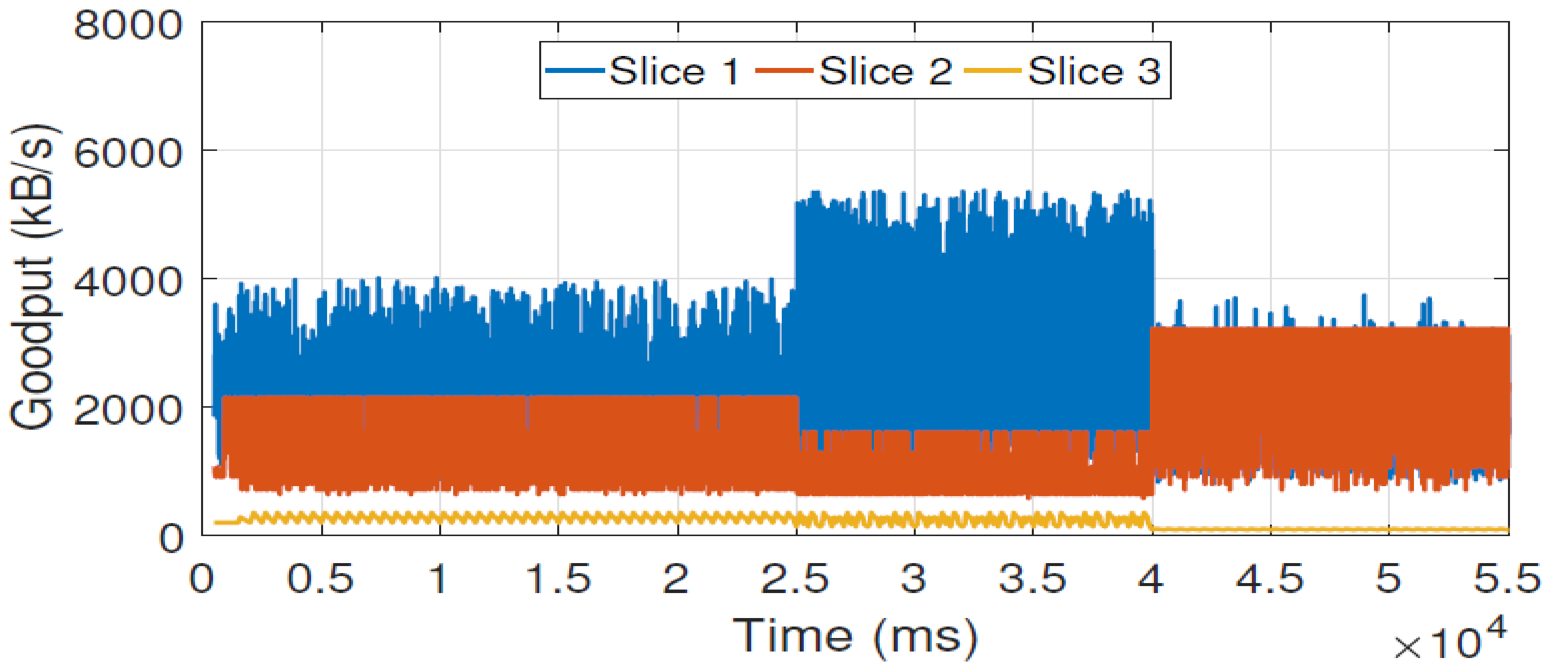
4 Steps to radio resources abstraction:

(1) Aggregation
(2) Partitioning
(3) Virtualization
(4) Polling
(5) Slice resource allocation
(6) Slice Scheduling & Accommodation
(7) Multiplexing/preemption



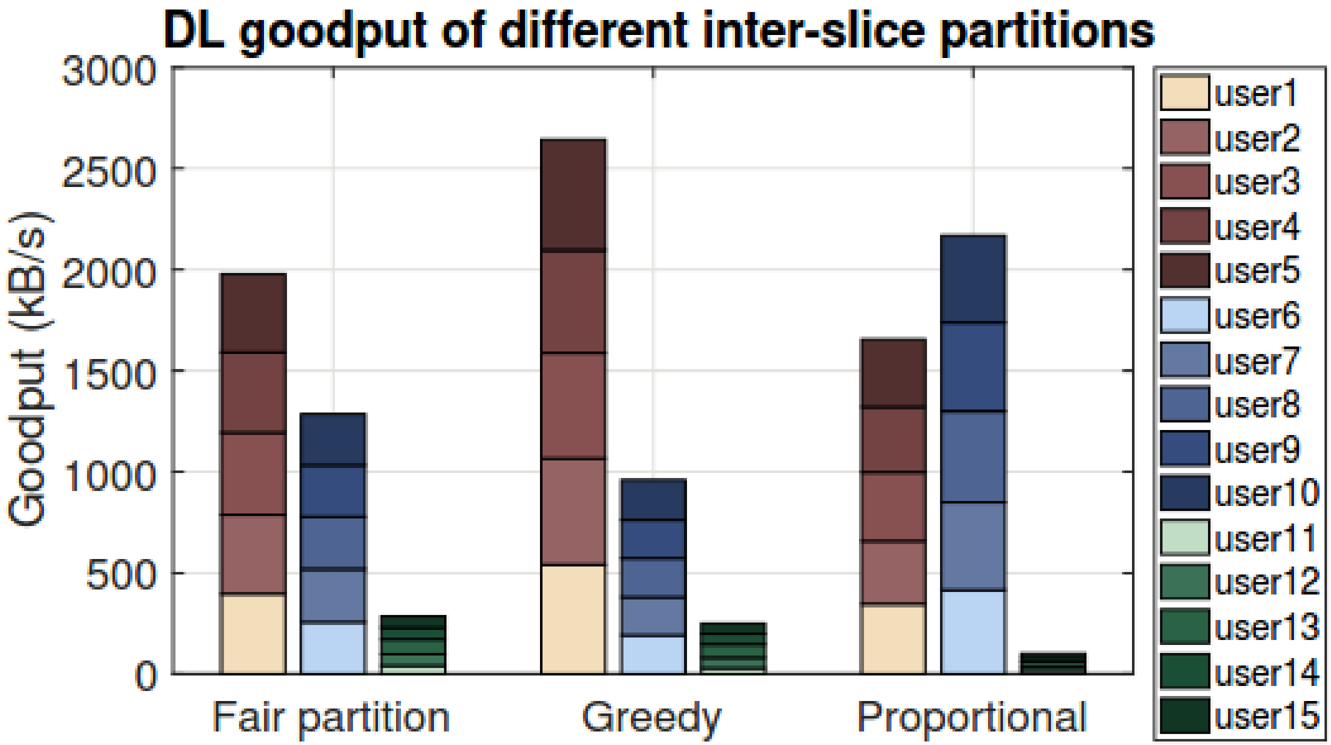
RAN Runtime

Inter-Slice Resource Partitioning and Polling



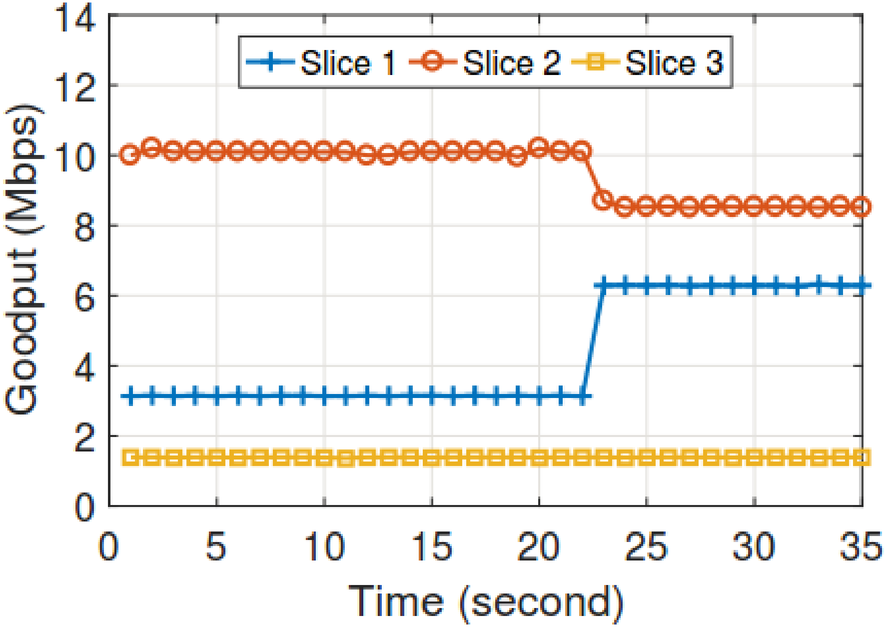
RAN Runtime

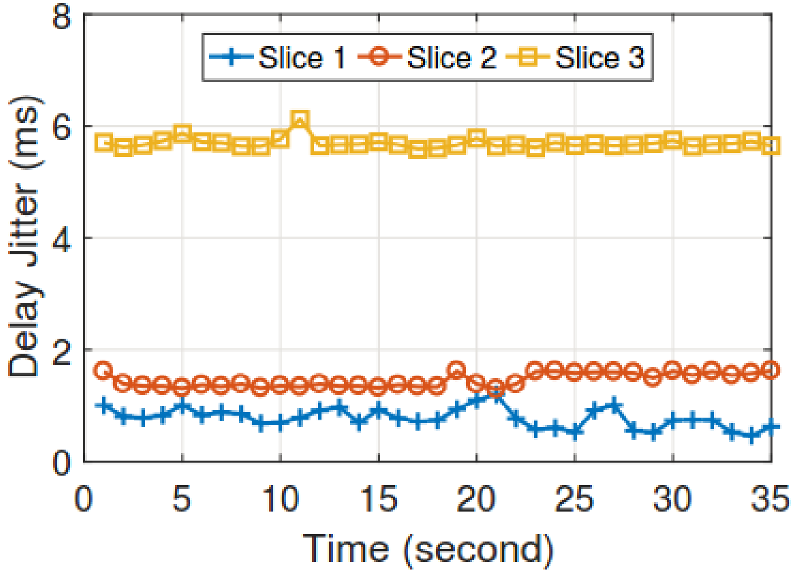
Decouple resource partitioning and accommodation from resource allocation



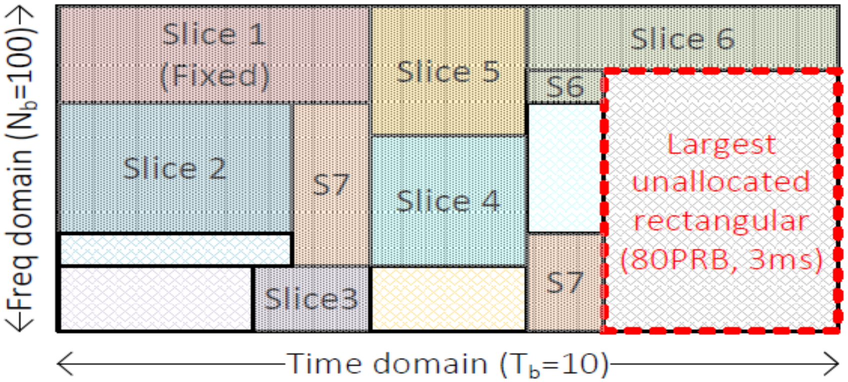
RAN Runtime

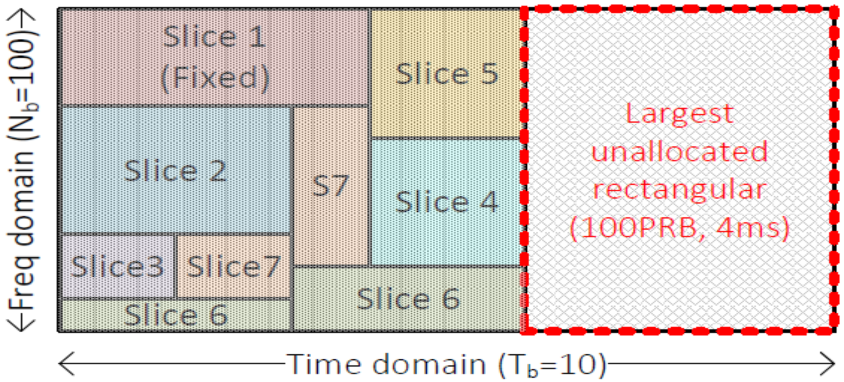
Slice QoS: Multiplexing/Preemption

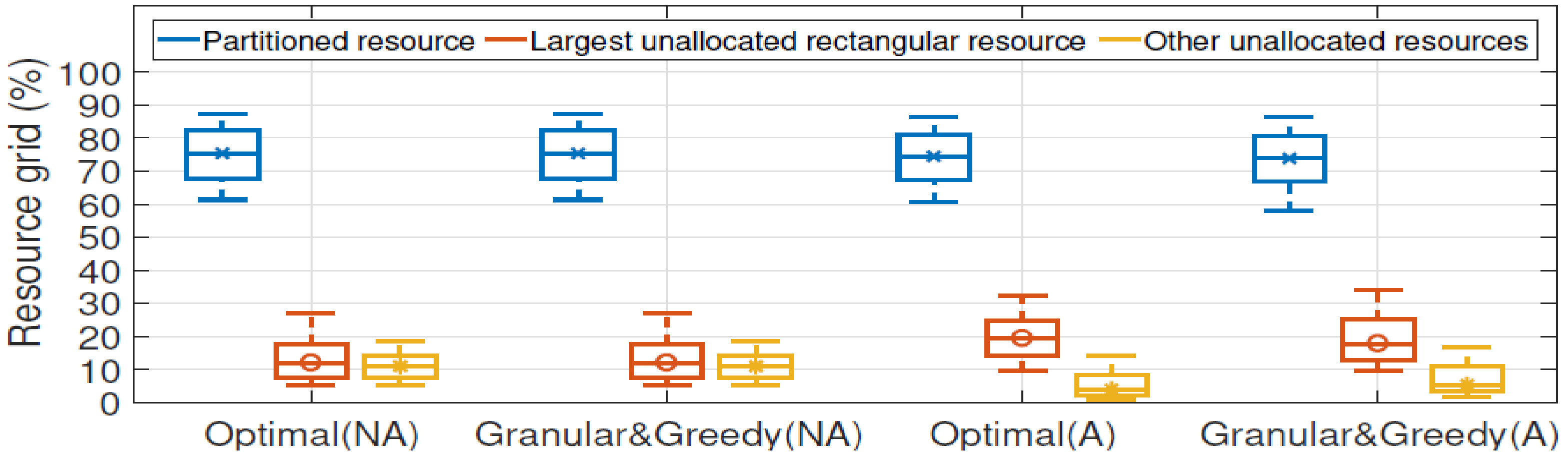




RAN Runtime

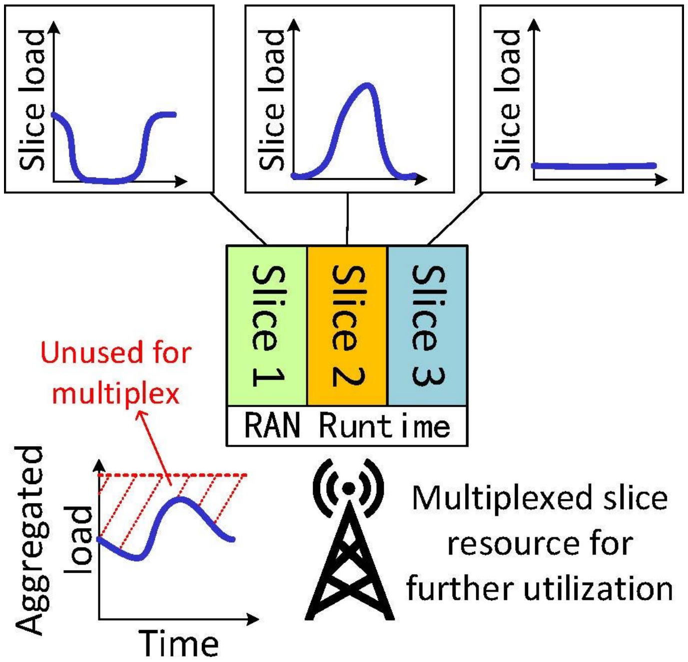

Slice 1
(Fixed)
Slice 5
Slice 6
S6
Slice 2
S7
Slice 4
Largest
unallocated
rectangular
(80PRB, 3ms)
S7
Slice3
←Freq domain (N_b=100)→
←Time domain (T_b=10)→

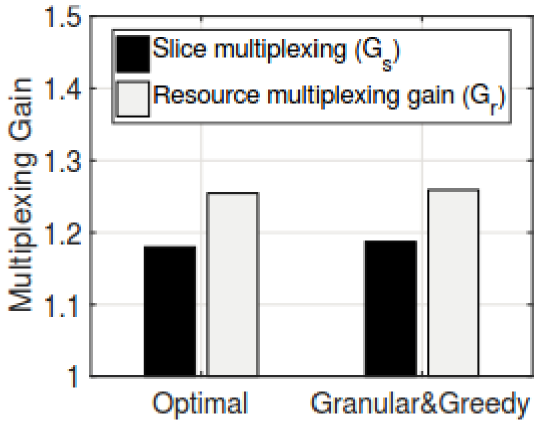

Slice 1
(Fixed)
Slice 5
Largest
unallocated
rectangular
(100PRB, 4ms)
Slice 2
S7
Slice 4
Slice3
Slice7
Slice 6
Slice 6
←Freq domain (N_b=100)→
←Time domain (T_b=10)→


Partitioned resource
Largest unallocated rectangular resource
Other unallocated resources
Resource grid (%)
100
90
80
70
60
50
40
30
20
10
0
Optimal(NA)
Granular&Greedy(NA)
Optimal(A)
Granular&Greedy(A)

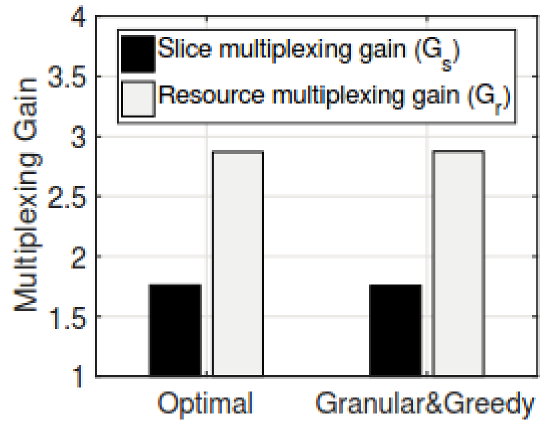
RAN Runtime

Multiplexing Gain





High traffic arrival rate



Low traffic arrival rate

RAN Runtime

Maximize the multiplexing gain

Isolate tenants resources

Customize tenant service

Benefit of Slicing

What is the typical number of slices?

What is the typical lifetime of a slice?

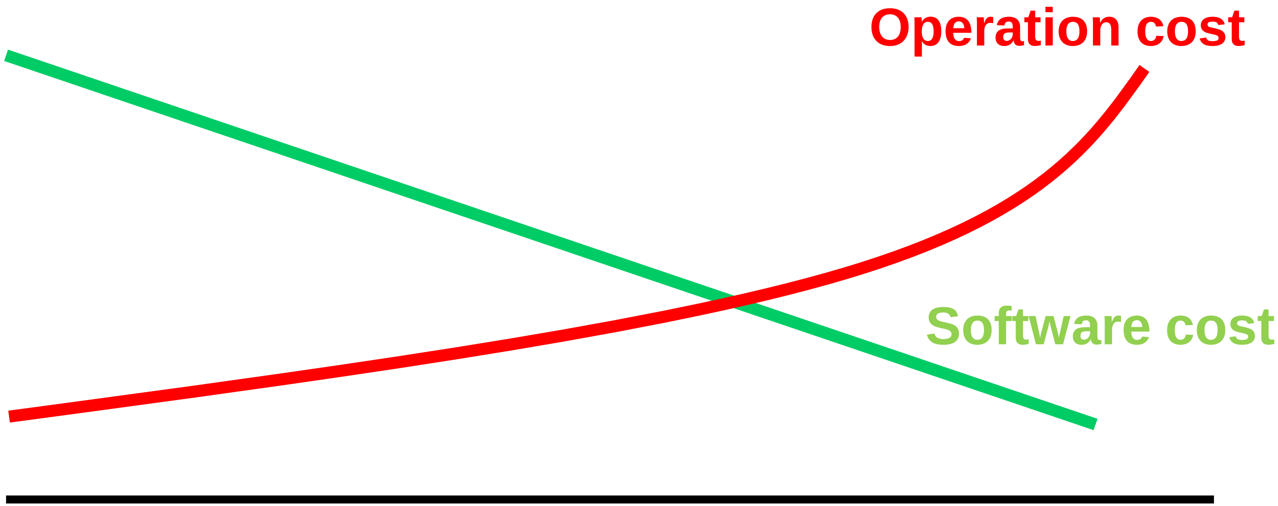
Two numbers in Slicing

Slice Orchestration

Automation-Orchestration

the phase change of modern software

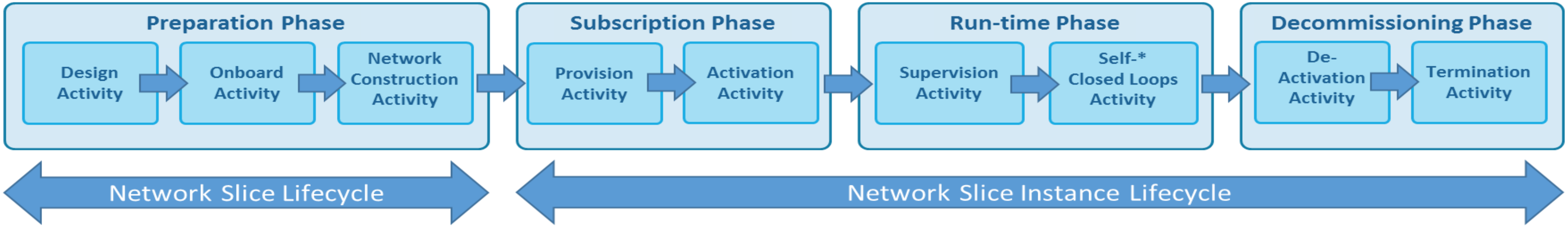
DevOps



Free software is becoming expensive

Automation-Orchestration

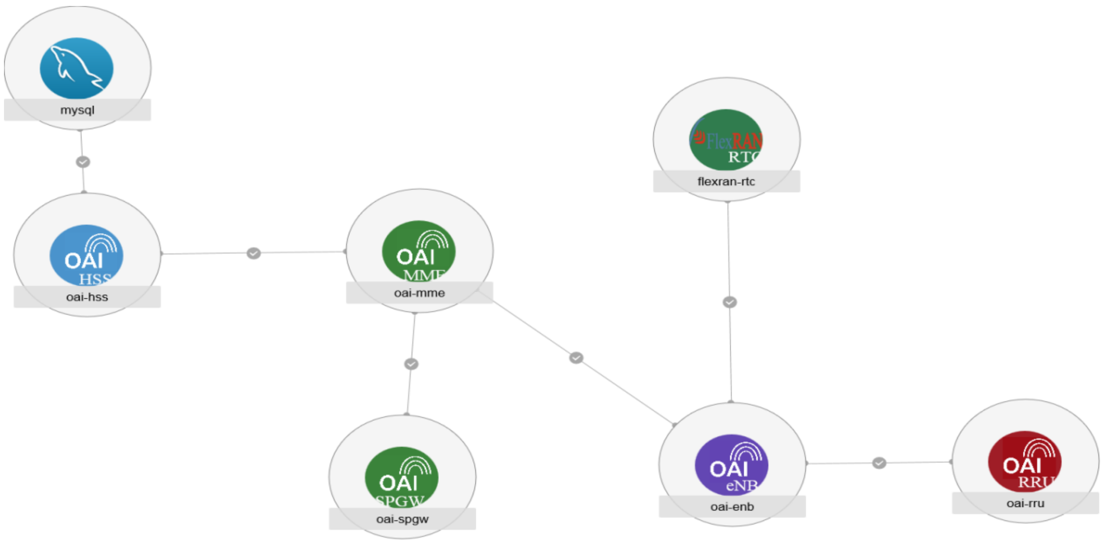
LifeCycle Management (Encapsulate operation)



Automation-Orchestration

LifeCycle Management (Encapsulate operation)

DevOps

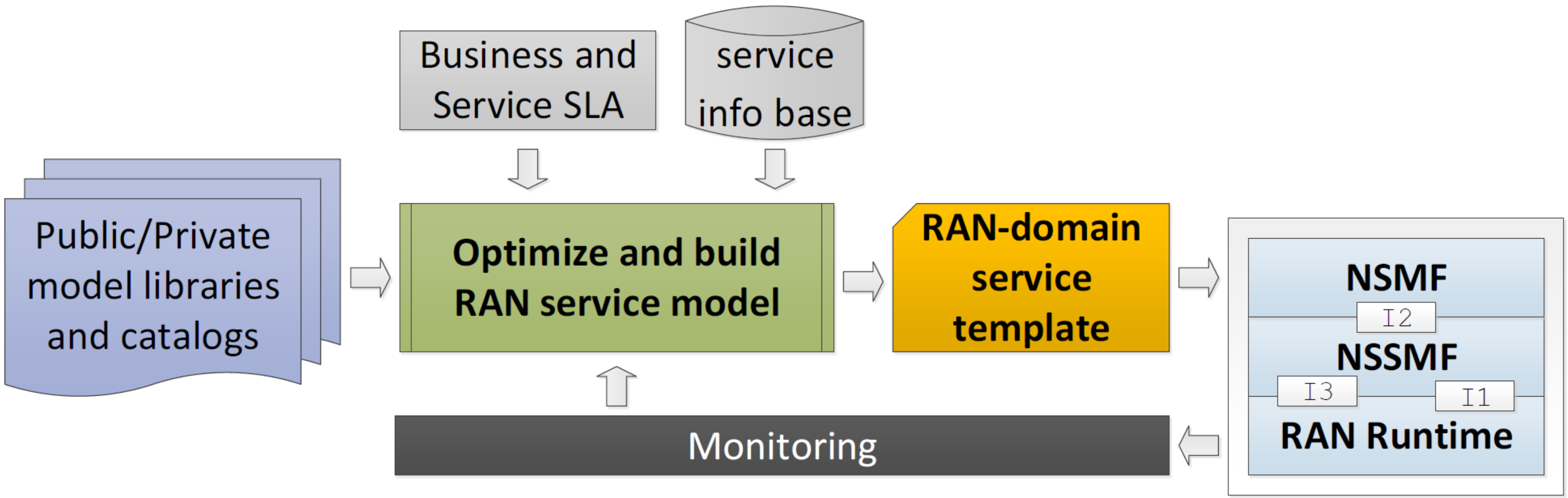


- → installation
- → configuration
- → connections
- → upgrades and updates
- → scale-out and scale-back
- → health checks
- → operational actions
- → benchmarks

https://jujucharms.com/q/oai

Automation-Orchestration

Dynamic Service updates to maintain and optimize performance/SLA

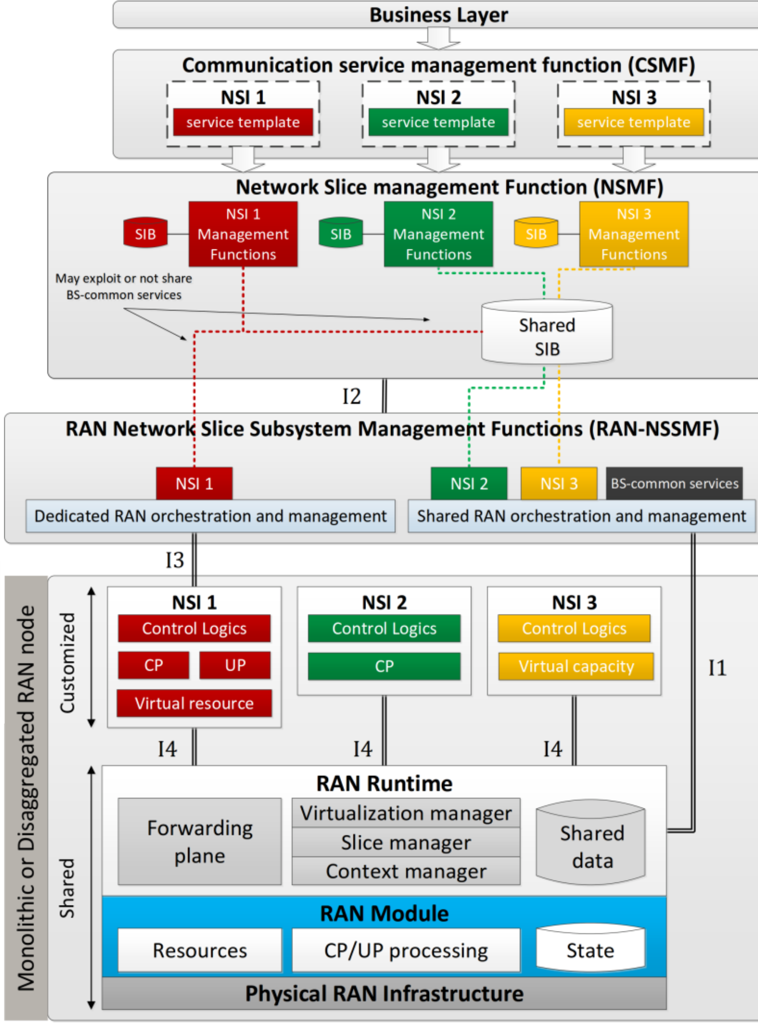


Automation-Orchestration

Slices life cycle management

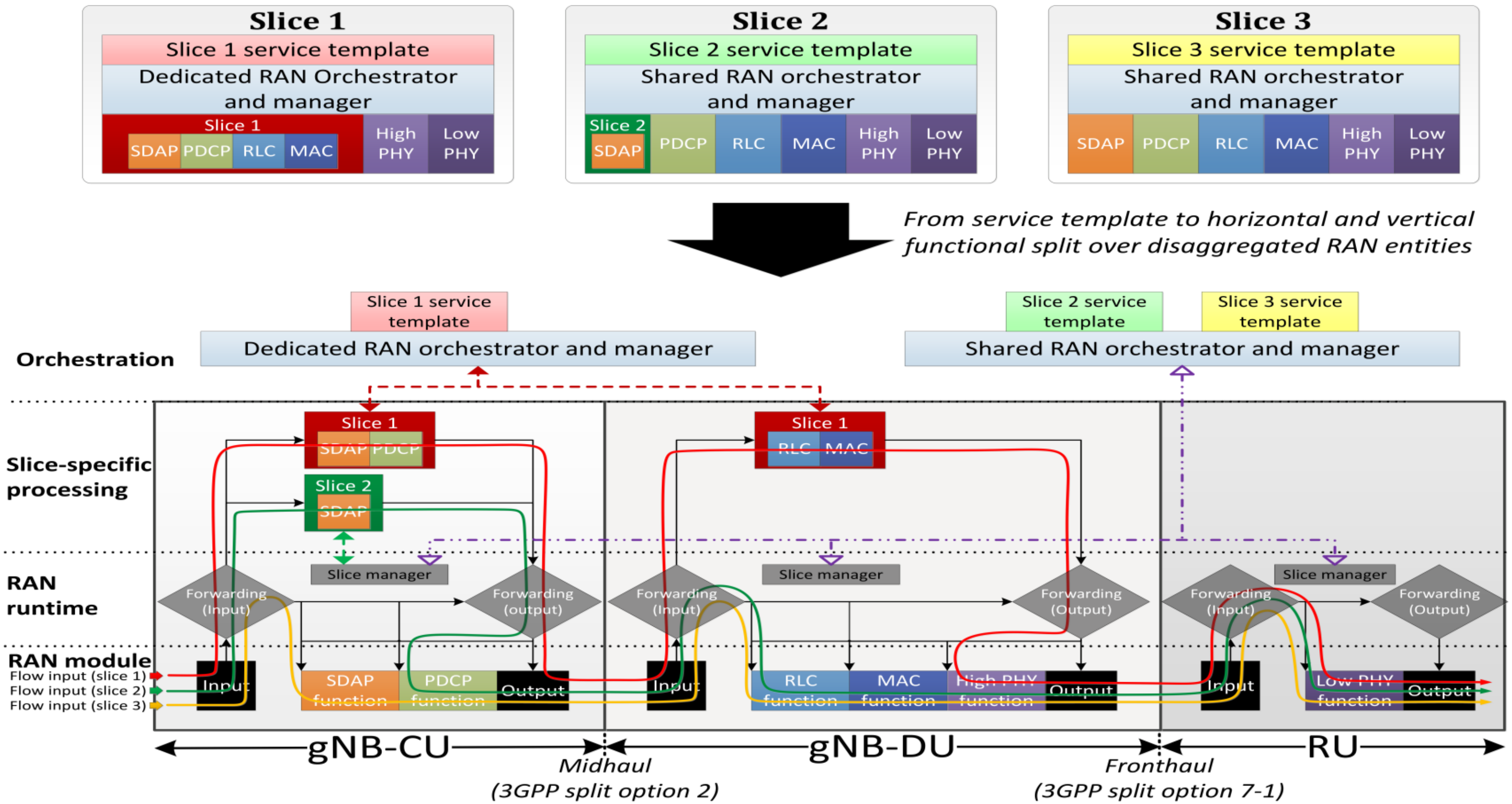
- Multi-service chaining of customized and/or shared NFs
- Multi-service chain placement
- Four interfaces I1 to I4

Interface	Description
I1	Expose active RAN runtime services and retrieve messages for monitoring and feedback.
I2	Subscribe for the slice-specific events and populate SIB accordingly.
I3	Customize management and monitoring for each slice and indicate any changes in underlay RAN.
I4	Register a slice to the RAN runtime and consume RAN runtime service as a separate process.



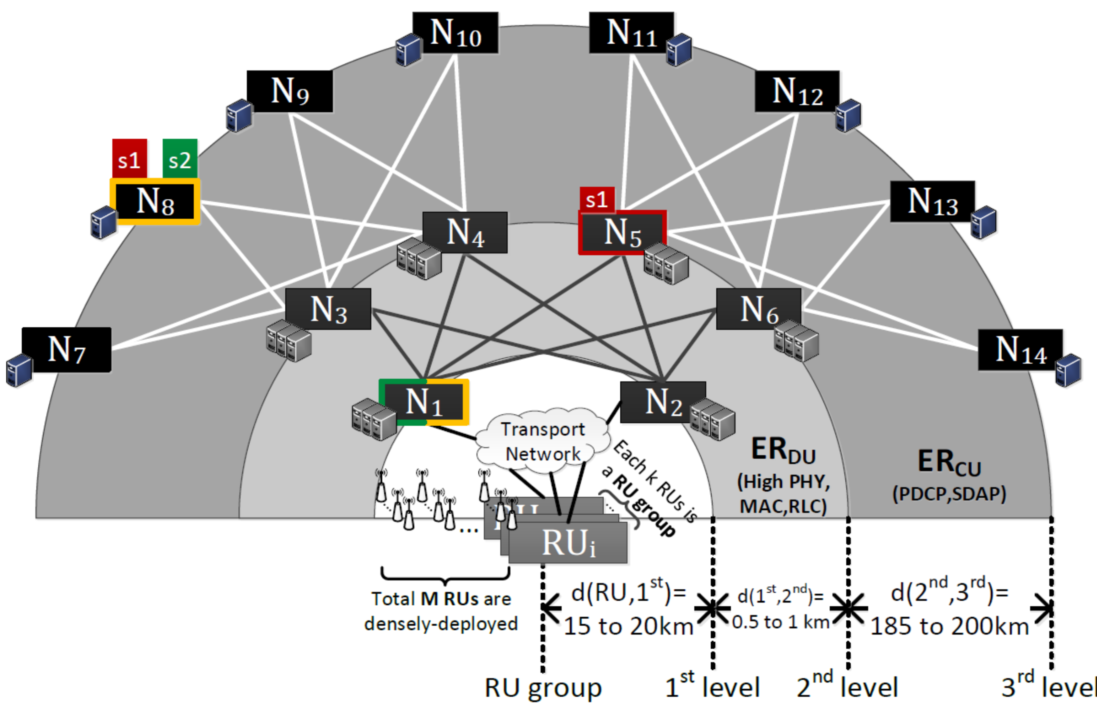
[1] C-Y. Chang, N. Nikaein, et al., Slice Orchestration for Multi-Service Disaggregated Ultra Dense RANs, IEEE Communication Magazine, 2018.
[2] O. Arouk, N. Nikaein, and T. Turletti, Multi-objective placement of virtual network function chains in 5G, IEEE CloudNet 2017.

Horizontal and vertical service composition



Multi-service chaining

Respect both service requirements and operator objective
2-stage placement algorithm: Shared → Customized



Step 1: $ER_{DU}=\{N_1,N_2,N_3,N_4,N_5,N_6\}$
$ER_{CU}=\{N_7,N_8,N_9,N_{10},N_{11},N_{12},N_{13},N_{14}\}$

Step 2: $CG_{DU}=\{N_1,N_2\}$
$CG_{CU}=\{N_7,N_8,N_9,N_{10},N_{11},N_{12},N_{13},N_{14}\}$

Step 3: $BN_{DU}=N_1$
$BN_{CU}=N_8$

Step 4: $BN_{DU,slice1}=N_5$
$BN_{DU,slice2}=BN_{DU,slice3}=N_1$
$BN_{CU,slice1}=N_8$
$BN_{CU,slice2}=N_8$
$BN_{CU,slice3}=N_8$

Overall path for each slice

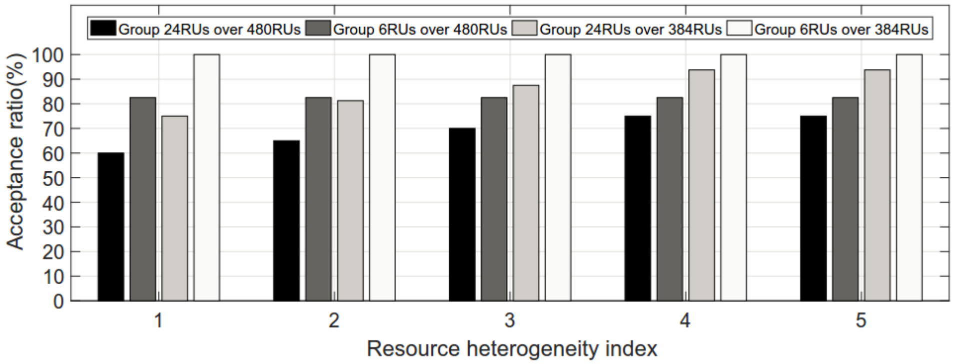
Slice 1: $N_8 \rightarrow N_1$ (via N_3 or N_4) $\rightarrow N_5 \rightarrow N_1 \rightarrow$ RU
Slice 2: $N_8 \rightarrow N_1$ (via N_3 or N_4) $\rightarrow$ RU
Slice 3: $N_8 \rightarrow N_1$ (via N_3 or N_4) $\rightarrow$ RU

Multi-service placement

Enhance acceptance ratio in multi-service placement

- Utilize a smaller group size of RUs (e.g. 6 RUs) as the placement granularity

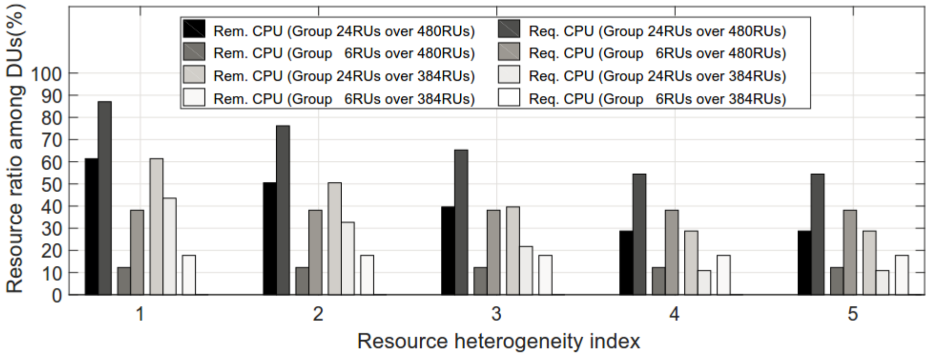
- Provision heterogeneous resources (i.e., index 2 to 5) base on service requirements



Actions

- Scale-up: Reallocate the unused resources to a subset of nodes (e.g., 384 RUs grouped in 24)

- Scale-out: Provision more nodes (e.g., 480 RUs)



Multi-service placement

Application Plane

Each slice is a composition of CP/UP processing coupled with a set of control applications responsible to control the behavior and manage the state

Control applications, shared or dedicated, may be chained together to perform the desired operation

(e.g. monitoring and load balancing)

Each control app can be **self-contained**: own execution environment, own process and lifecycle, and own instance of SDK,

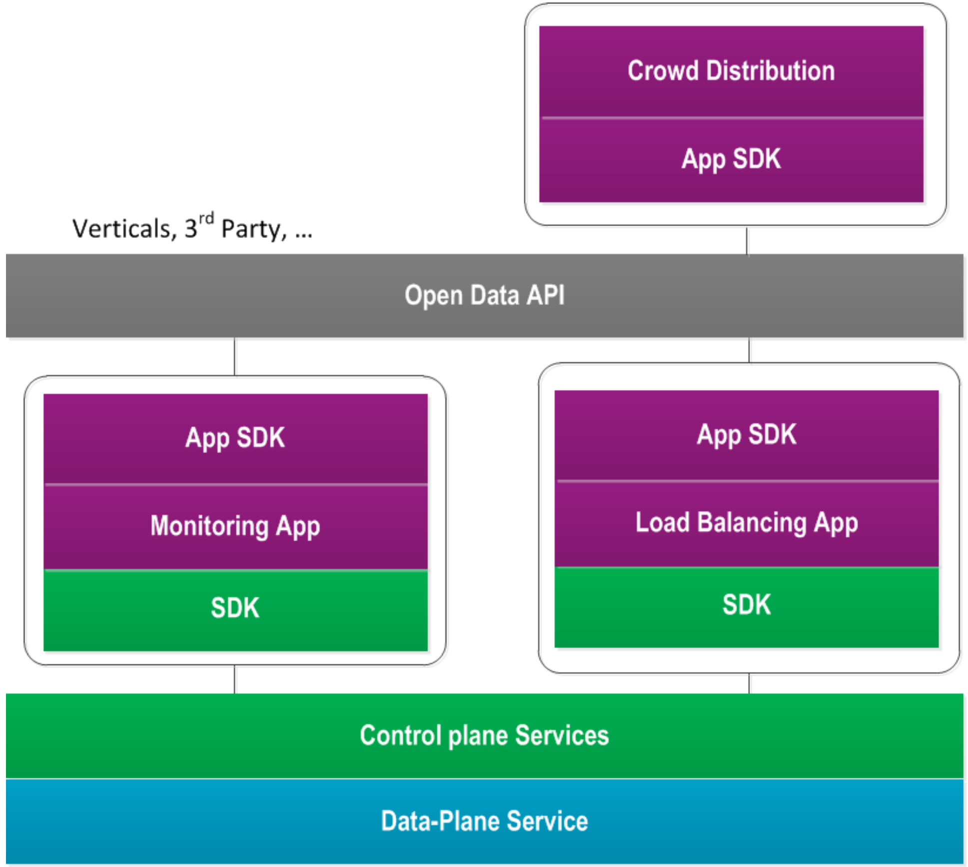
Slice programmability

Application Plane

(1) Enable automation and extendibility of the network control operations

(2) Improves decision making process across different slices

(3) Network graph DB and partitioning for multiple substrates



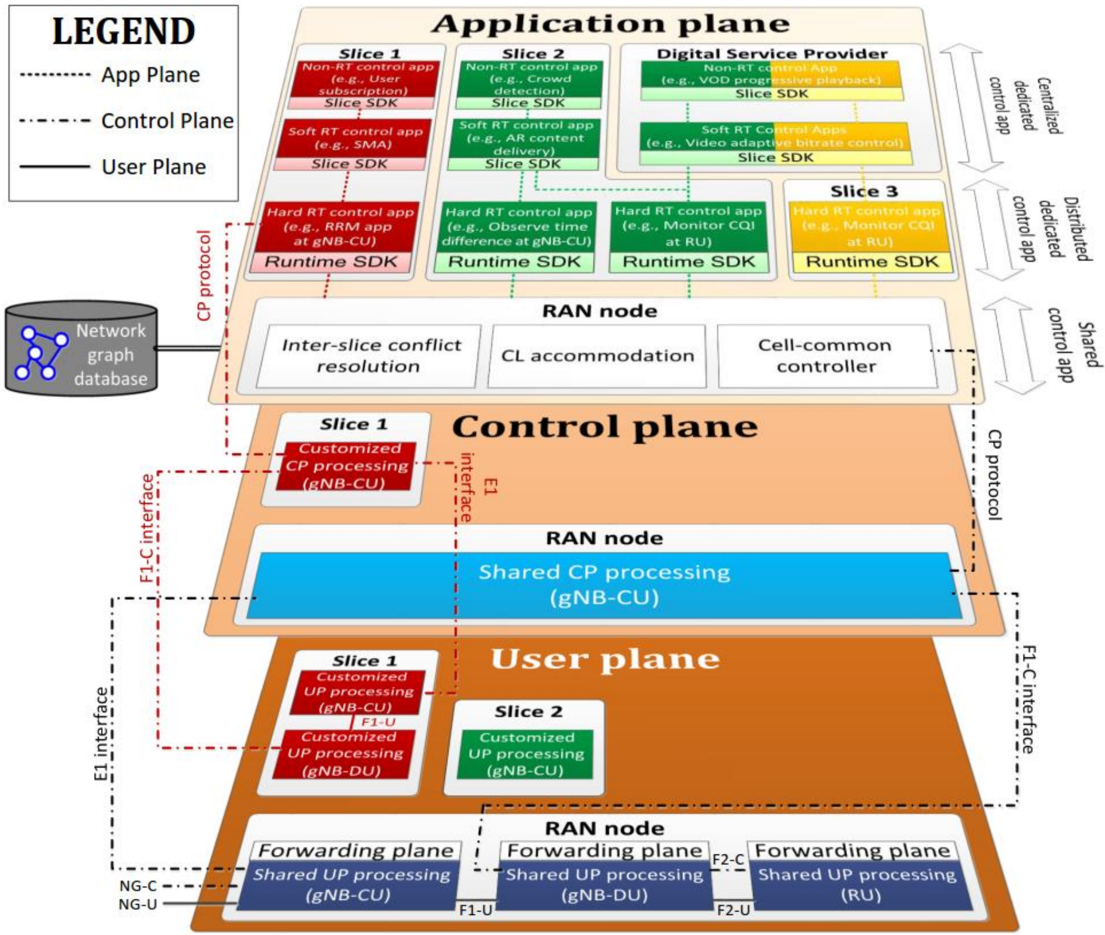
Slice programmability

Application Plane

(1) Enable automation and extendibility of the network control operations

(2) Improves decision making process across different slices

(3) Network graph DB and partitioning for multiple substrates



[1] C-Y. Chang, N. Nikaein, Enabling Network Application for Multi-Service Programmability in a Disaggregated RAN, IEEE Communication Magazine, 2018.

Slice programmability

Plug and play application plane

Chaining shared and dedicated control Apps on per slice basis

Two level of Abstractions and SDKs

Slice programmability

Abstract the underlying network and controller by means of high-level technology-agnostic level APIs

Decouple control logic from data plane actions following SDN principles

Aggregated and structured network config, status, topology information in form of instantaneous network graphs

Facilitate the development of network control apps (extendable, coordination

Semantic modeling of the underlying networks

Software-development kit

Chaining of Slice User Management and RRM Apps

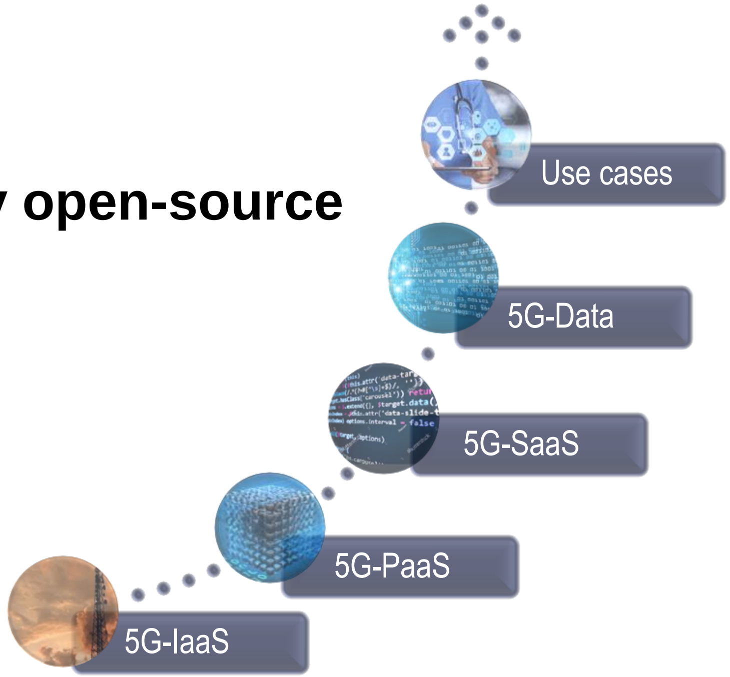
Super user | Platinum user | Gold user | Silver user | Bronze user

Slice programmability

OpenSource Platforms

Need for agile network service delivery platforms and use-cases for 4G-5G R&D

5G Innovations empowered by open-source



Opensource Platforms

Agile network service delivery platforms

LL-MEC

A Low Latency SDN-based MEC Platform

FlexRAN

A Flexible & Programmable SD-RAN Platform

JOX

An event-driven juju-based service orchestrator core

Store

Network function & application distribution Repository

Open5G Lab

Access to 4G/5G network facilities and perform experiments

My Project?

Create a project and build your use-case

Mosaic-5G.io Ecosystem

Agile network service delivery platforms

Open5G Lab

JOX Orchestrator

Store

FlexRAN Controller

LowLatency MEC

OpenAirInterface RAN

OpenAirInterface CN

MOSAIC 5G

OPEN AIR INTERFACE

Mosaic-5G.io Ecosystem

MWC 2016, 2017

ITU, FG-13, 2016, 2017

ETSI 2016, 2017

EUCNS 2015, 2016, 2017

OPNFV 2016

Mobicom 2014,2016,2017

Success Stories

Mail : contact@mosaic-5g.io

Website : mosaic-5g.io

Linkedin:
https://www.linkedin.com/in/mosaic-5g

Twitter: @mosaic5g

Info

Conclusion

5G and beyond is not only New Radio and verticals, it is also evolution in computing for wireless networks: Central offices becoming data-centers

Centralized computing and storage using more general-purpose equipment (Intel servers)

More and more software technologies from cloud-computing (NFV,SDN,MEC, etc.) jointly with radio signal processing

Applicable to lesser extent for existing and evolving 4G radio
$\longrightarrow$ Fusion of Information and Cellular technologies

Conclusion

Increased interest in understanding (managing?) the role of open-source communities by ITU, NGMN, ETSI

Main issue: patent-pool licensing

Conclusion

Facebook and Google are quickly entering the datacenter Telco space

Value-chain of Telecom is under siege and may become very different because of this

Example: https://telecominfraproject.com/

- Low-cost equipment for rural areas
- Federating open-source developers

Conclusion

RAN slicing is an on-going research with several challenges Isolation, Sharing, Customization

Satisfy requirements from both slice owner and operator

Two main solutions: ORION and RAN runtime slicing systems

Conclusion

Slice orchestration/management for multi-service

Interfaces between RAN and 3GPP management functions

Auto-scaling operations to enhance acceptance ratio in chaining & placement

Runtime SDK and slice SDK for 2-level abstractions

Single/cross-domain control application chaining

Conclusion

Realtime control and coordination in RAN and CN

Tradeoff between slice isolation and resource sharing

Security control across many logical networks and abnormally detection

Pattern recognition and correlation to support QoS-QoE

Predict network behavior if a given control logic is applied

Automate failover and network health monitoring and prediction

Dynamic guarantees as a function of cost /adaptive/probabilities

Example Research Areas

Why such a big complexity to support slicing?

How the net neutrality principles be retained ?

Two questions in Slicing


Q
A

Email: navid.nikaein@eurecom.fr

Website: http://www.eurecom.fr/~nikaeinn/

Linkedin: https://www.linkedin.com/in/navidnikaein

Tel: +33.(0)4.93.00.82.11

Contact Information