A Primer on 5G Network Slicing:
Concepts, Algorithms, and Practice

Pablo Serrano

http://www.it.uc3m.es/pablo/

uc3m | Universidad Carlos III de Madrid

Outline

- 5G Network Slicing in 3 slides
- Why we need network slicing
- Enablers & Ancestors
- Network Slicing novelties
- “Definition” of 5G
- Network slicing in practice
- Research problems

About the speaker

Dr. Pablo Serrano
Associate Professor, Univ. Carlos III de Madrid
Mail: pablo@it.uc3m.es TW: @pbluc3m
Web: http://www.it.uc3m.es/pablo/
Past visiting positions at: Univ. Massachusetts Amherst, Univ. of Edinburgh, Trinity College Dublin, Telefónica R+D Barcelona, Univ. Brescia.

Current projects:

- 5G-VINNI: 5G Verticals INNovation Infrastructure
- 5G-MoNArch: 5G Mobile Network Architecture for diverse services, use cases, and applications in 5G and beyond
- TIGRE5-CM: Tecnologías Integradas de gestión y operación de Red 5G

Recent Past projects:

- Flex5Gware: Flexible and efficient hardware/software platforms for 5G network elements and devices
- FLAVIA: FLexible Architecture for Virtualizable future wireless Internet Access

What is 5G?

Wikipedia
The Free Encyclopedia

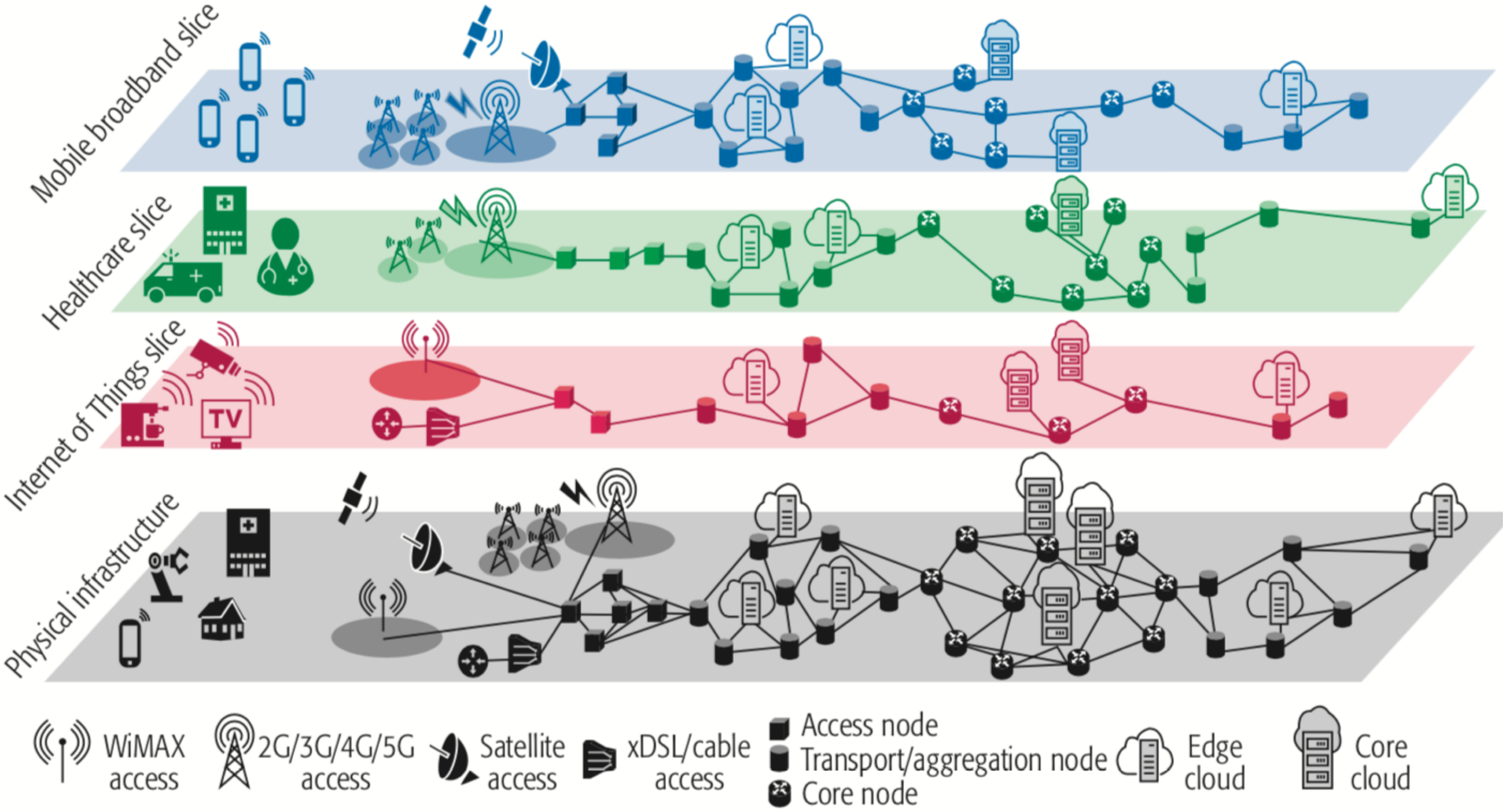
"***5G*** is a marketing term for some new mobile technologies.[*according to whom?*] Definitions differ[*citation needed*] and confusion is common[*citation needed*]. The ITU IMT-2020 standard provides for speeds up to 20 gigabits per second and has only been demonstrated with millimeter waves of 15 gigahertz and higher frequency. The more recent 3GPP standard includes any network using the NR New Radio software." (Sept. 4th, 2018)

What is Network Slicing?

- **Network Slice:** A set of infrastructure resources and service functions that has attributes specifically designed to meet the needs of an industry vertical or a service
- **Network Slicing:** A management mechanism that Network Slice Provider can use to allocate dedicated infrastructure resources and service functions to the user of the Network Slice

From: "Architecture of Network Slicing Management," L. Geng et al, IETF 102

What is 5G Network Slicing?



From: J. Ordonez-Lucena, P. Ameigeiras, D. Lopez, J. J. Ramos-Munoz, J. Lorca and J. Folgueira, "Network Slicing for 5G with SDN/NFV: Concepts, Architectures, and Challenges," in *IEEE Communications Magazine*, vol. 55, no. 5, pp. 80-87, May 2017.

WHY WE NEED SLICING

Traffic trends

- *Essential* services will continue to proliferate
- **Novel very-demanding** services (augmented reality) delivered over mobile
- 1000x traffic increase over the next decade [1]
- Plus: **50 billion connected devices** by 2020 [2]

[1] "Global Mobile Data Traffic Forecast Update", *2010-2015 White Paper*, Feb. 2011.
[2] "More than 50 Billion Connected Devices", *White Paper*, Feb. 2011

From: 5G-MoNArch Project video
https://www.youtube.com/watch?v=y6b9FNniPuQ

Requirements

- **Stringent** latency and reliability requirements
  – Verticals: Health-care, security, logistics, automotive, mission-critical control
- **Wide range** of data rates, up to 10 Gbps, and 100 Mbps with high availability and reliability
- Network **scalability and flexibility** to support a large # of devices with very long lifetime
- And the ability to **quickly instantiate** novel services anywhere

In Figures

	4G IMT-A		5G IMT-2020
Practical user transfer speed	10Mbps	High-res video streaming (HD, 4K, 8K...)	100Mbps~1Gbps
Peak transfer speed	1Gbps	High-capacity file download (DVD → BD class)	20Gbps
Max. device movement speed	350km/h	High-speed transportation (e.g. maglev/linear-motor vehicles)	500km/h
Permissible latency (wireless areas)	10ms	Even lower latency (AR, fighting games, remote-controlled medical treatment)	1ms
Device connection density	$10^5/km^2$	Popularization of IoT (17% annual growth rate)	$10^6/km^2$
Area traffic capacity	$0.1Mbps/m^2$	Increase in traffic (40%+ annual growth rate)	$10Mbps/m^2$

Full HD 1080 ▶ 4K ▶ 8K

6Mbps 25Mbps 100Mbps

High-res video streaming

DVD ▶ Blu-ray Disc

4.7GB 50GB

Online content 10 times the quality

SCMaglev (500km/h)

Low delay (fighting games, AR)

Usage scenarios

- **Enhanced Mobile Broadband** to deal with hugely increased data volumes, overall data capacity and user density
- **Massive Machine-type Communications** for the IoT, requiring low power consumption and low data rates for very large numbers of connected devices
- **Ultra-reliable and Low Latency Communications** to cater for safety-critical and mission critical applications

From: https://www.etsi.org/technologies-clusters/technologies/5g

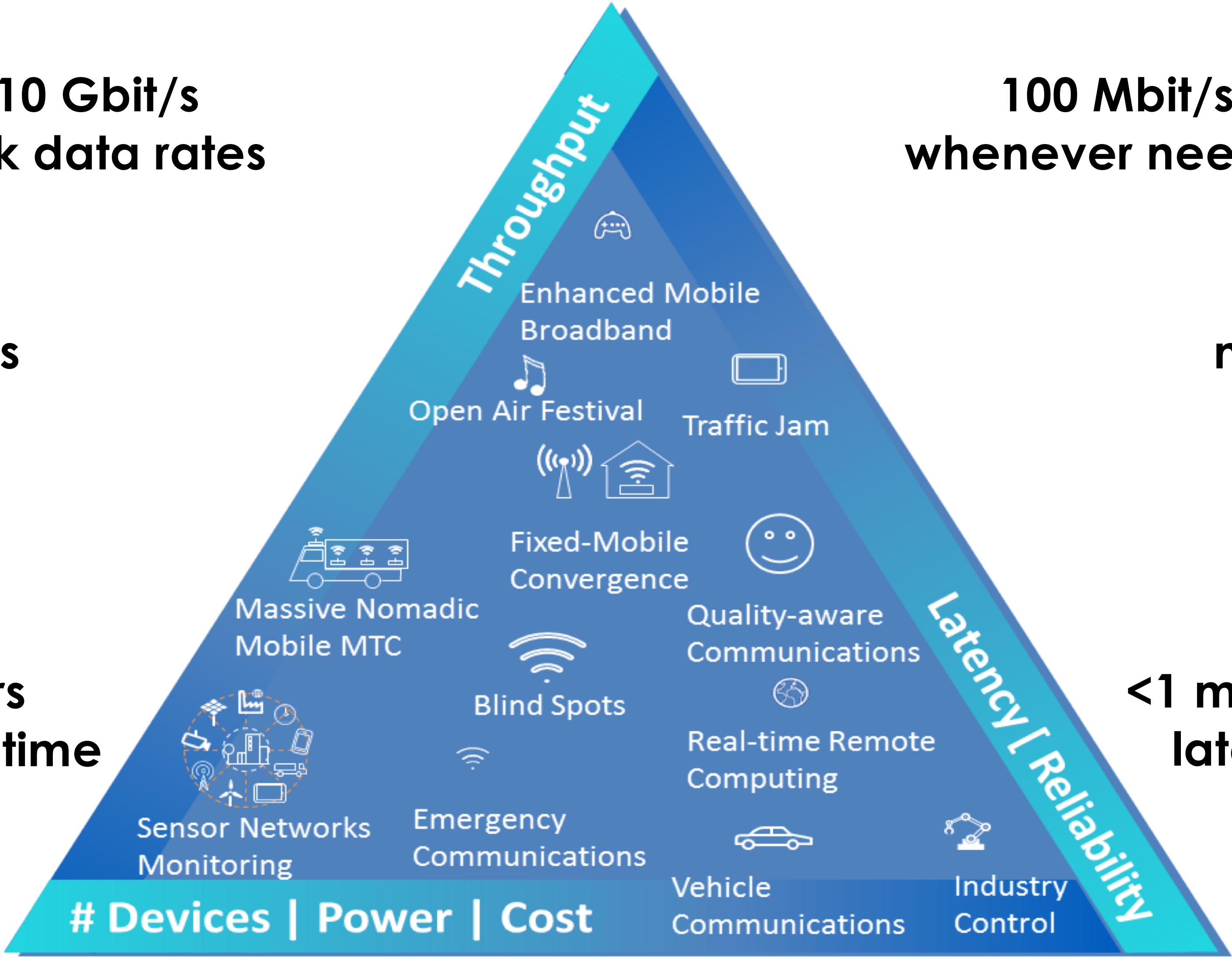
Requirements: Heterogeneity

>10 Gbit/s
peak data rates

10 – 100 x
more devices

M2M
ultra low cost

10 years
battery lifetime



100 Mbit/s
whenever needed

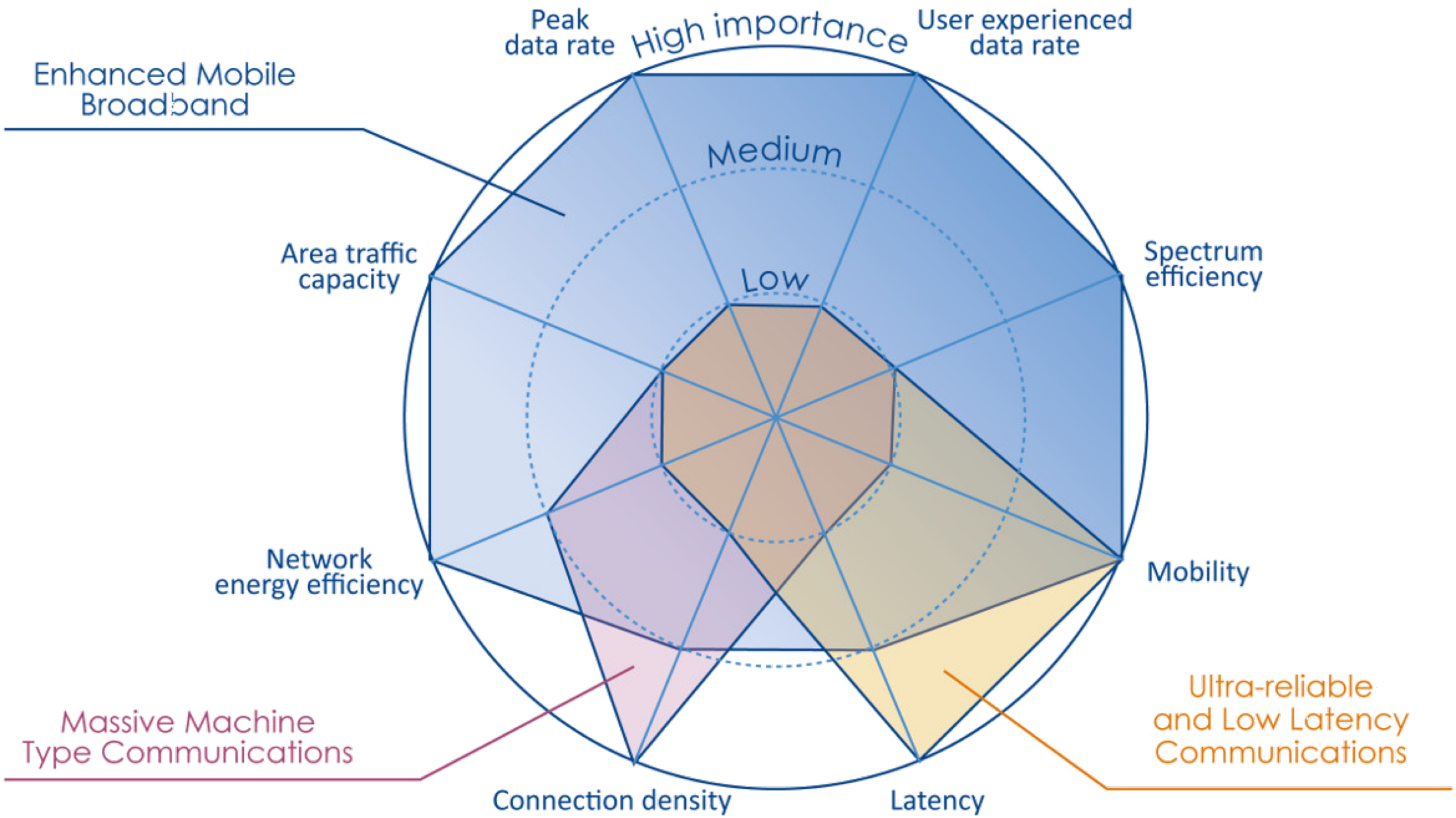
10 000 x
more traffic

ultra
reliability

<1 ms radio
latency

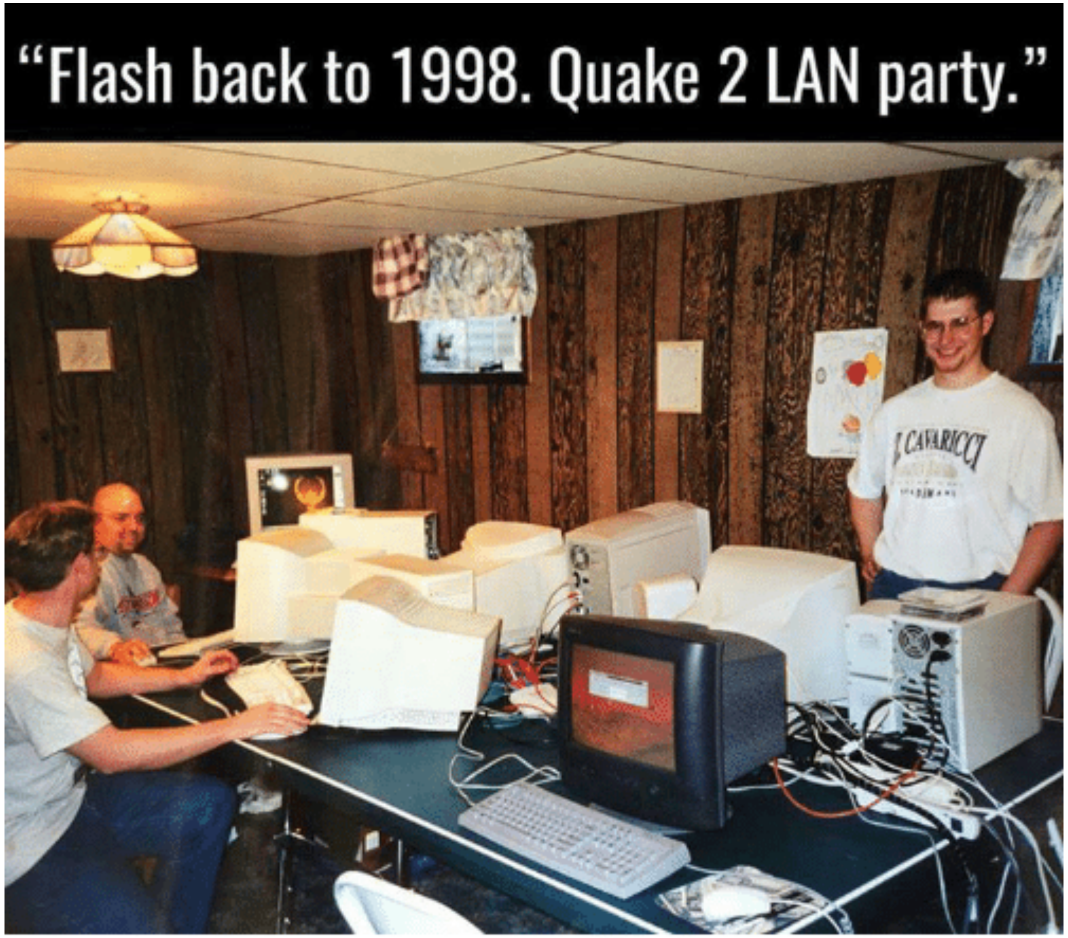
From: 5G-MoNArch Project video
https://www.youtube.com/watch?v=y6b9FNniPuQ

Usage Scenarios

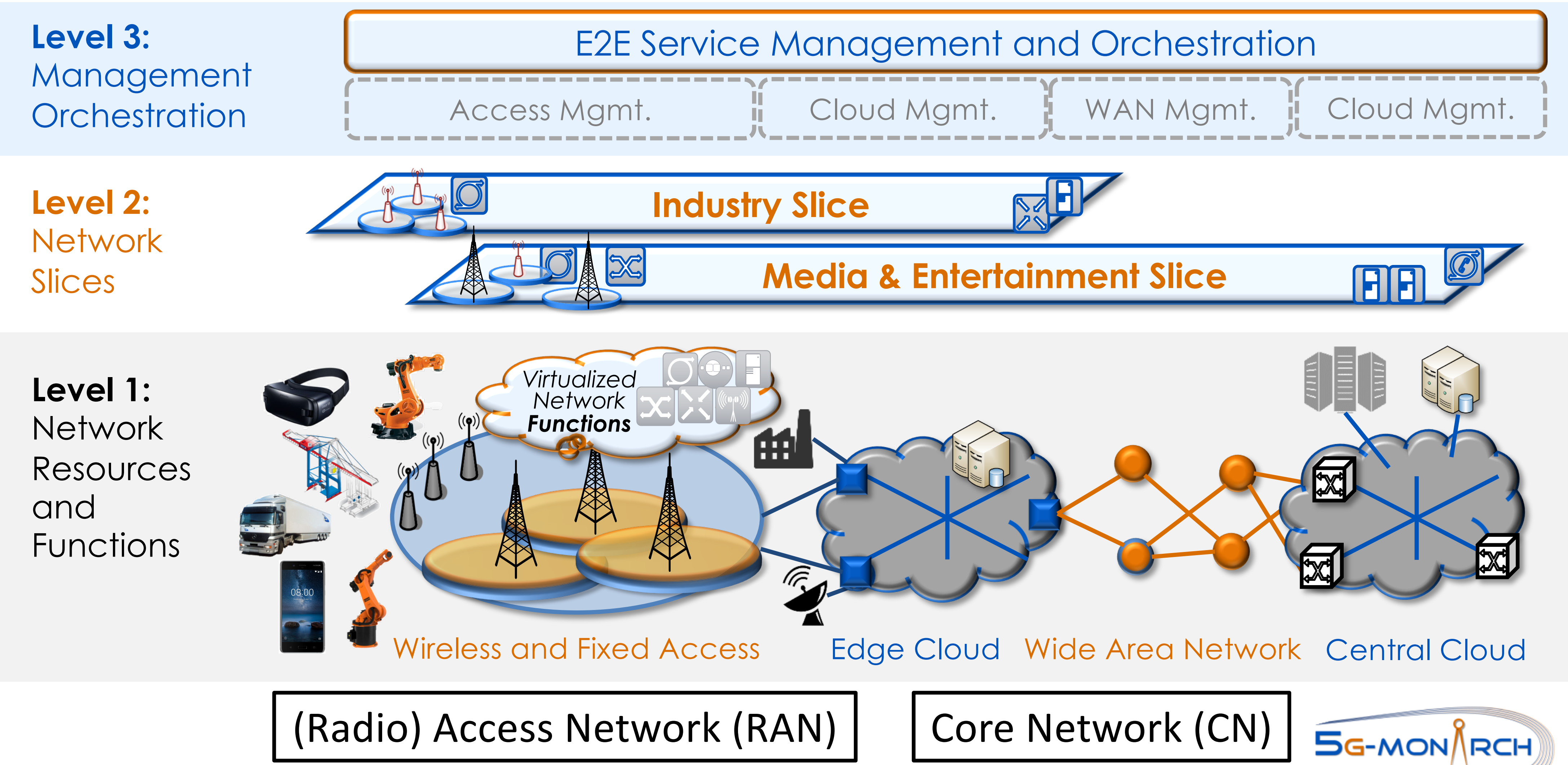


Challenge: in one network

- It's "easy" to deploy ad hoc networks



Customization over one network



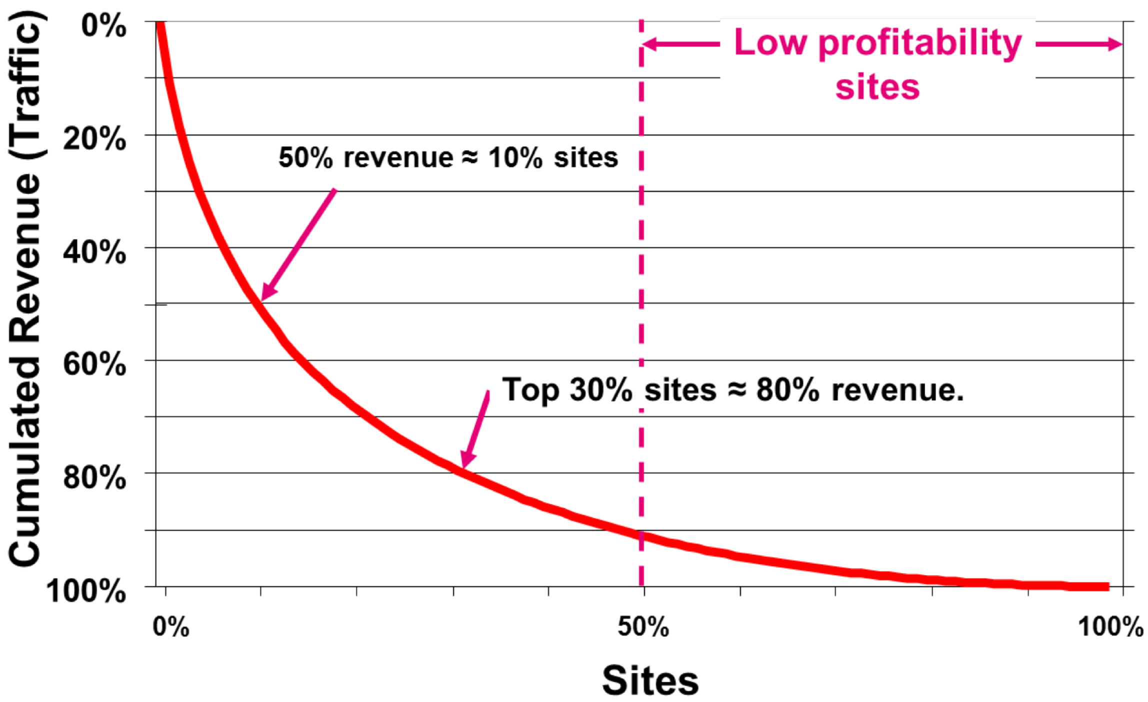
From: 5G-MoNArch Project video
https://www.youtube.com/watch?v=y6b9FNniPuQ

NS: The 3GPP definition

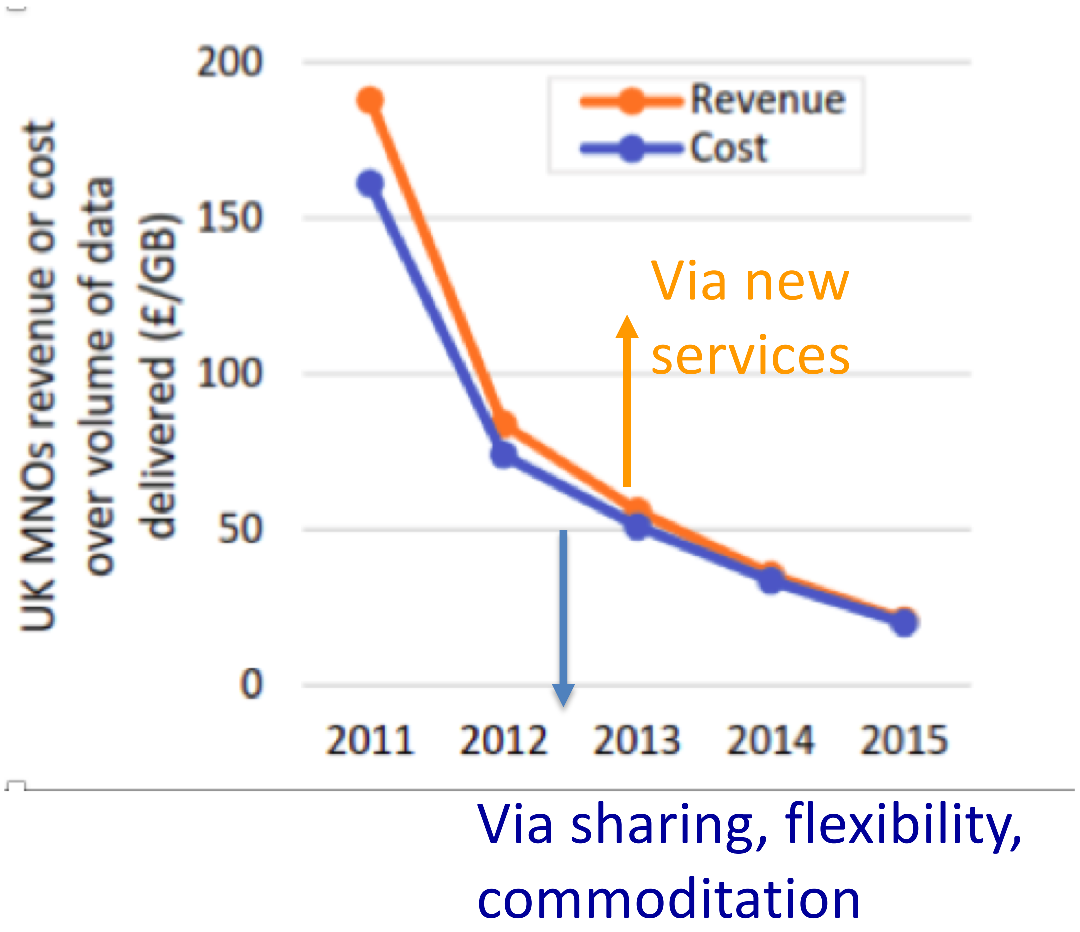
- “A logical **network** that provides specific network capabilities and network characteristics”
- “A network created by the operator **customized to provide an optimized solution** for a specific market scenario which demands specific requirements with end to end scope”
- Implemented by “slice instances”
- Created from a “network slice template”

Motivation

- Share the underutilized resources
  - 50% of revenue created by < 10% of sites
  - Diverse traffic patterns/mobility cause **resource underutilization**



Motivation



From: Source: 5GPPP 5G-NORMA project

ANCESTORS/ENABLERS: SHARING THE NETWORK

Sharing the Network

- The concept of separated (virtual) networks deployed over a single network is not new
  - However, there are specificities
- Virtualization of Networks
  - Software Defined Networking: SDN
- Virtualization of Machines
  - Network Function Virtualization: NFV

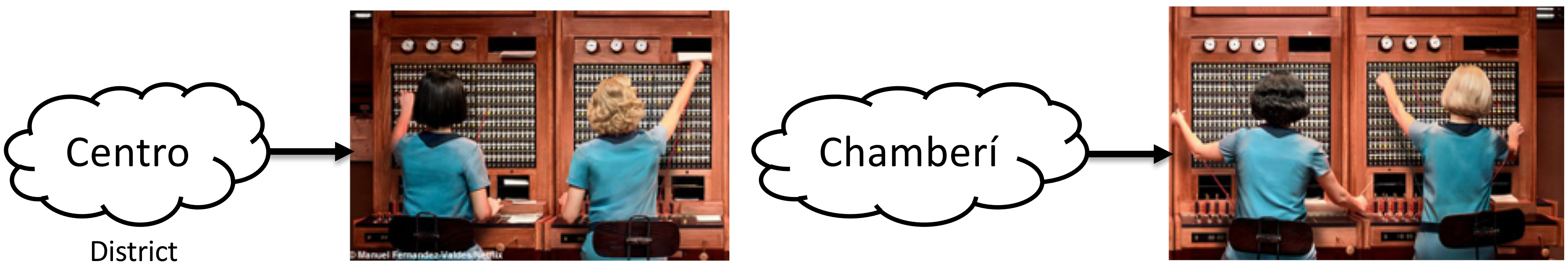
Back in the day: Circuit Switching

One service (voice)
Two users

Multiple users
Shared resources

Sharing gains

- Performance (blocking probability, service times) improves when resources are merged

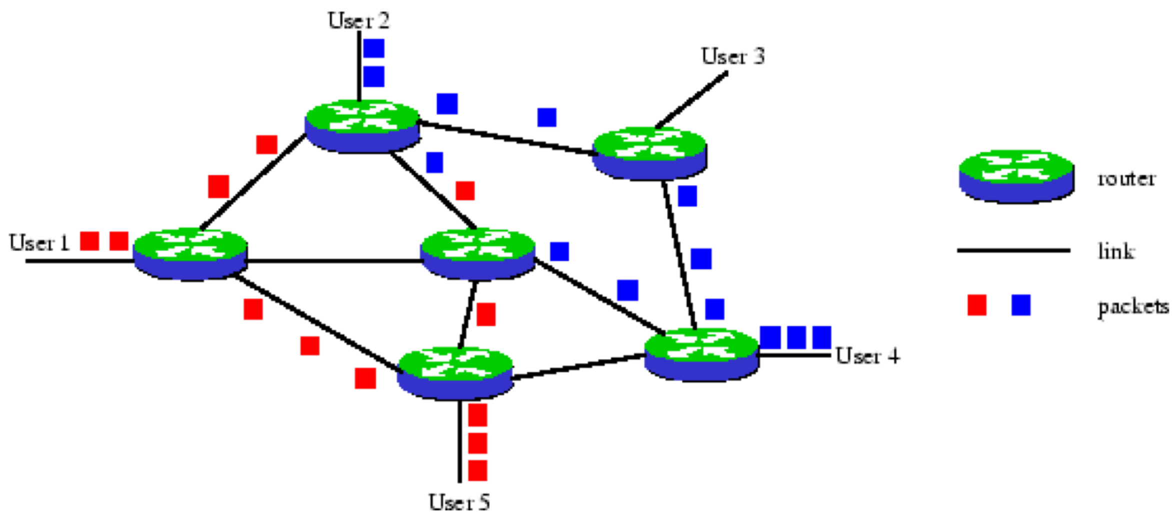


Circuit switching: additional notes

- Specialized equipment
  - PSTN: Rotary 7 A 2
  - GSM: Ericsson MSC-S
  - *De facto* vendor lock-in
- “Simple” ecosystem
  - One/few services
  - One network = provider
- Slow service development

Packet switching

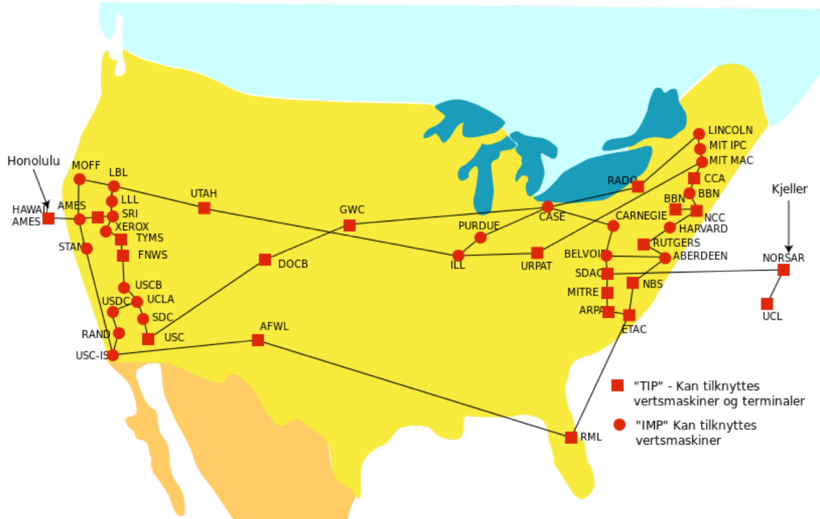
- A stream is divided into packets



- A channel is shared by a number of data streams (statistical multiplexing)
  – Variable-bitrate streams: flexibility

Packet switching

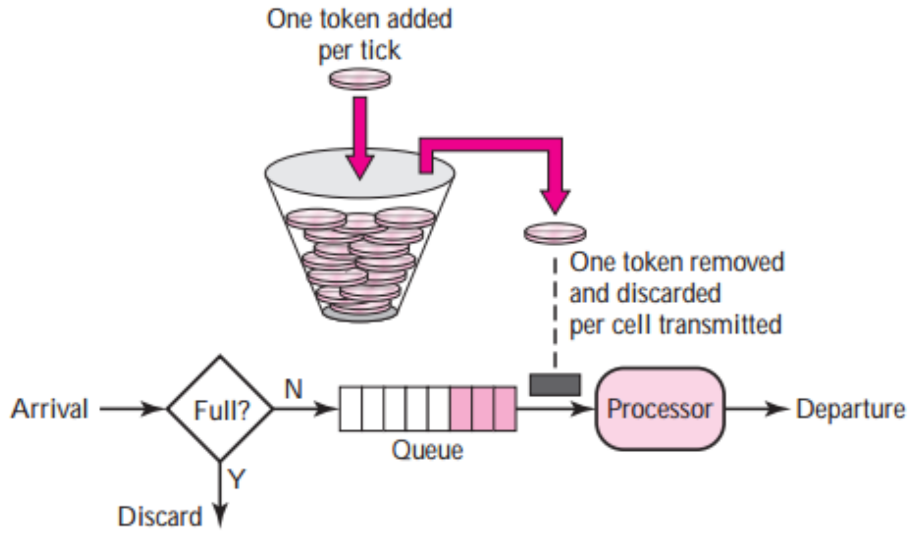
- ARPANET: 1969
- TCP/IP: 1982
- (Etc.)
- **Multi-service**



- Best effort delivery: no guarantee (packet may be lost, duplicated, re-ordered)
  - OK for non real-time services (email, ftp, web)
  - What about real-time traffic?

Providing QoS to Het. Serv.

- TCP/IP: two proposals
  - Diffserv: coarse-grained
    - Low, medium, high priority.
  - Intserv: fine-grained
    - E2E reservations



- ATM: much focus on traffic engineering
  - Unspecified, Variable, Constant Bit Rate.
  - Peak, Sustainable Cell Rate
  - Maximum Burst Size

Many research papers!

Quality of Service: challenges

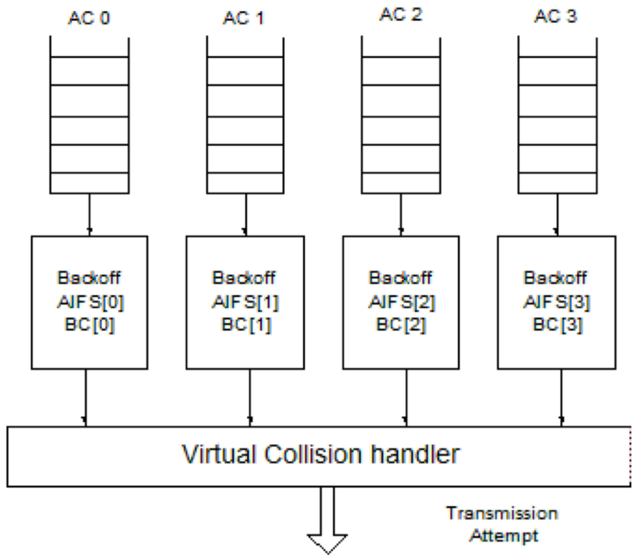
- On the Taxonomy
  - Many categories: complexity (few implemented)
  - Few categories: not useful
- On the Management
  - Intra-network might work
  - Inter-network QoS: guarantees, monitoring
- Usual Approach: add more bandwidth
  - Ethernet vs. ATM

Andrew Odlyzko, “Data networks are lightly utilized, and will stay that way,”. AT&T Labs October 7, 1998.

For Wireless Networks

- 3GPP defines QoS to differentiate services
  - 15 classes (e.g., conversational voice, gaming, V2X)

- IEEE 802.11 – WiFi
  - DCF vs. PCF (not used)
  - EDCA, little adoption
    - Best effort, Background, Voice, Video



QCI

- QCI: QoS Class Identifiers
  - vendors can all have the same understanding
  - 3GPP Technical Specification 23.203, Policy and charging control architecture (Release 8)

QCI	RESOURCE TYPE	PRIORITY	PACKET DELAY BUDGET (MS)	PACKET ERROR LOSS RATE	EXAMPLE SERVICES
1	GBR	2	100	10^{-2}	Conversational voice
2	GBR	4	150	10^{-3}	Conversational video (live streaming)
3	GBR	5	300	10^{-6}	Non-conversational video (buffered streaming)
4	GBR	3	50	10^{-3}	Real-time gaming
5	Non-GBR	1	100	10^{-6}	IMS signaling
6	Non-GBR	7	100	10^{-3}	Voice, video (live streaming), interactive gaming
7	Non-GBR	6	300	10^{-6}	Video (buffered streaming)
8	Non-GBR	8	300	10^{-6}	TCP-based (for example, WWW, e-mail), chat, FTP, p2p file sharing, progressive video and others
9	Non-GBR	9	300	10^{-6}	

Sharing: Wrap-up

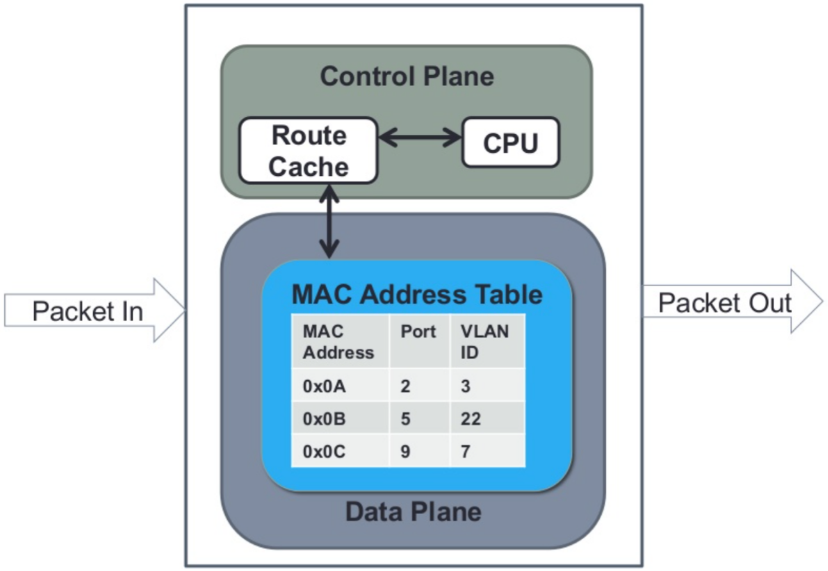
- Sharing common resources across multiple users is **inherent** to networking
- Circuit switching: multi-user
  - Guaranteed service (reserved resources)
- Packet switching: multi-service
  - Multiple services, no QoS (statistical multiplex.)
- Challenges
  - Service definition, inter-domain,
  - **Adoption** ("Why bother?")

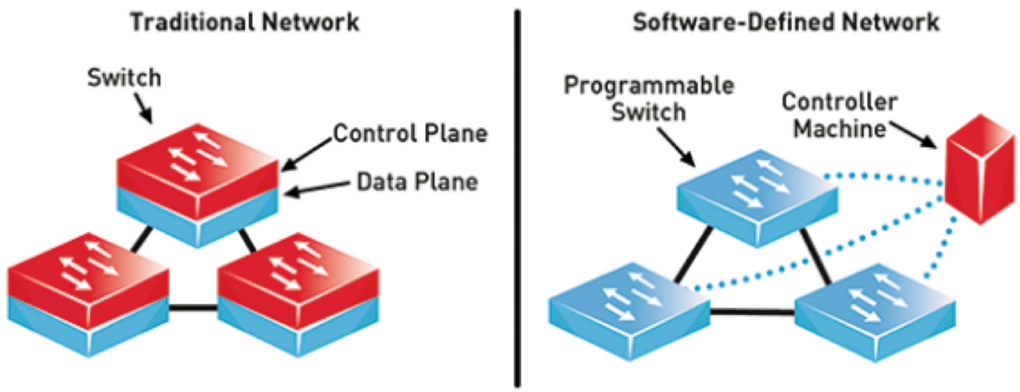
“Why Software is eating the world” Marc Anderseen, WSJ, 2011

ANCESTORS/ENABLERS: SOFTWARIZATION

Software Defined Networking

- "OpenFlow: a way for researchers to run experimental protocols in the networks they use every day."
  - Motivation: large-scale trials

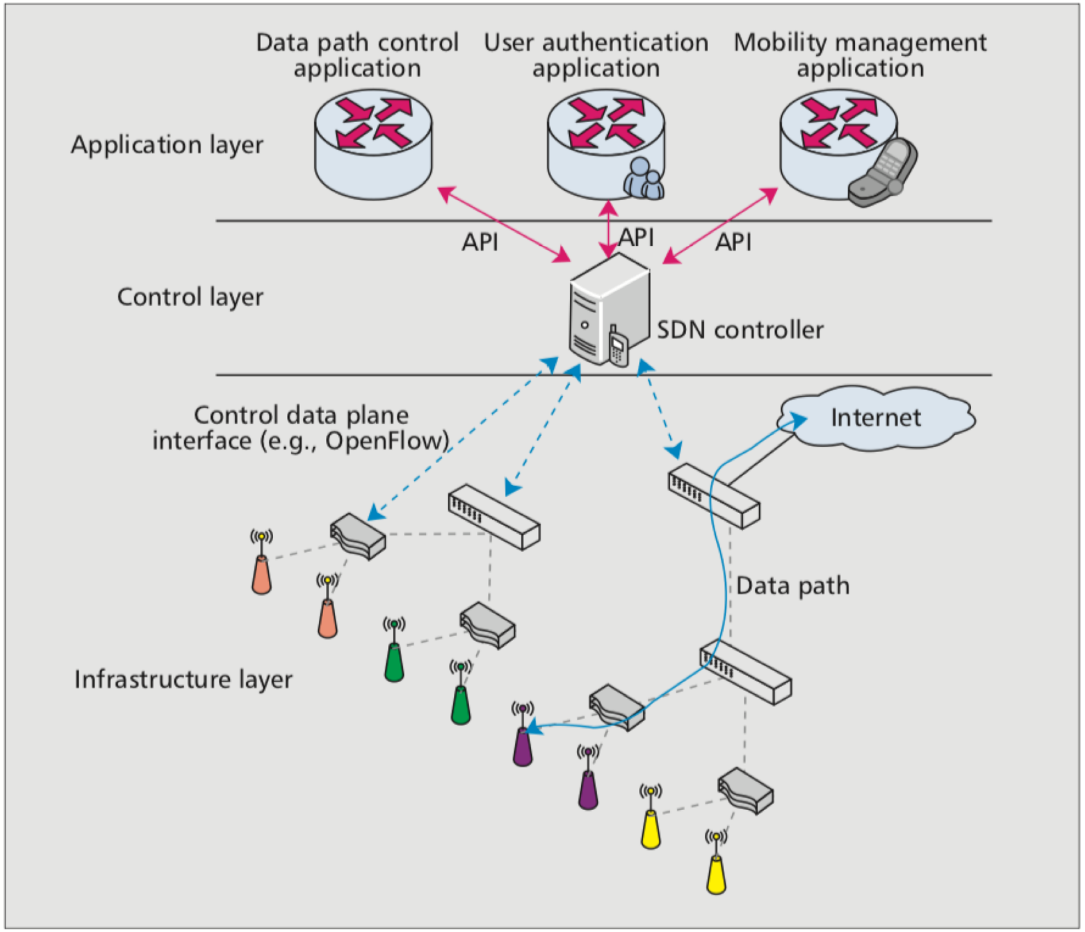




From: McKeown et. al "OpenFlow: Enabling Innovation in Campus Networks," ACM CCR 2008

SDN

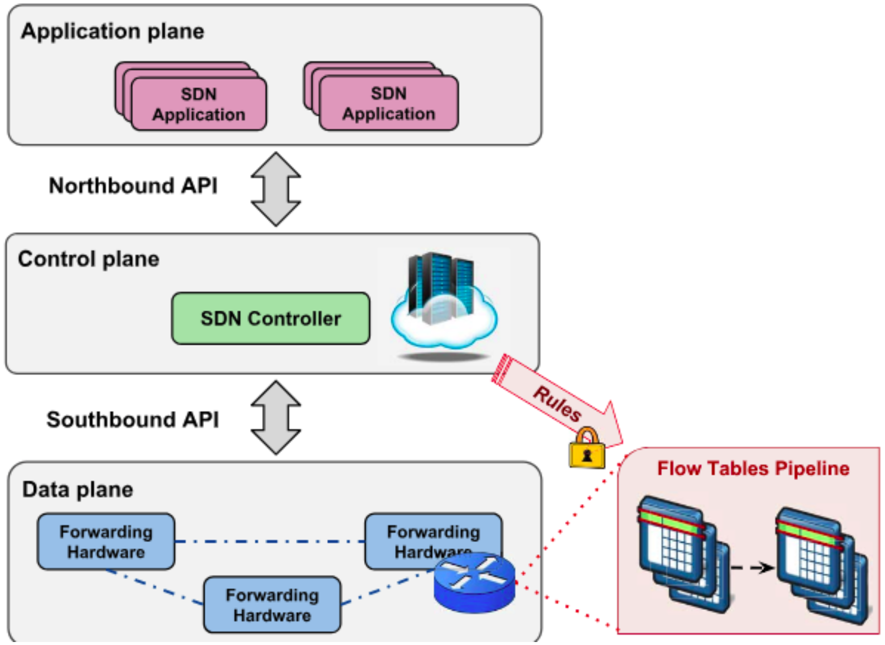
- Alternative to
  - Distributed approaches
  - Specialized hardware
  - Standardized protocols
- Interfaces
  - Southbound (OF)
  - Northbound

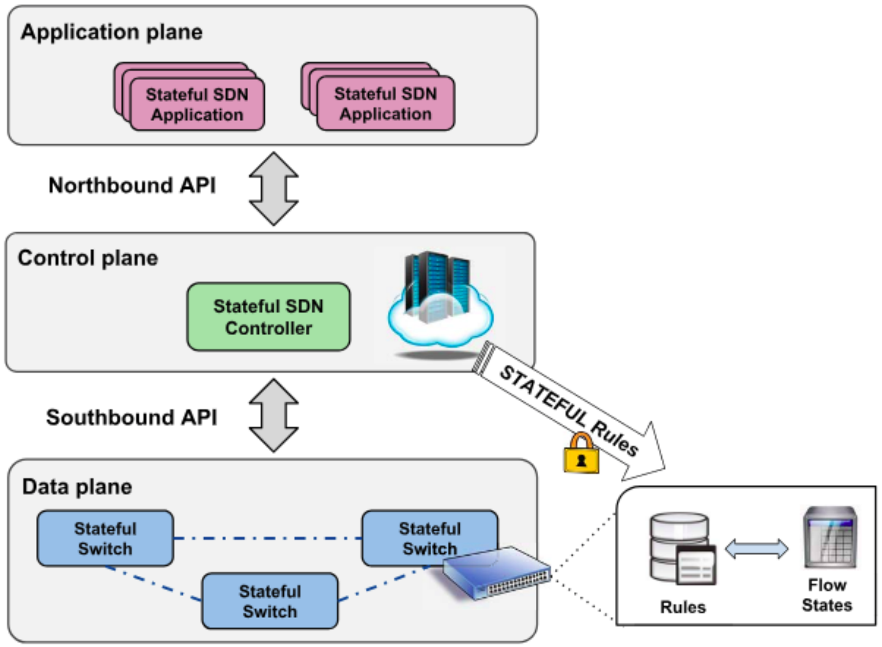


From: Bernardos et. al "An Architecture for Software Defined Wireless Networking," IEEE Wir. Comm. Mag. 2014

SDN: Control/Data split

- SDN: Breaks the vertical integration
  - Forwarding (switching/routing), Intrusion Detection Systems (IDS)
  - Upcoming stateful SDN switches [1], i.e., more applications



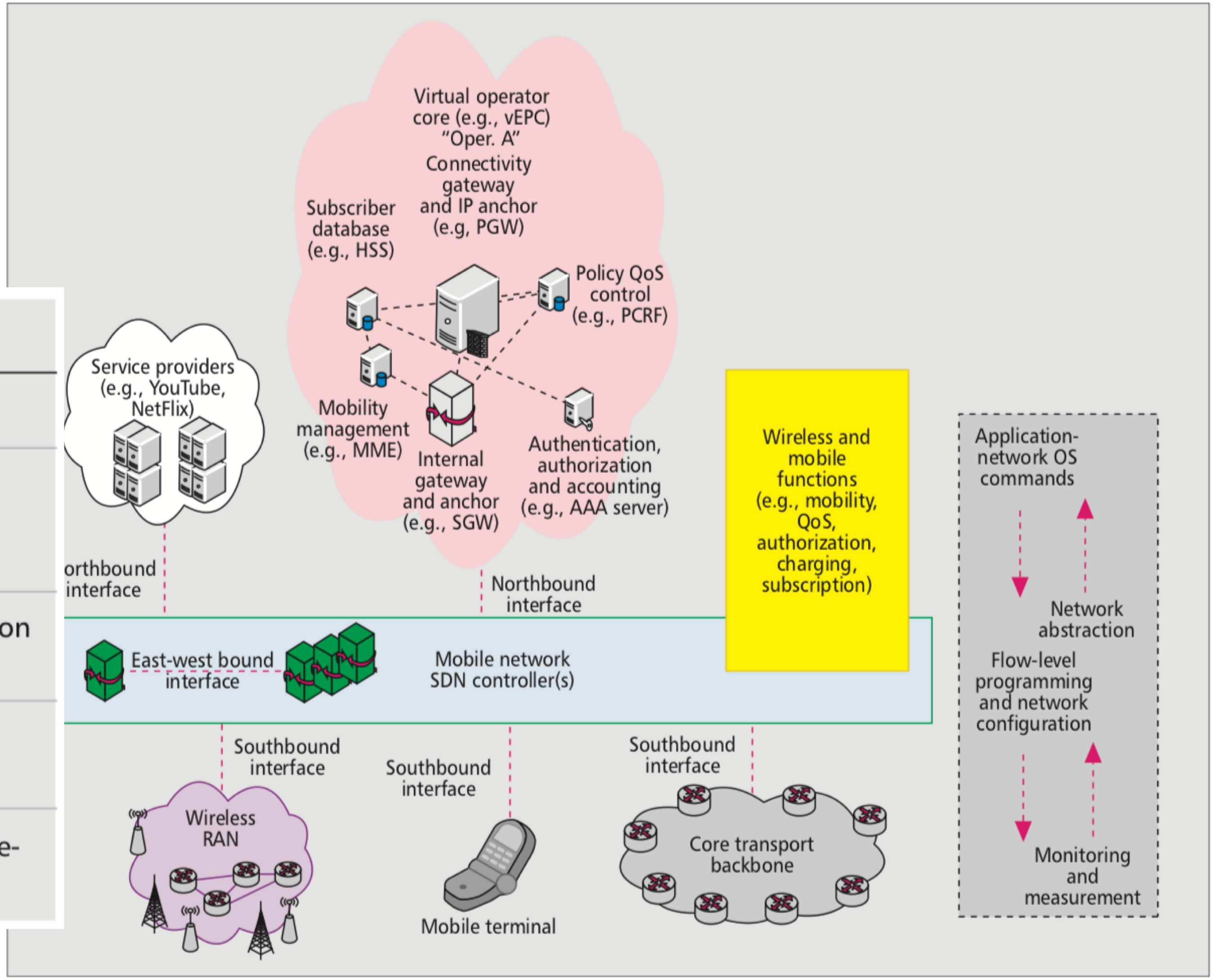


[1] T. Dargahi, A. Caponi, M. Ambrosin, G. Bianchi and M. Conti, "A Survey on the Security of Stateful SDN Data Planes," in IEEE CS&T, vol. 19, no. 3, pp. 1701-1725, thirdquarter 2017.

SDN for Wireless

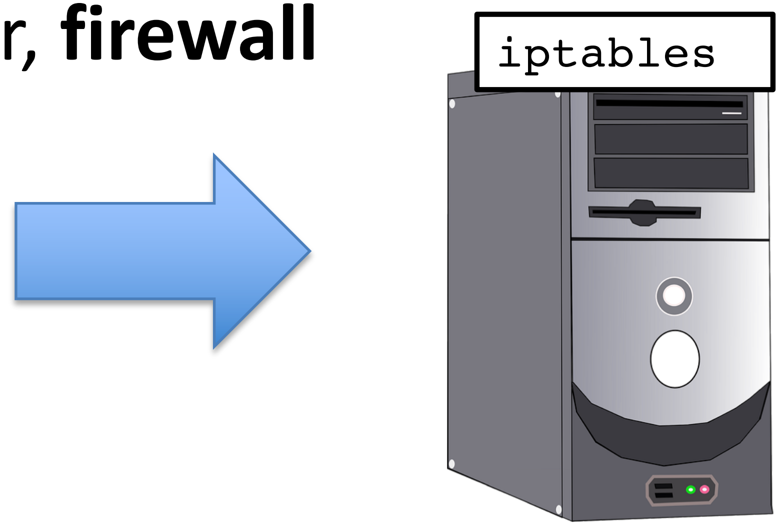
Key benefits
Easier deployment of new services
Reduced management and operational costs of heterogeneous technologies
Efficient oper infrastructure
Increased acc differentiatio
Continuous a enhancement

Key challenges
Specification of the interfaces
Need to integrate scheduled-based and contention-based systems
Harmonization of the standardization efforts
Verifiable security and privacy architecture
Operation and management of wireless networks is more complex



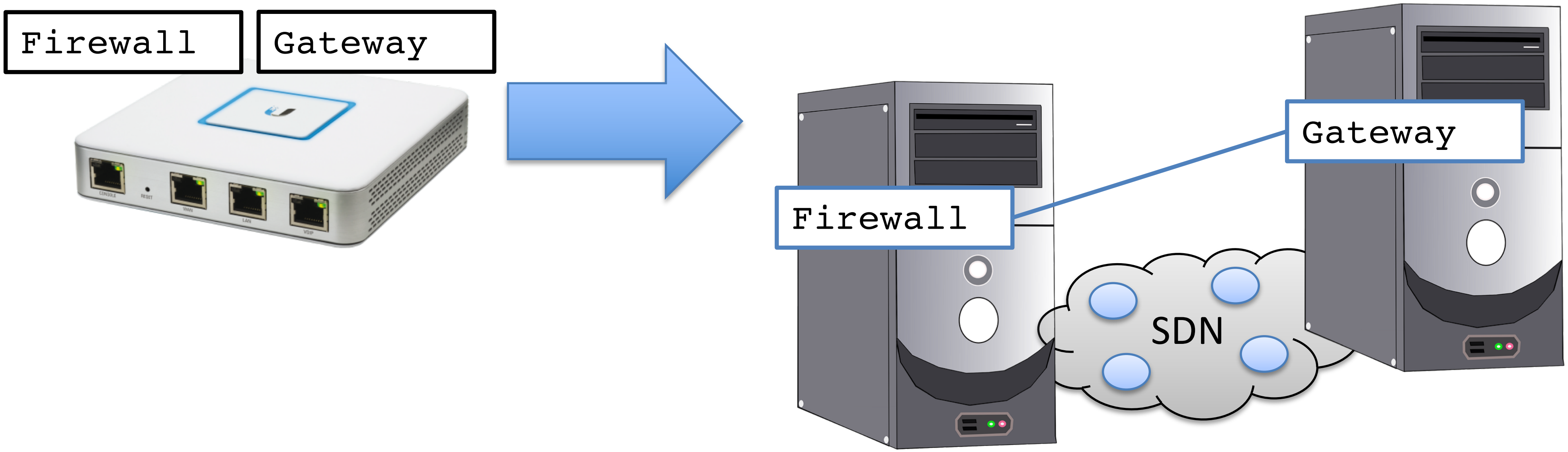
Network Function Virtualization

- Often considered part of SDN
- **But:** Complementary technology which depends on SDN to deliver its benefits
- Network Function: Building block of a communication service
  - E.g., gateway, load balancer, **firewall**

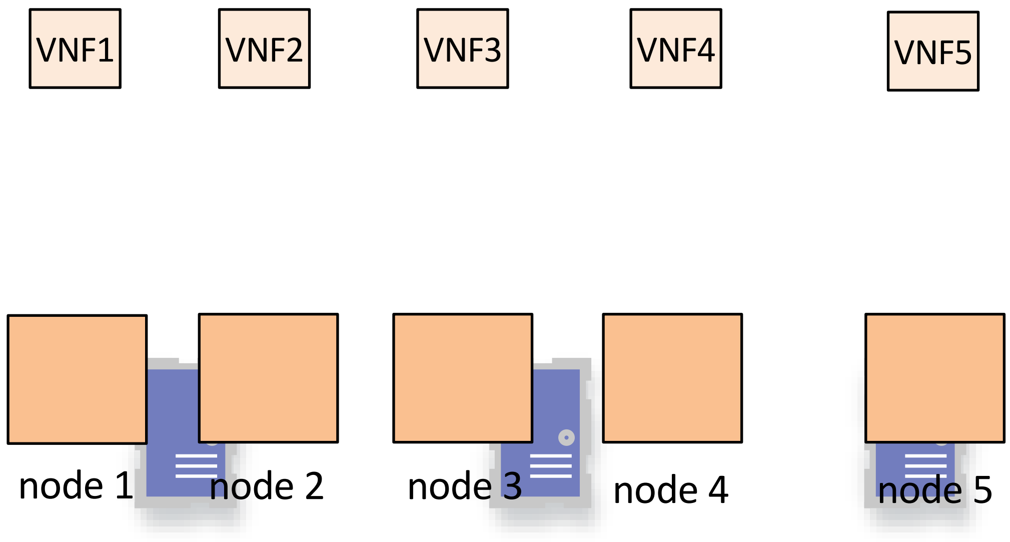


NFV: definition

- To use technologies of IT virtualization to virtualize and connect network node functions



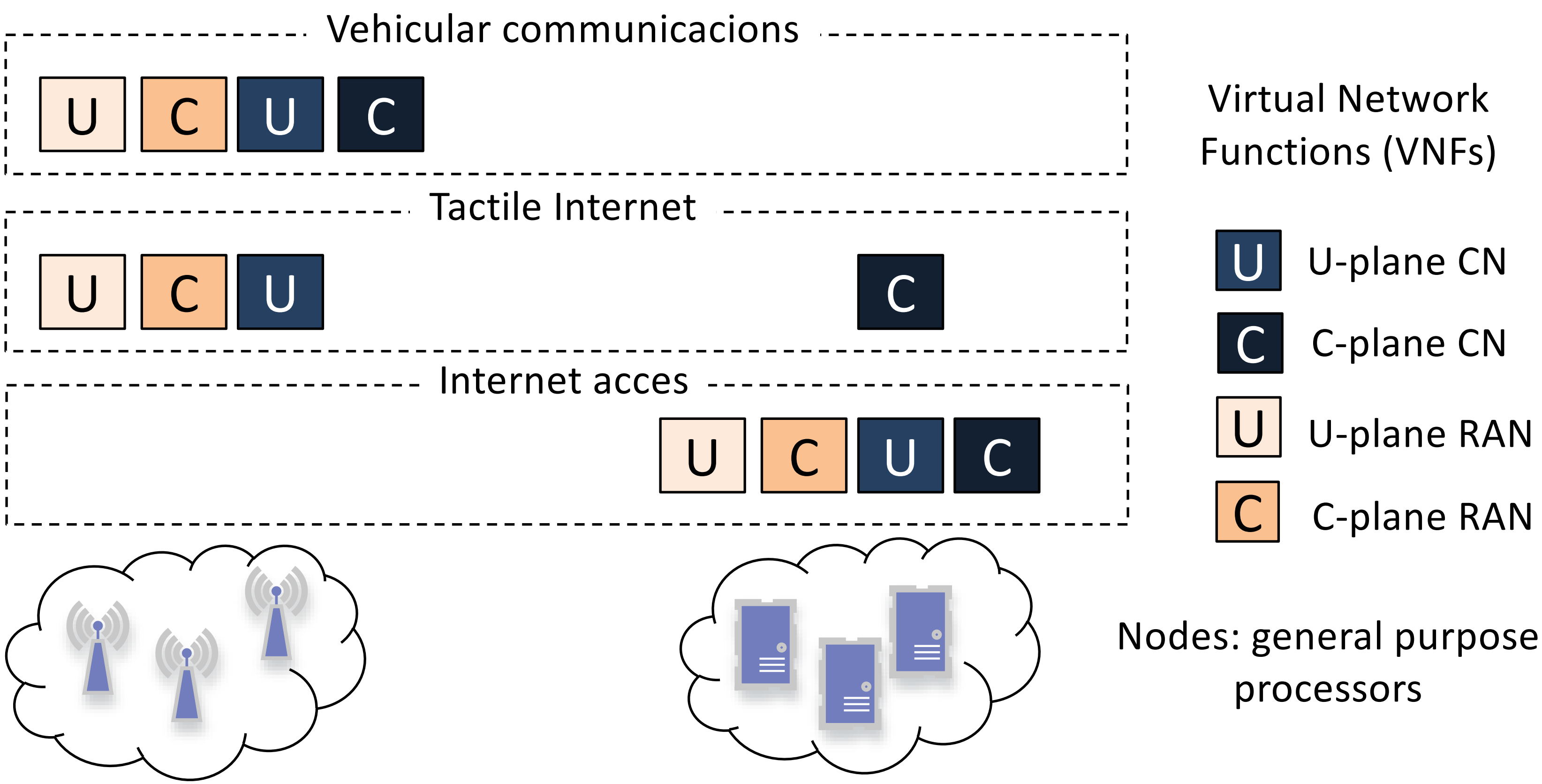
Virtualizing Network Functions



Virtualized network functions (VNFs)

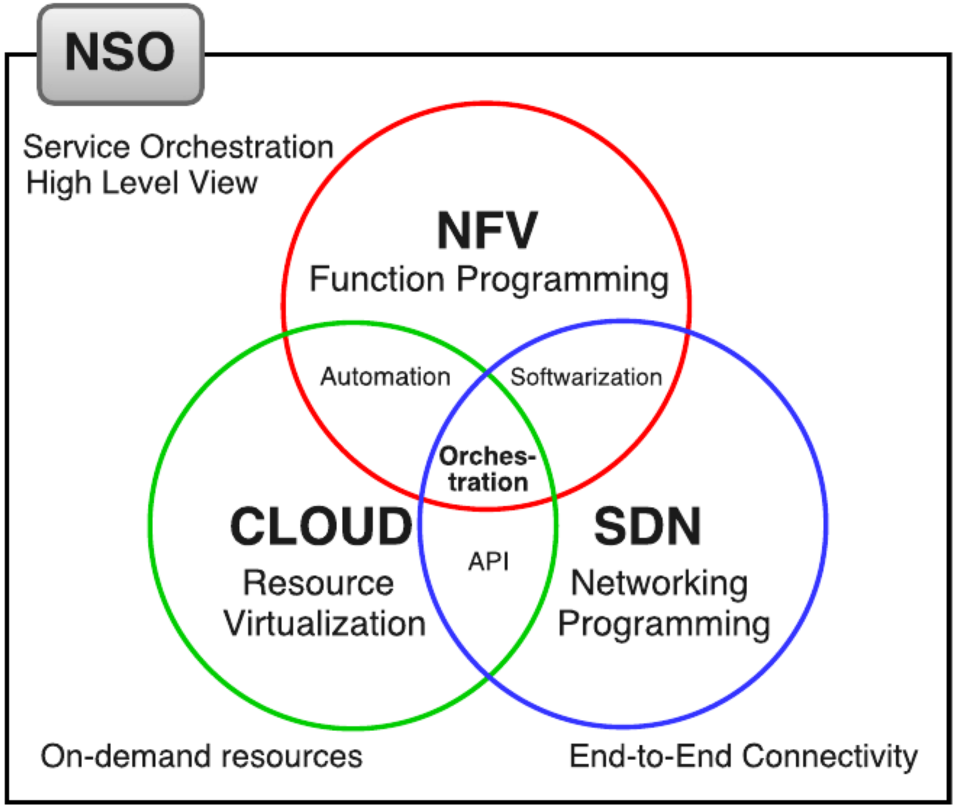
Nodes: general purpose processors

Orchestration of VNFs



Orchestration

- Coordinated set of activities to automate the control of physical and virtual resources



From: Source: de Sousa, N. F. S., Perez, D. A. L., Rosa, R. V., Santos, M. A., & Rothenberg, C. E. (2018). Network Service Orchestration: A Survey. arXiv preprint arXiv:1803.06596. 2018

Orchestration

- Automatic selection and control multiple resources, services, and systems to meet certain objectives
- Approx. plan
  - 2017: based on analytics
  - 2018: analytics + prescription
  - 2019: semi-automated operation (some supervision)
  - 2020: fully automated

Softwarization: Wrap Up

- Centralization of the intelligence
  - Scalability and reliability concerns
    - Time and CPU complexity
- Hardware agnosticism
  - Interfaces to control network elements
  - Running specialized functions
- Target: fully-automated operation
  - ETSI: Industry Specification Group 'Experiential Network Intelligence' (ENI)
  - ETSI: Zero touch network & Service Management (ZSM)

NETWORK SLICING: WHAT’S NEW

NS: Some Early Attempts

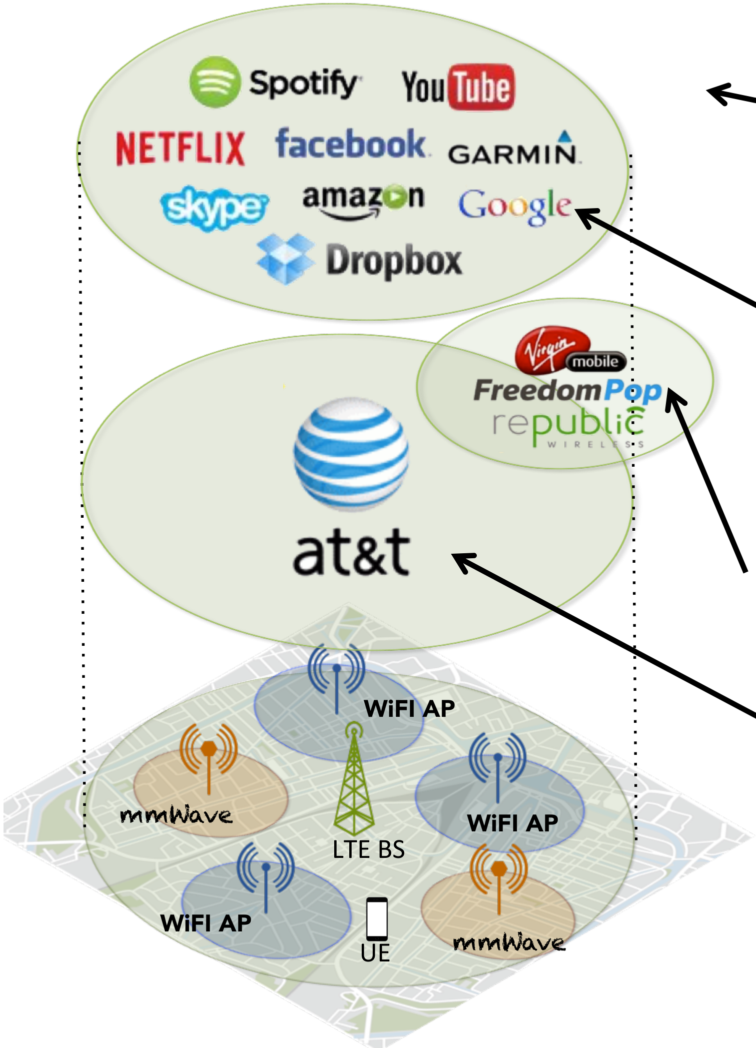
- Active / Programmable Networks research: node operating systems & resource control frameworks (1995 -2005)
- Federated Testbed research : Planet Lab USA (2002), PlanetLab EU (2005), OneLab EU (2007), PlanetLab Japan (2005), OpenLab EU (2012)
- GENI (2008): Shared network testbed i.e. multiple experimenters may be running multiple experiments at the same time.

From: A. Galis, K. Makhijani, “Network Slicing Tutorial,” IEEE NetSoft 2018, June 2018

Main novelties

- **Ecosystem**: Different players, domains, modules and interfaces
- **Life cycle**: [Deploy new services] ”From 90 hours down to 90 minutes”
- **Customization**: “Underlays / overlays supporting all services equally (‘best effort’ support) are not fully representing a Network Slice”

Current Mobile Netorks



OTTs with new services

i. Internet of Things
ii. Vehicular communications

Over-the-Top (OTT)

i. Getting lion's share of revenues, but
ii. Can't customize due to Net Neutrality

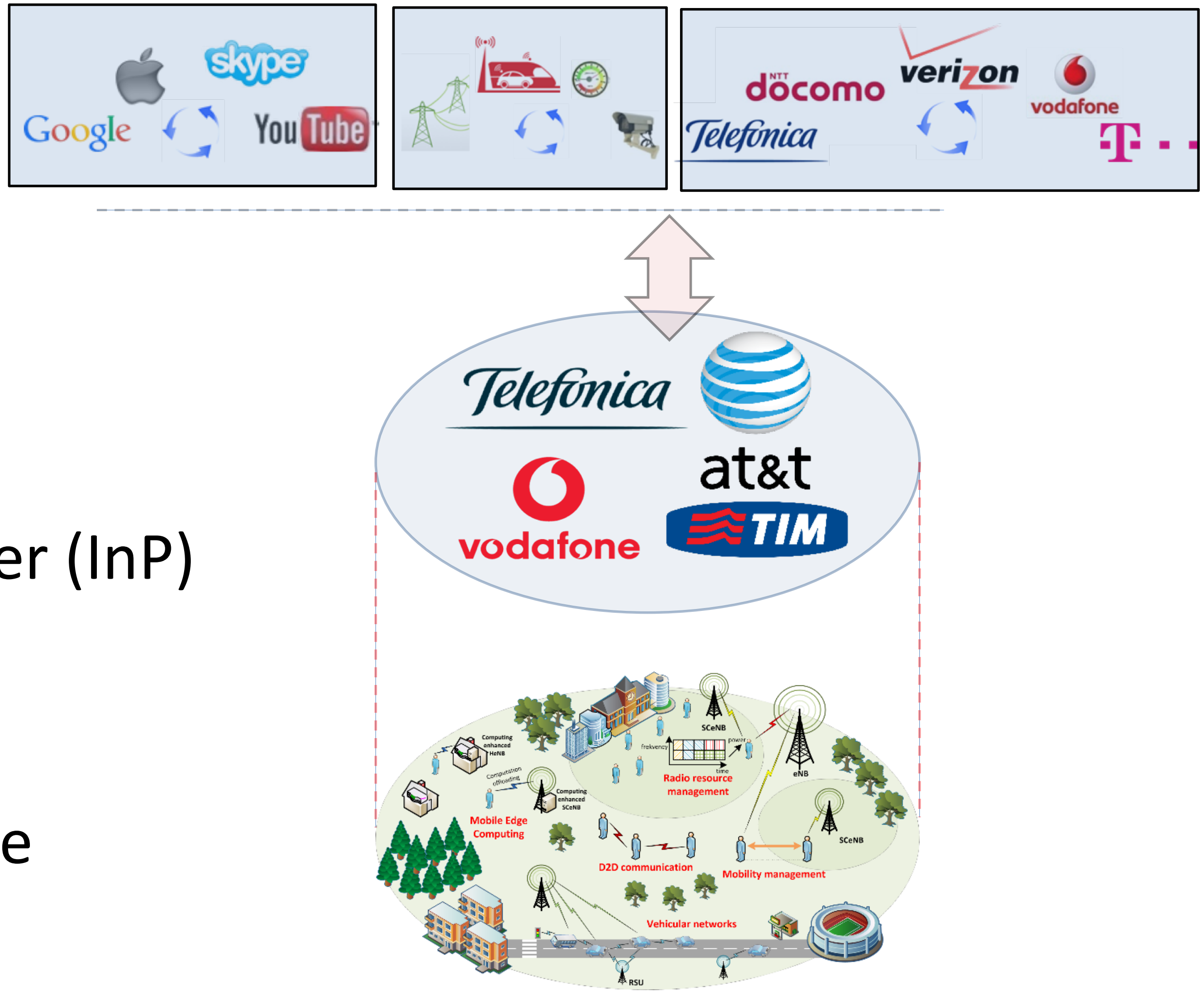
Mobile Virtual Network Operators (MVNO)

Mobile Network Operators (MNO)

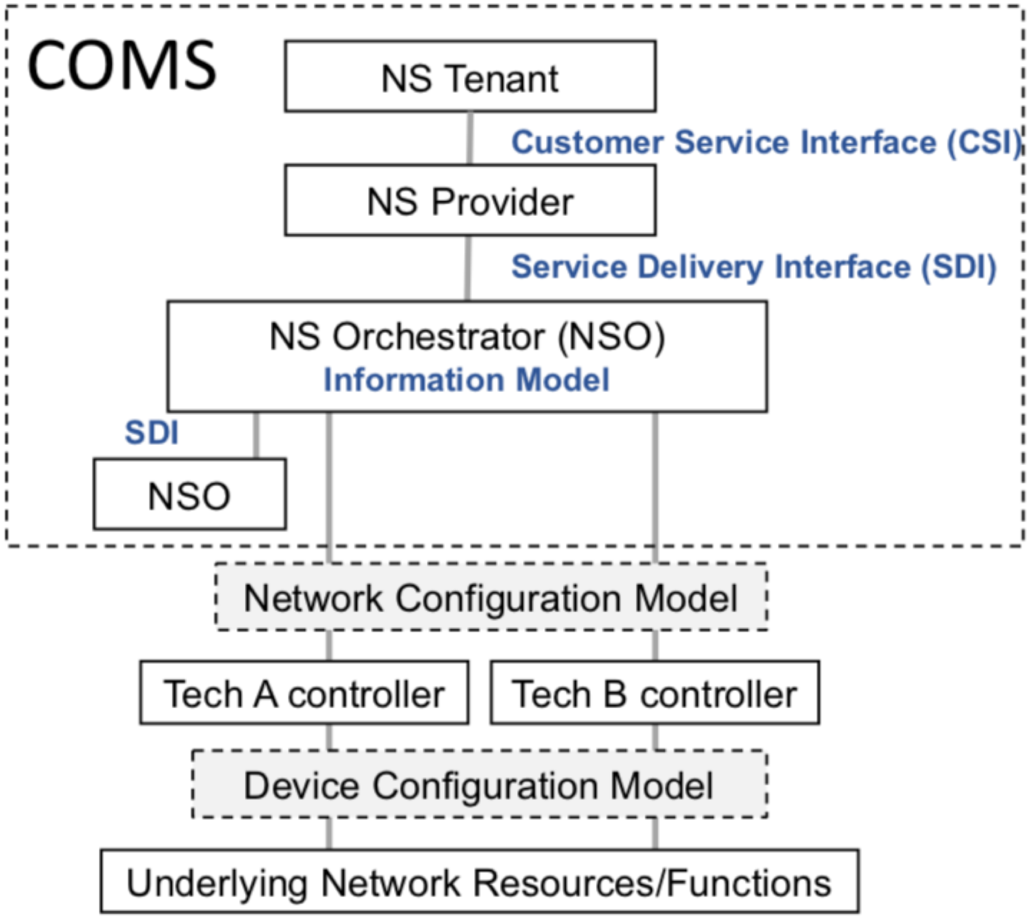
i. Deal with infrastructure costs
ii. Can't generate new revenue streams

Novel Players

- Tenants
  - Over-the-Top (OTT)
  - Vertical industries
  - Network Operators
- Infrastructure provider (InP)
- Physical infrastructure



Novel Language (IETF)



- CSI: management tasks within their slice instance under certain policies
- SDI: describe a NSaaS in network language. Also used between orchestrators, enabling hierarchical management.
- COMS: Information model, network slice entities in terms of resource components and characteristic attributes (in progress)

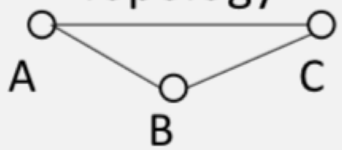
From: Geng et al. "Architecture of Network Slicing Management", IETF 102, July 2018

Top-Down Example

A NST is requesting a NSaaS

CSI)

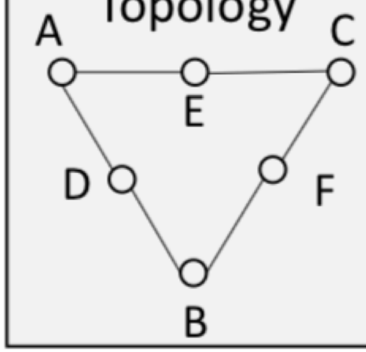
Topology



100 Mbit/s assured BW
20T CDN network
20 VMs at DC 1

DI)

Topology



A-E BW Policy
E-C BW Policy
....
CDN network topology
CDN resource attribute
VM configurations

CDN configuration Domain 1	A-D-B-F Domain 2	F-C-E-A Domain 3	MANO Domain 4

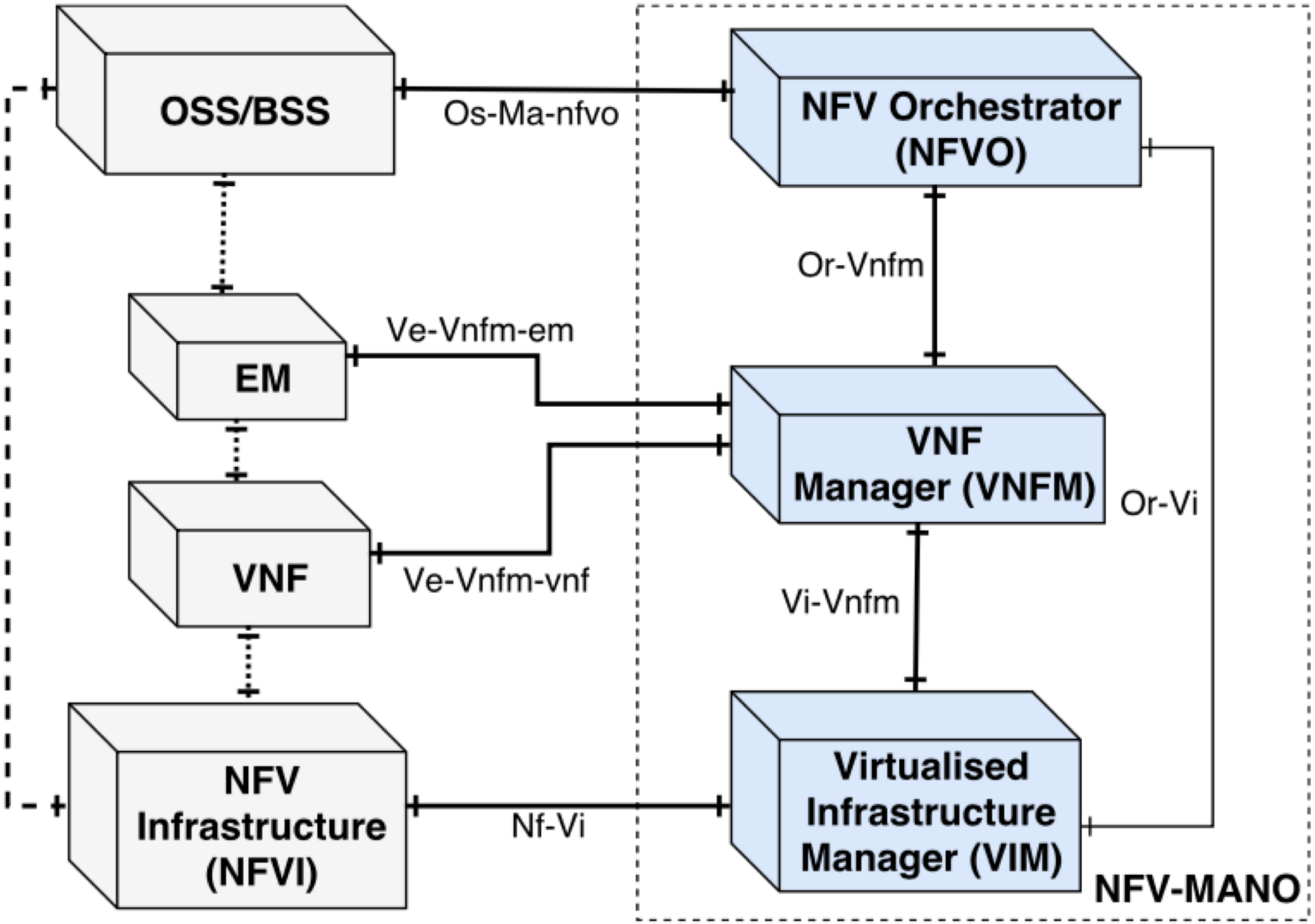
Someone is building a house

I want a 3-storey house with 5 en-suite bedrooms and a living room. With a size of 400 m^2

Exact floor plan
Exterior and interior design
Building material lists
....

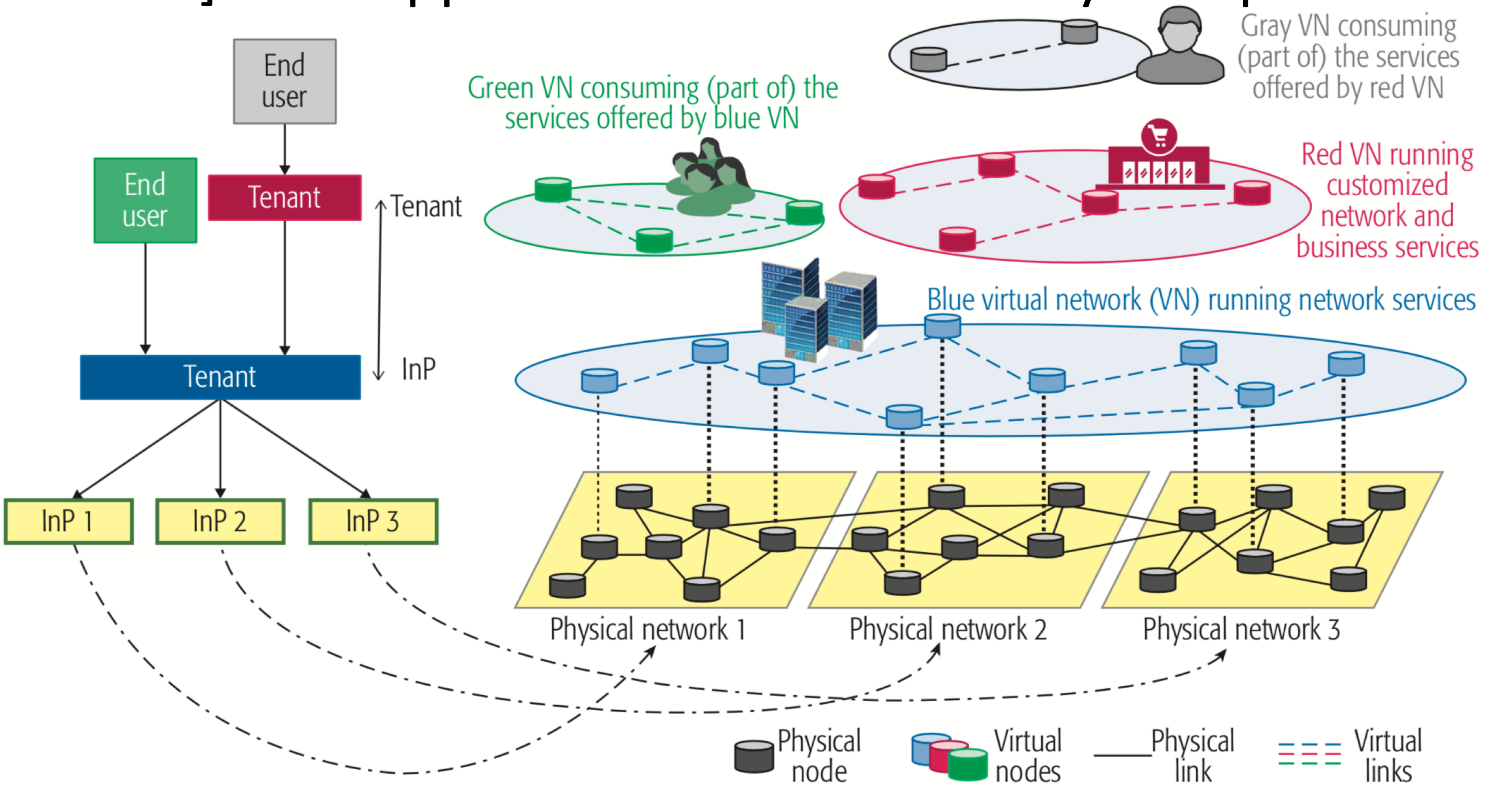
Building Design Studio	Construction Contractor	B&Q

ETSI NFV-MANO



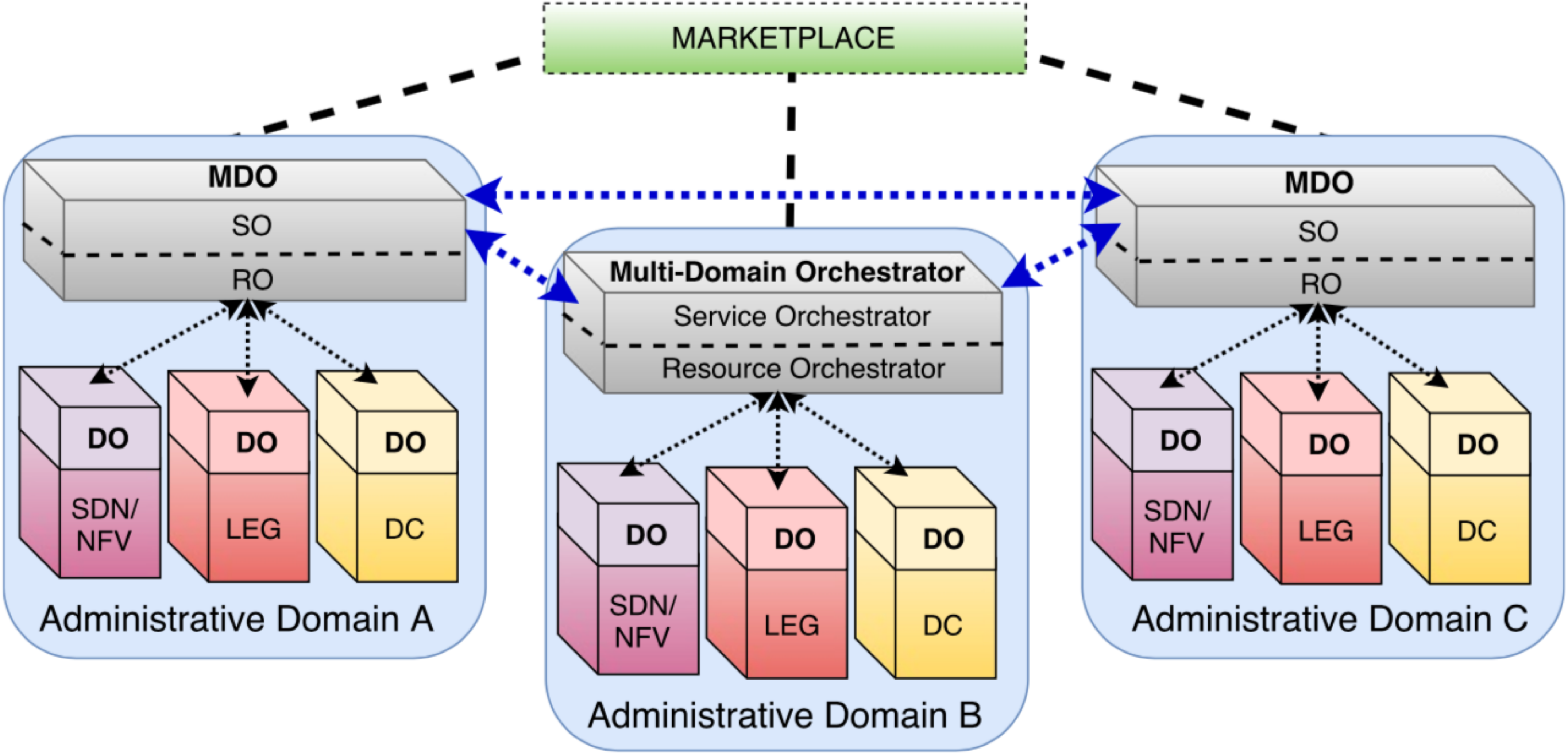
Ecosystem: Recursive

- “virtualization is naturally recursive, and [provider-tenant] can happen in a vertical multi-layered pattern”

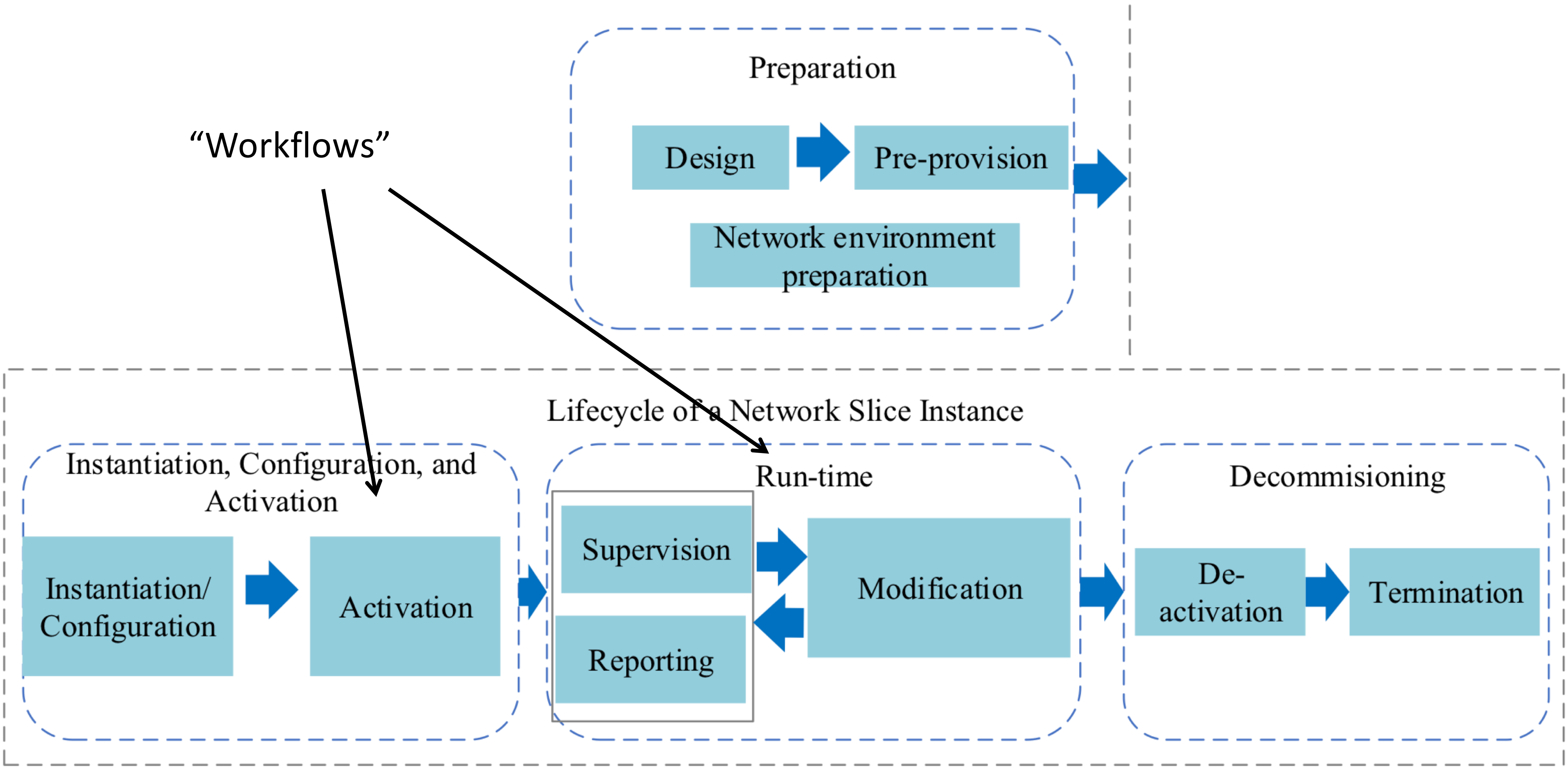


From: J. Ordonez-Lucena et al., “Network Slicing for 5G with SDN/NFV: Concepts, Architectures, and Challenges,” IEEE Comm. Magazine, May 2017

Ecosystem: Multi-Domain



Lifecycle of a NSI

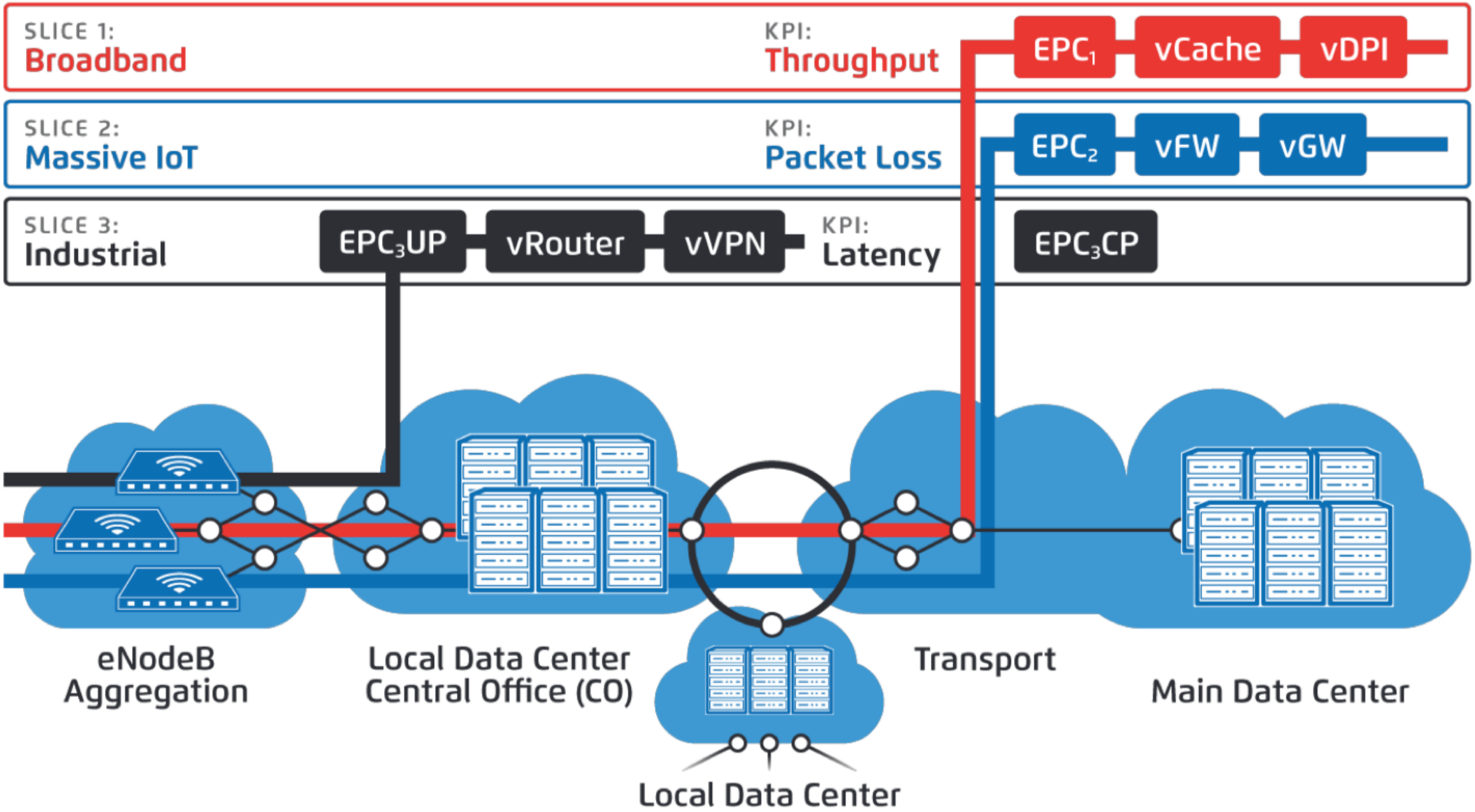


From: 3GPP TR28.801

Orchestration and Lifecycle

- Service Orchestration: service composition and decomposition.
  - Interaction with OSS/BSS
- Lifecycle Orchestration: management across service components
  - Service Level Agreement
- Resource Orchestration: mapping service requests to resources (virtual/physical)

Service Heterogeneity

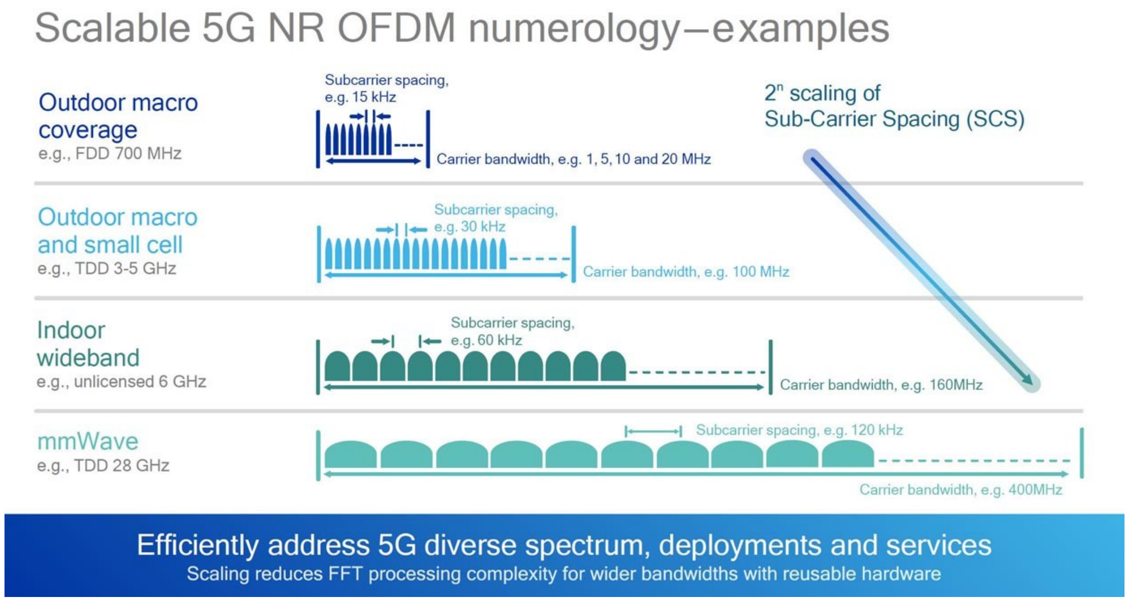


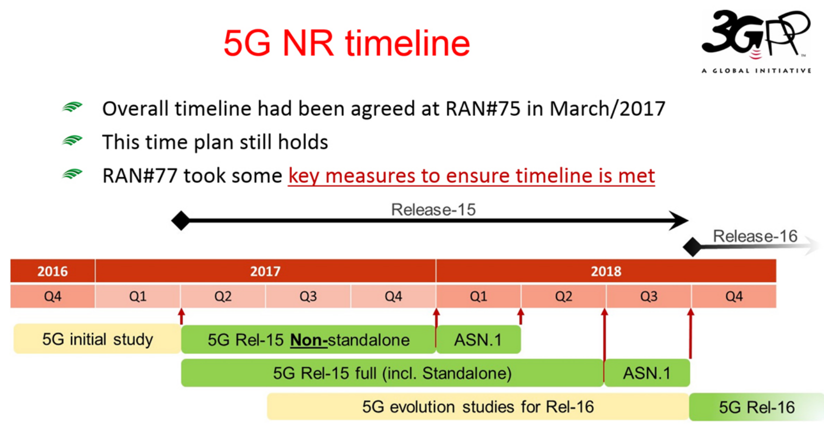
From: ETSI OSM PoC #1

WHAT IS 5G?

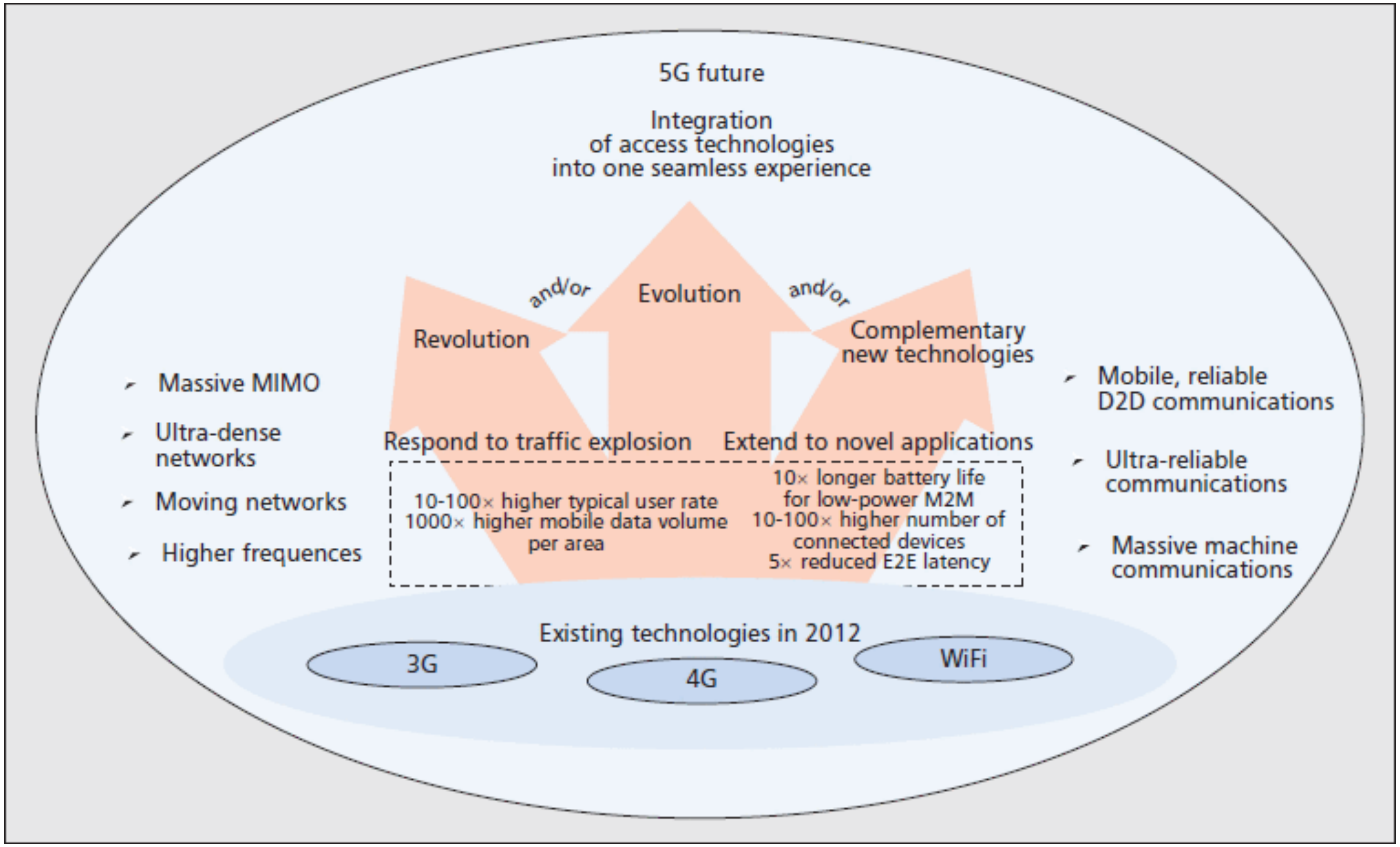
5G New Radio

- OFDM - Flexible Numerology
  - The subcarrier spacing is no longer fixed to 15 kHz
  - A mini-slot is shorter in duration than a standard slot and can start at any time





5G Access



From: A. Osseiran *et al.*, "Scenarios for 5G mobile and wireless communications: the vision of the METIS project," in *IEEE Communications Magazine*, vol. 52, no. 5, pp. 26-35, May 2014.

5G System

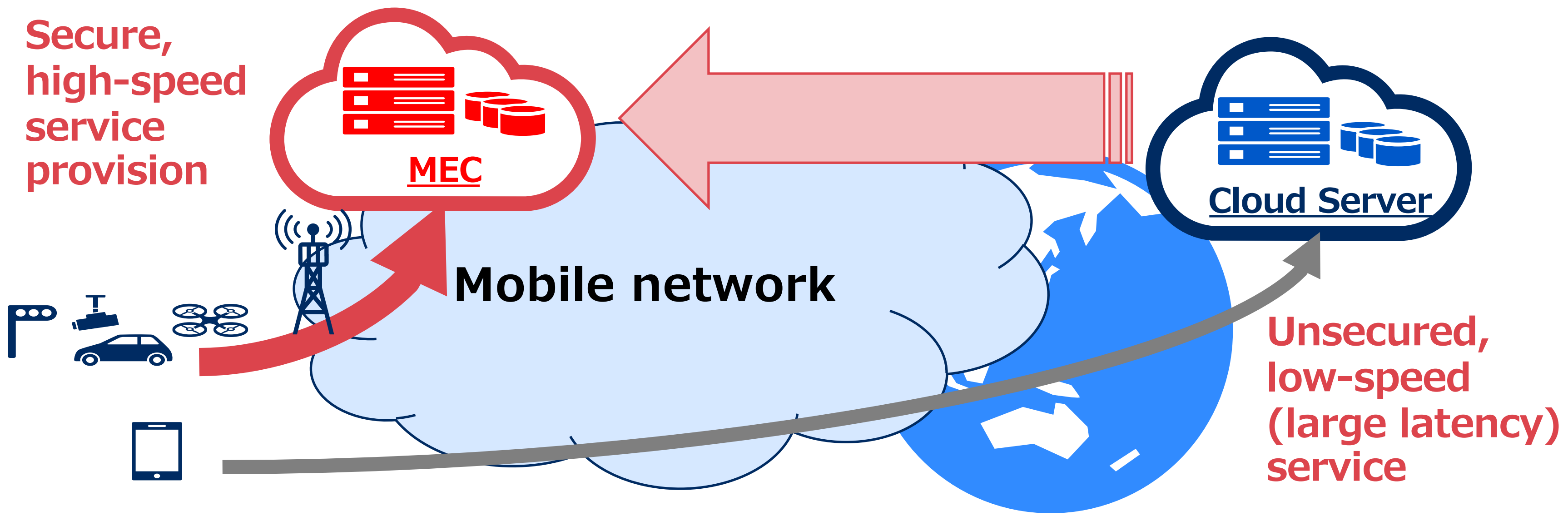
- Performance criteria for 5G systems set by ITU in their IMT-2020 Recommendation. ITU-R M.2083 describes three overall usage scenarios for 5G systems:
- **Enhanced Mobile Broadband** to deal with hugely increased data volumes, overall data capacity and user density
- **Massive Machine-type Communications** for the IoT, requiring low power consumption and low data rates for very large numbers of connected devices
- **Ultra-reliable and Low Latency Communications** to cater for safety-critical and mission critical applications

ETSI's 5G Building Blocks

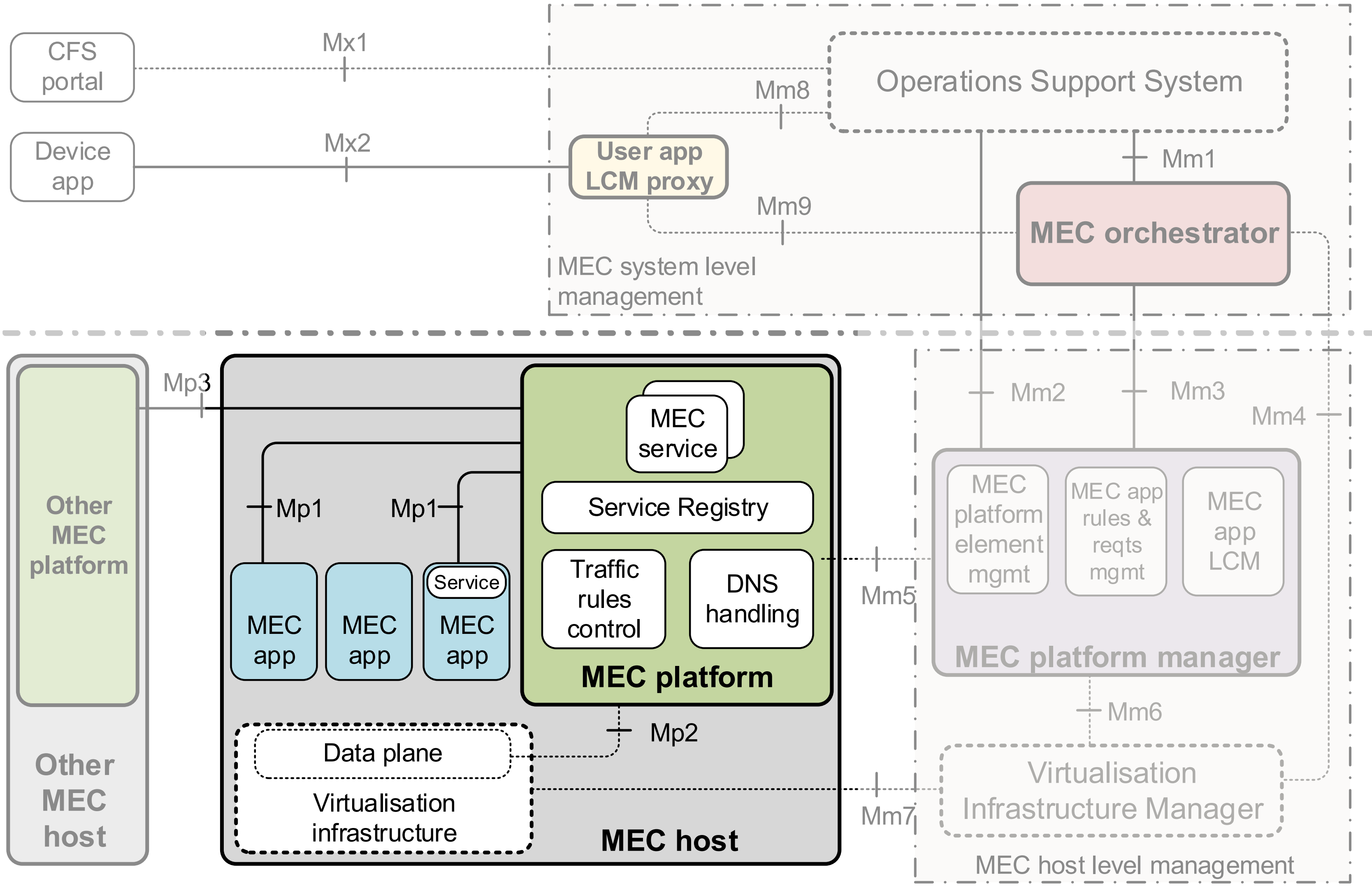
- ETSI has a number of component technologies which will be integrated into future 5G systems:
  - Network Functions Virtualization(NFV),
  - Multi-access Edge Computing (MEC),
  - Millimetre Wave Transmission (mWT)
  - Next Generation Protocols (NGP).

Mobile Edge Computing (MEC)

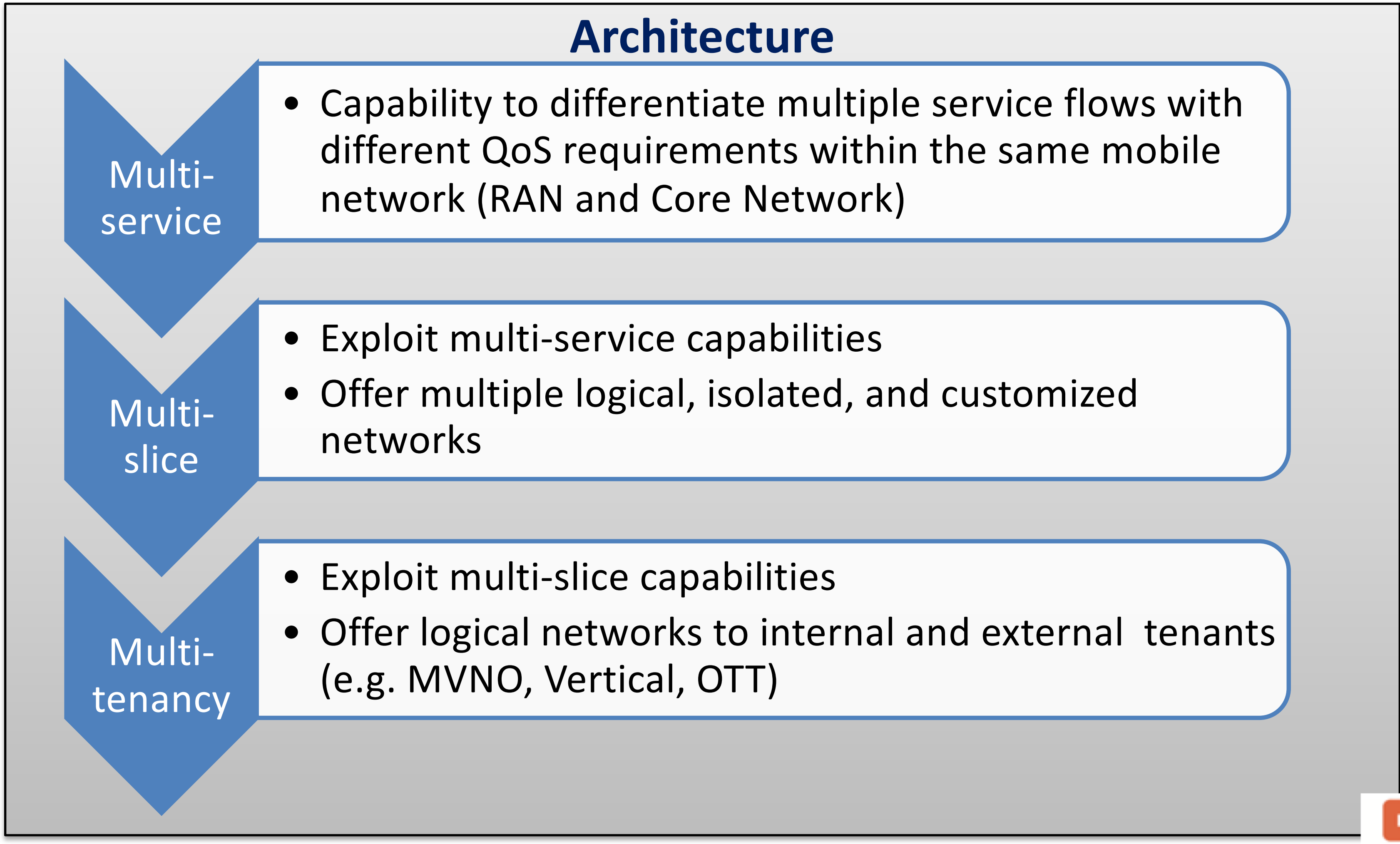
- Now: Multi-Access Edge Computing
  - Increases responsiveness from the edge
  - Context related services, more information about the consumer (e.g., radio)



MEC Architecture (GS MEC 003)



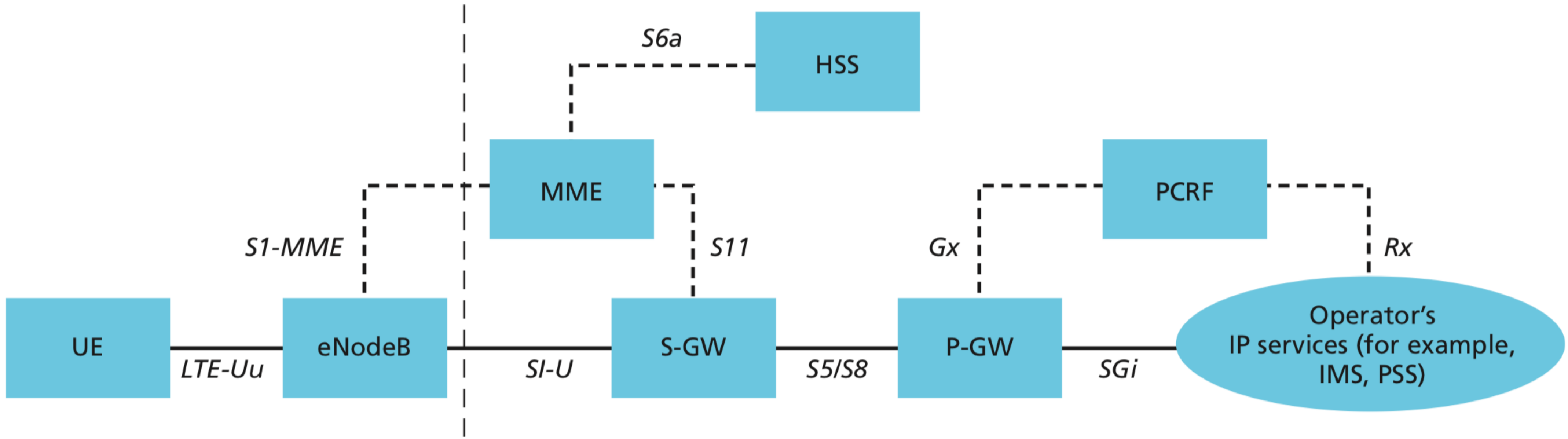
5G Architecture



5G ARCHITECTURE: BACKGROUND

LTE Network

- Access network (E-UTRAN)
  - Evolved Universal Terrestrial Radio Access Network
  - Essentially: the evolved NodeB (eNodeB)
- Core Network (Evolved Packet Network)
  - Many logical nodes



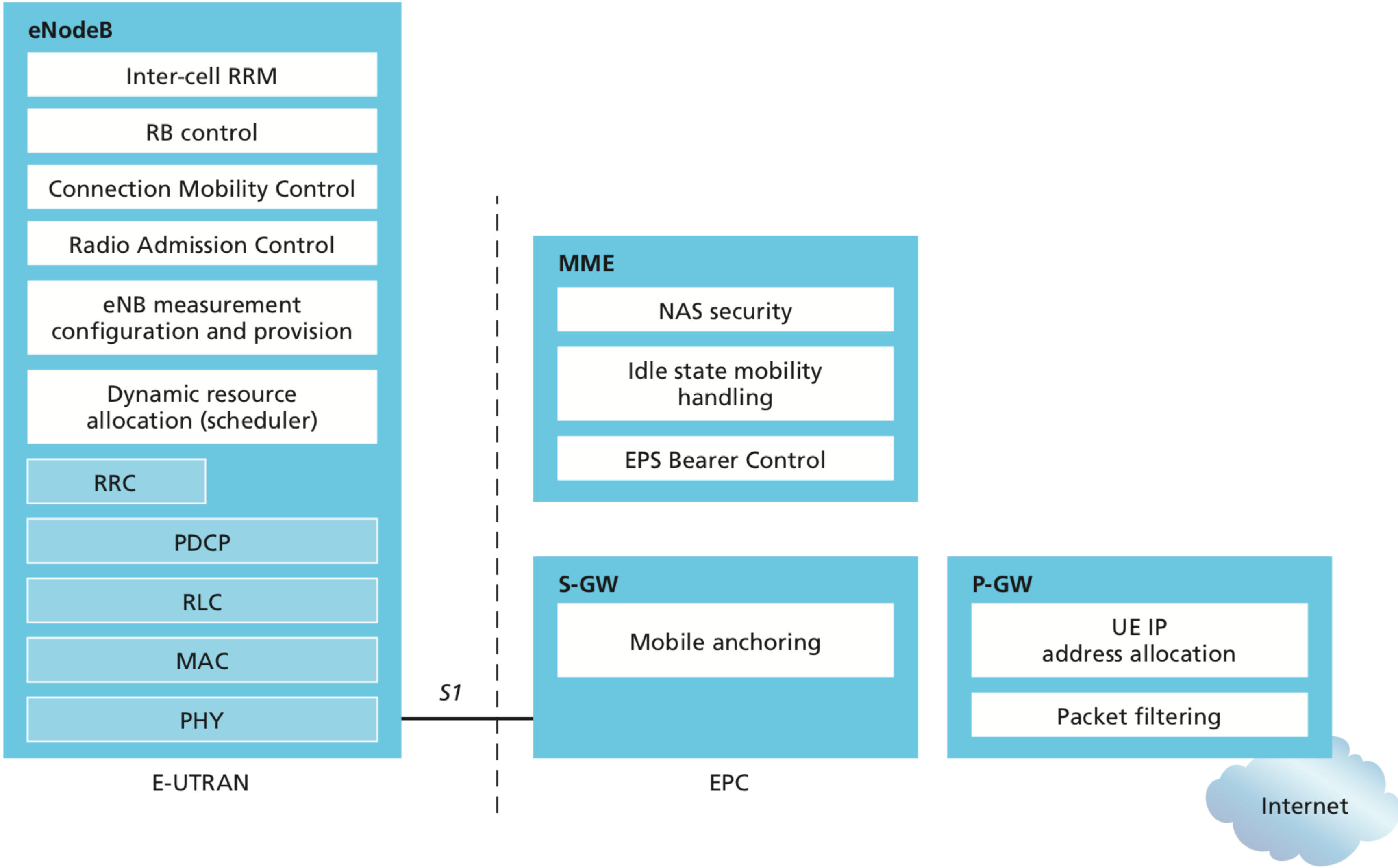
E-UTRAN

- E-UTRAN is responsible for all radio-related functions:
  - Radio resource management (RRM)
    - Radio bearer control, Radio admission control
    - Radio mobility control, Scheduling (both uplink and downlink.)
  - Header Compression: compressing the IP packet headers (especially useful for VoIP)
  - Security: All data sent over the radio interface is encrypted.
  - Connectivity to the EPC: signaling toward MME and the bearer path toward the S-GW.

EPC (Evolved Packet Core)

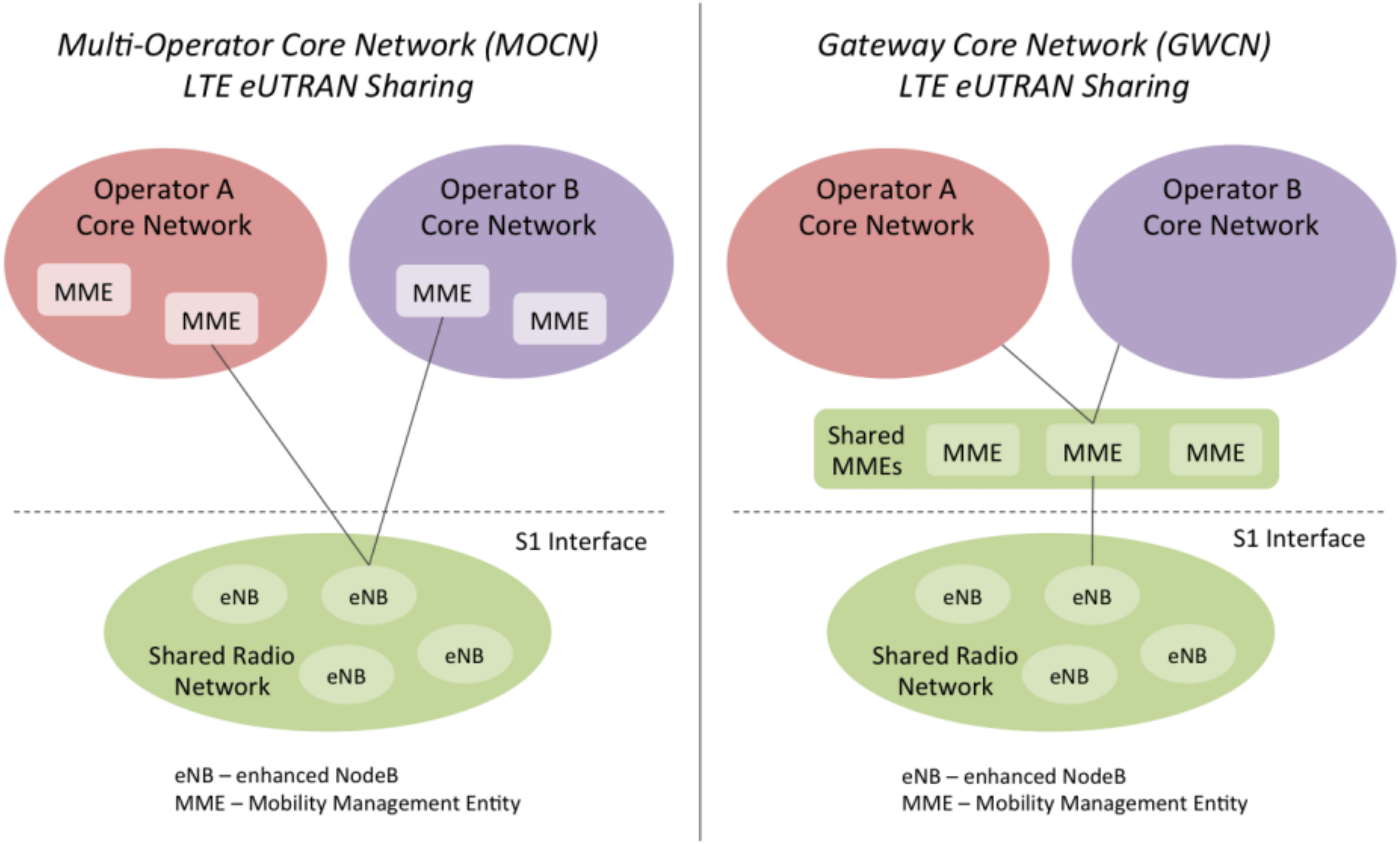
- PDN Gateway (P-GW)
- Serving Gateway (S-GW)
  - Mobility, roaming
- Mobility Management Entity (MME)
  - Processes the signaling between the UE and the CN (Non Access Stratum)
  - Bearer management and connection management
- Home Subscriber Server (HSS)
  - QoS profile, information about the PDNs to which the user can connect
- Policy Control and Charging Rules Function (PCRF)
  - Policy control decision-making
  - Controlling the flow-based charging functionalities in the Policy Control Enforcement Function (PCEF), which resides in the P-GW.

Functional Split



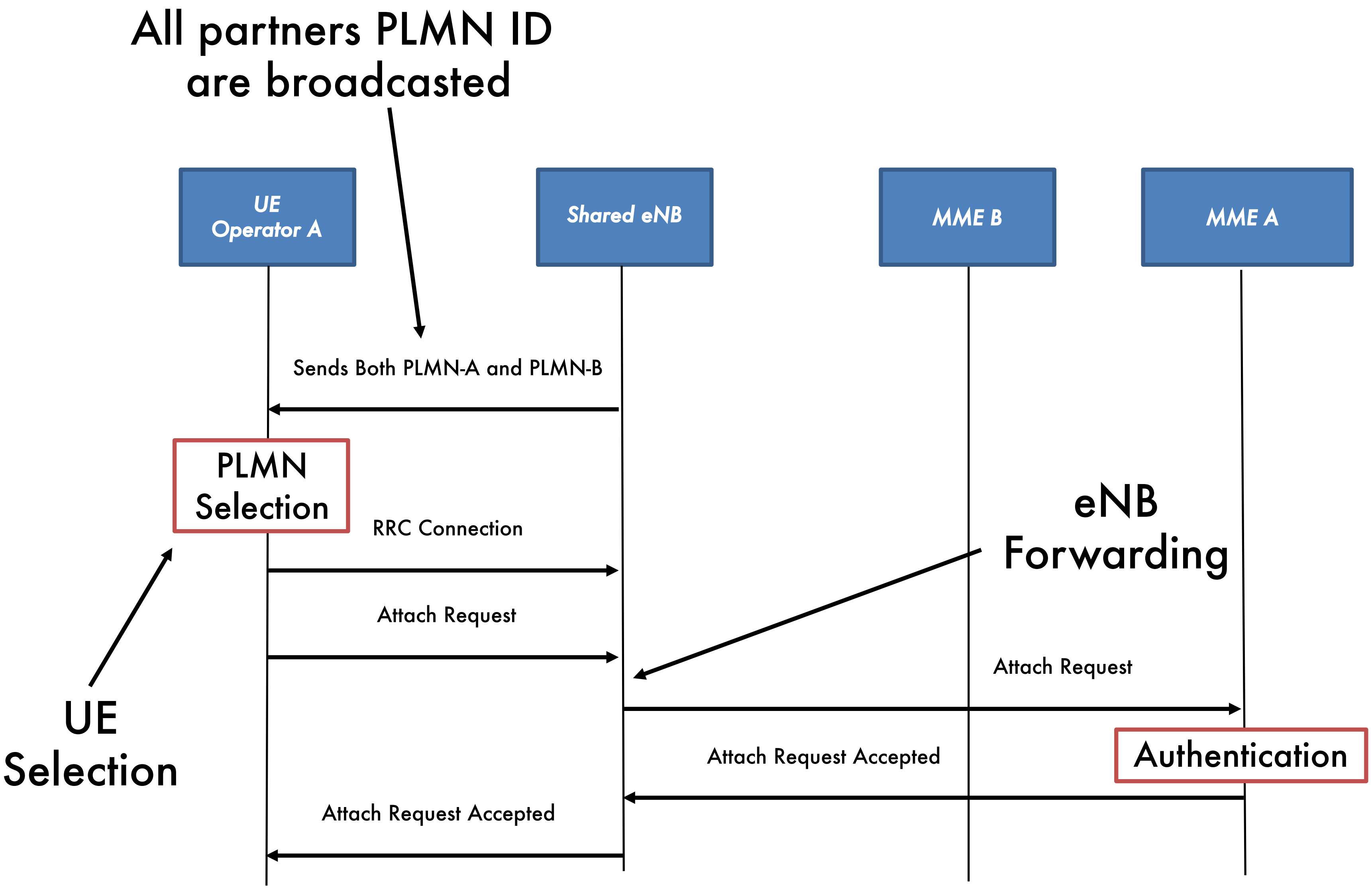
Multi-Tenancy

- Also known as RAN sharing



- Also: DECOR, eDECOR

E.g., MOCN

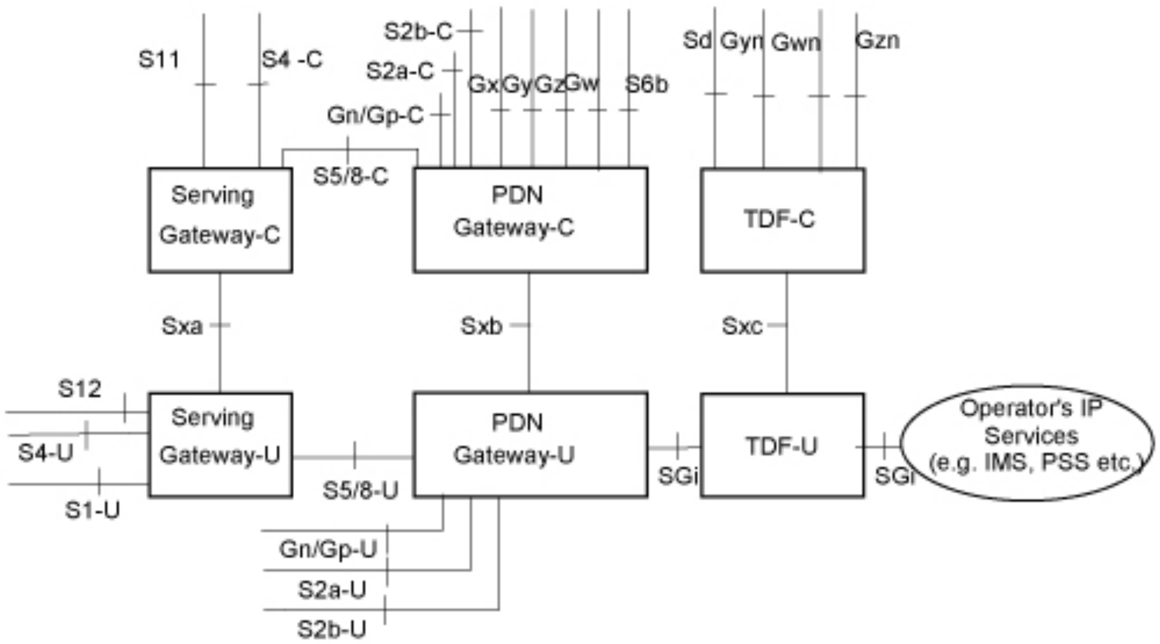


TS 23.251 TR 22.951

5G ARCHITECTURE: EVOLUTION

Softwarization: C/U Split

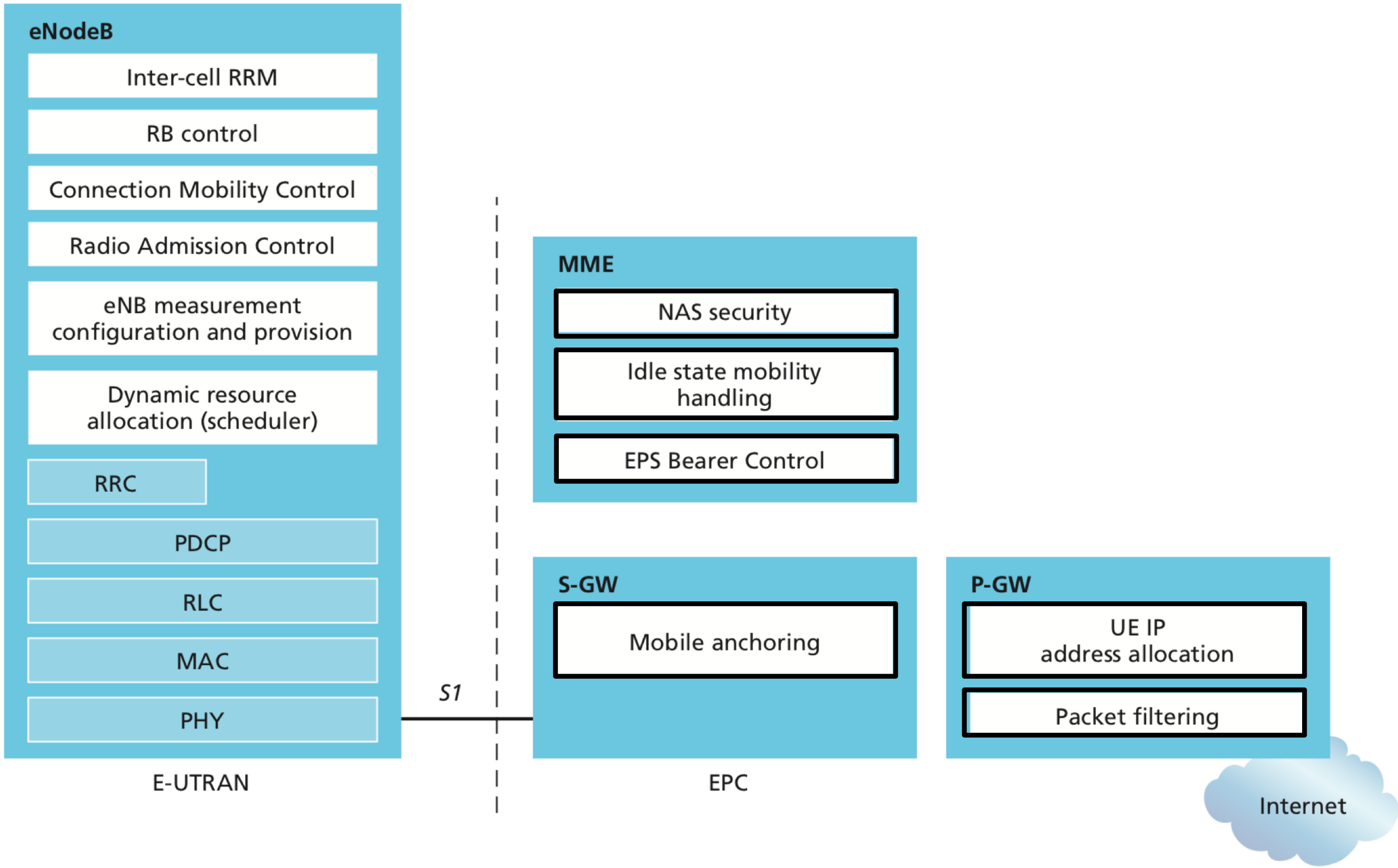
- R14: complete control user plane split (CUPS)
  - Interfaces, Sxa, Sxb and Sxc between the CP and UP functions of the SGW, PGW and TDF (Traffic Detection Fun), respectively.



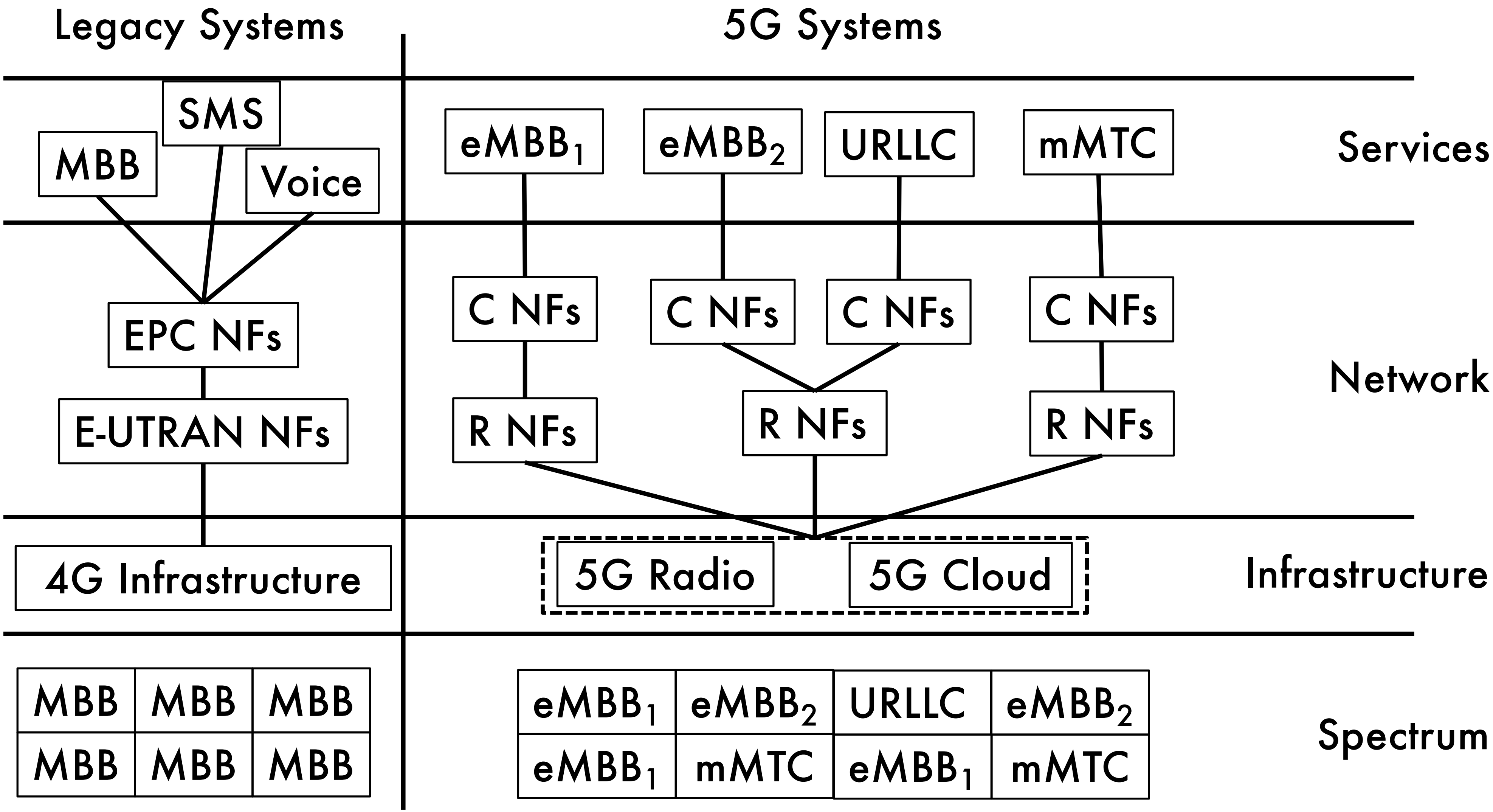
http://www.3gpp.org/cups

Modularization

- Consequence of softwarization

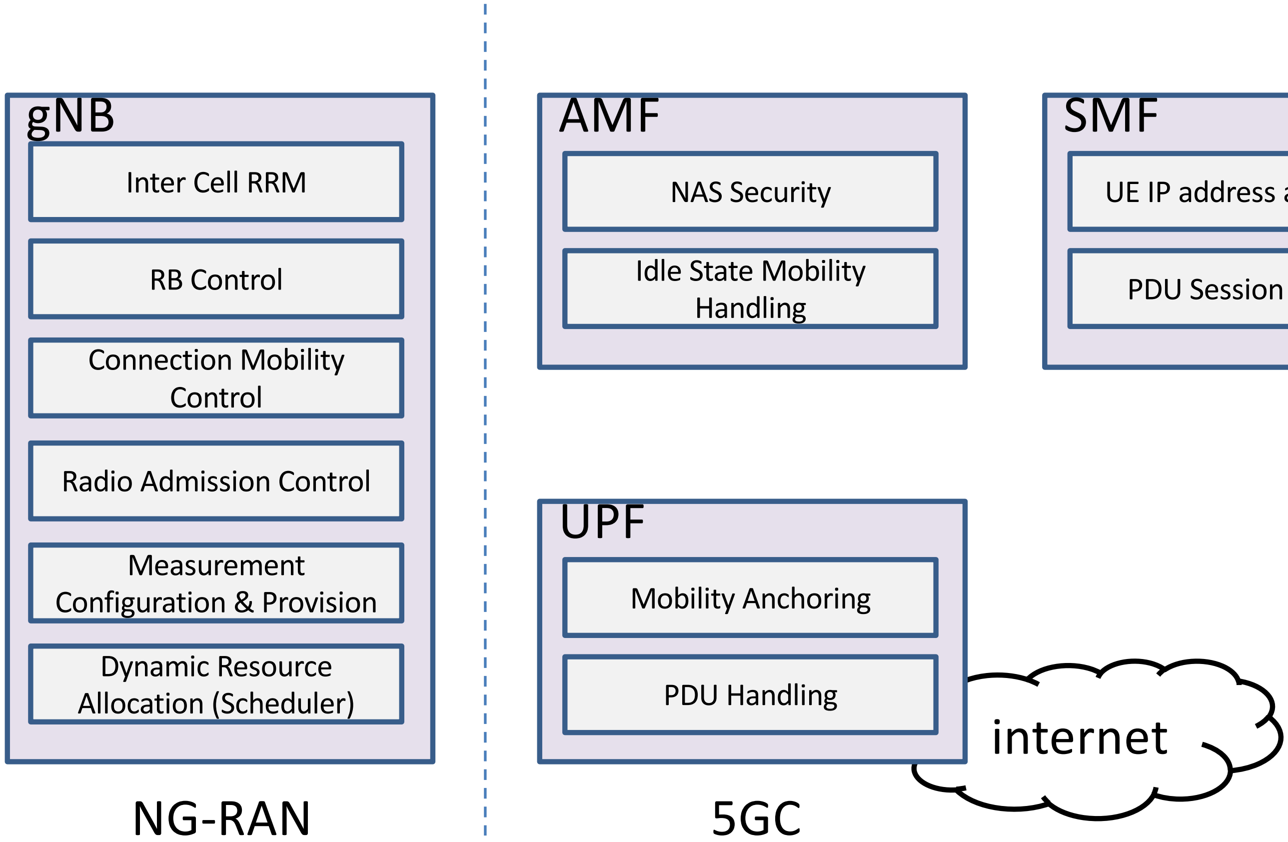


4G vs 5G

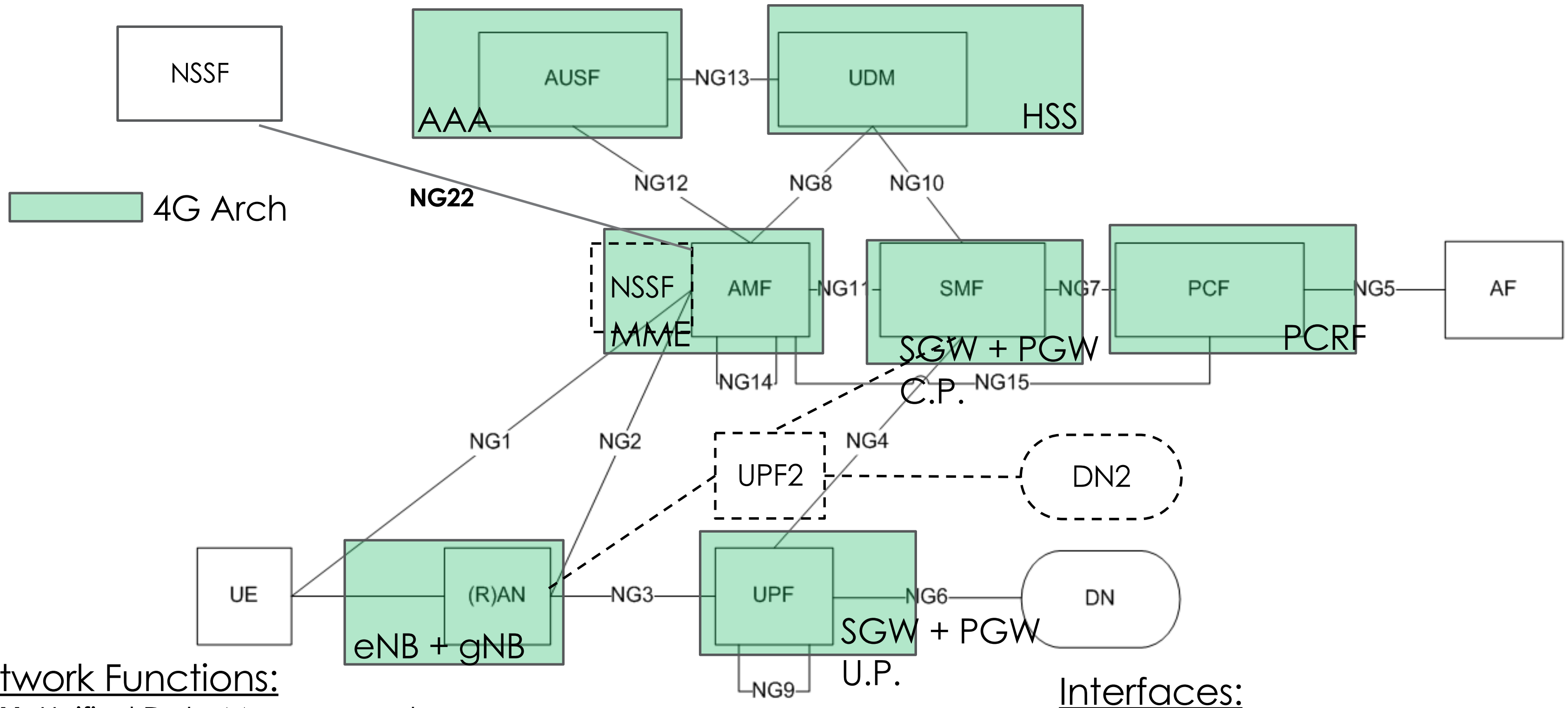


Adapted from "5G System design book"

Release 15


gNB
Inter Cell RRM
RB Control
Connection Mobility Control
Radio Admission Control
Measurement Configuration & Provision
Dynamic Resource Allocation (Scheduler)
AMF
NAS Security
Idle State Mobility Handling
SMF
UE IP address allocation
PDU Session Control
UPF
Mobility Anchoring
PDU Handling
internet
NG-RAN
5GC

3GPP SA2 Network Slicing Architecture



Network Functions:

- **UDM**: Unified Data Management
- **AUSF**: Authentication Server Function
- **PCF**: Policy Control Function
- **AMF**: Core Access and Mobility Management Function
- **SMF**: Session Management Function
- **UPF**: User plane Function
- **DN**: Data network, e.g., operator services, Internet access
- (*) **NSSF**: Network Slice Selection Function

Interfaces:

- **NG1**: NAS
- **NG2**: AN-CN C-plane
- **NG3/NG9**: per PDU Session tunnelling
- **NG7/8/10-15**: C-Plane **Service-based interface**

New functions

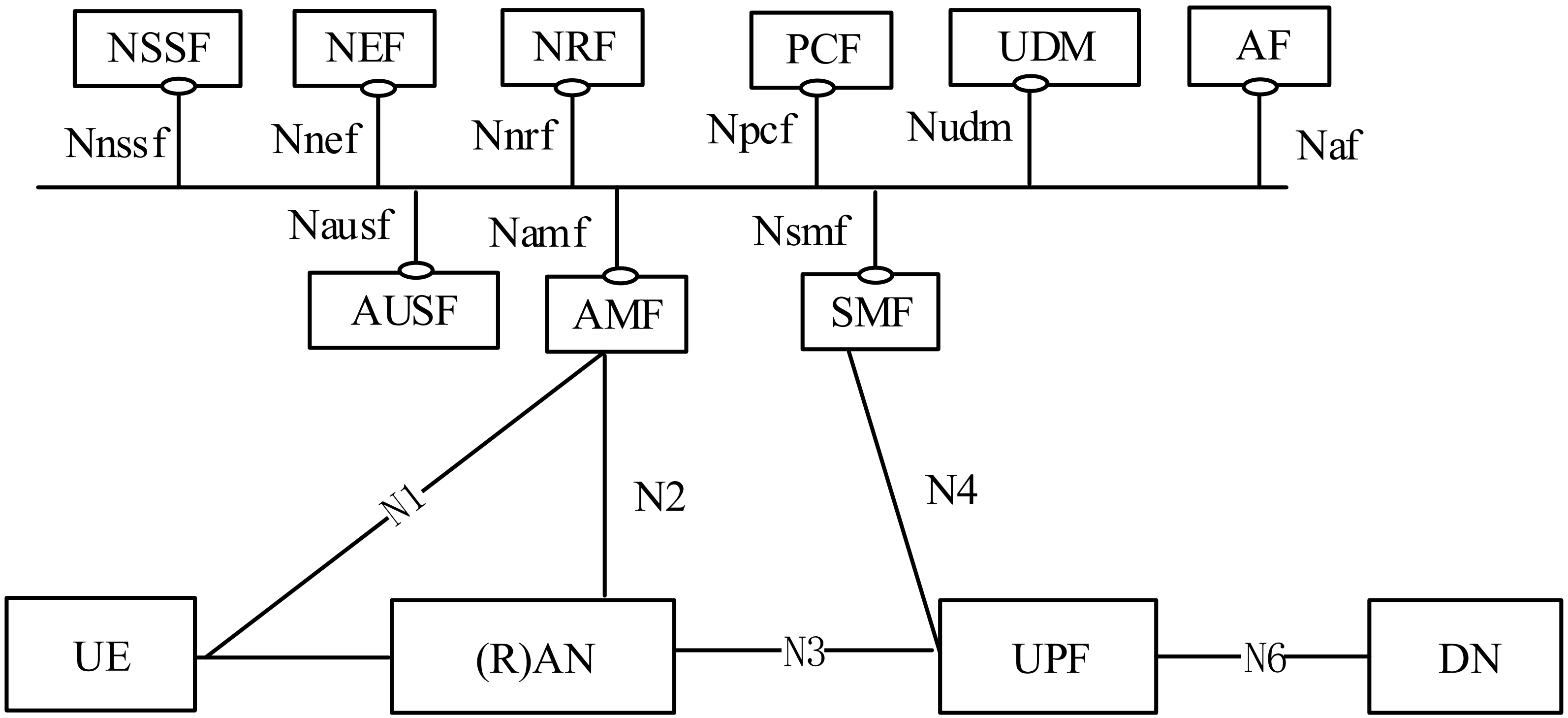
- Network Exposure function (NEF)
  - exposure of capabilities and events
  - secure provision of information from external application to 3GPP network
  - translation of internal/external information
- NF Repository function (NRF)
  - service discovery function, maintains NF profile and available NF instances
- Network Slice Selection Function (NSSF)
  - selecting of the Network Slice instances to serve the UE, determining the allowed NSSAI, determining the AMF set to be used to serve the UE

Network Data Analytics Function (NWDAF)

- It provides to the core the capability of collecting per-slice aggregated data
- It supports network optimization through the interaction with PCF
- Through push/subscribe mechanisms
- Included in Release 15, features definition ongoing

Service Based Architecture

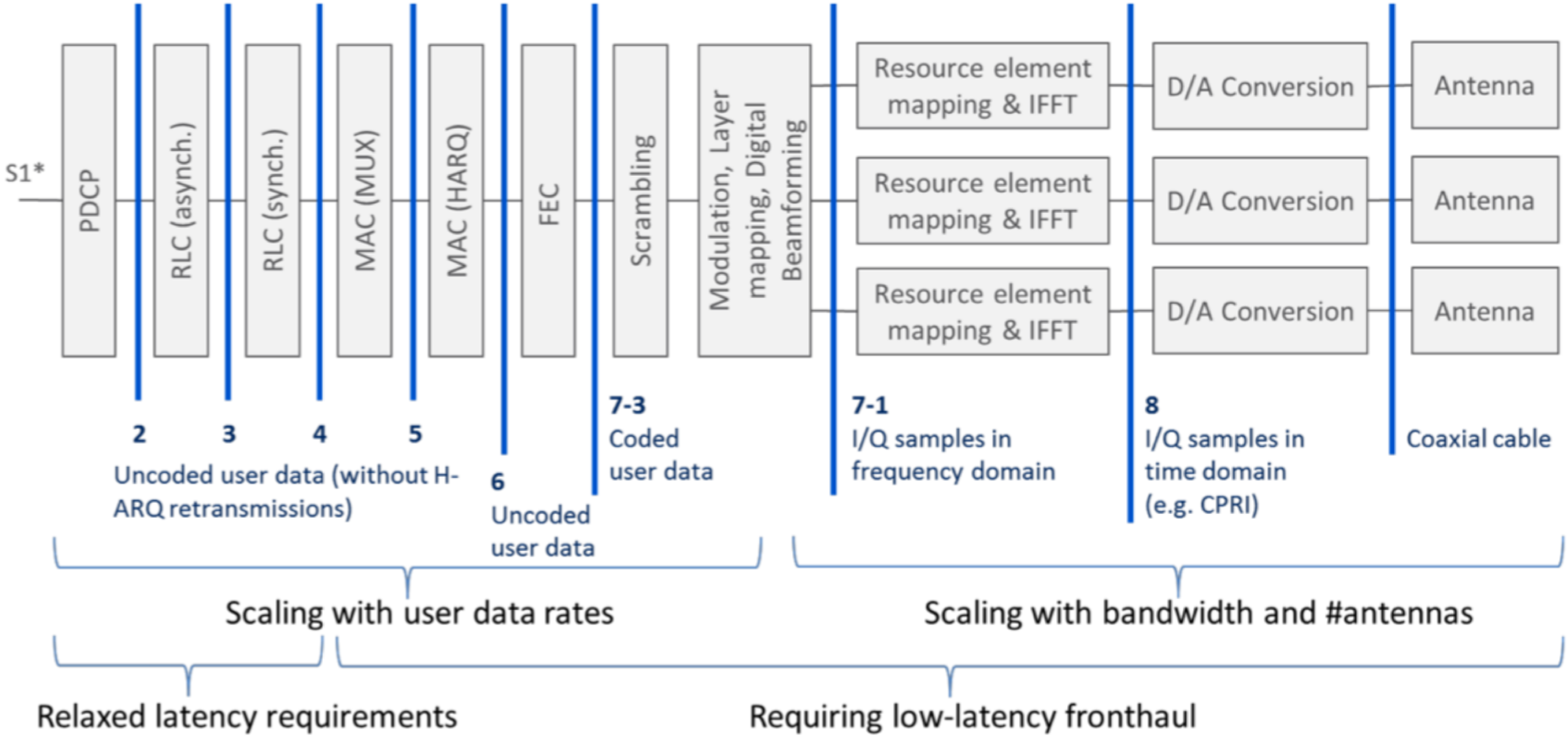
- Consequence of softwarization & modularization
  - ”Explosion” of interfaces between entities
- *Microservice* development techniques



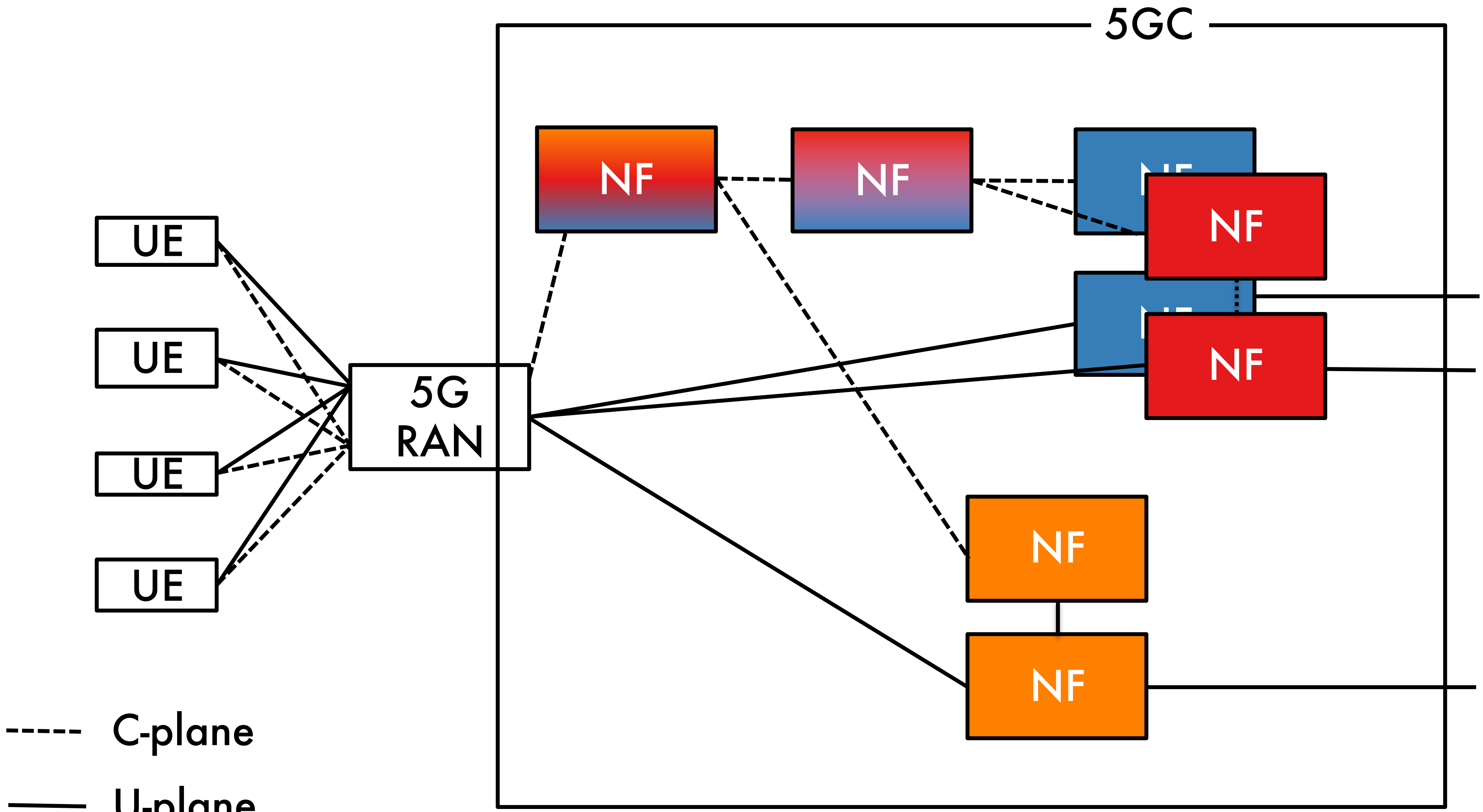
On the RAN/Core split

- Traditionally, the mobile network architecture has been divided into access and core network
- With 5G network softwarization, this division will not be as sharp as it is now
- Some functionality can run in the cloud
  – Edge or core (CPU gains)
  – Increased communication costs
- Some functionality will be shared across slices

Split options



Shared NFs across slices



Adapted from "5G System design book"

Wrap-up

- Some NFs are evolved versions of EPC functions (like MME, SGW, PGW, HSS & PCRF)
  - AMF, SMF, PCF, UDM, AF, UPF or AUSF,
  - Unlike EPC functions, AMF, SMF and UPF should be access independent
- Some NFs are new
  - Slicing (i.e., NSSF), to support handling multiple slices,
  - SBA itself (i.e., NEF & NRF) to support subscription, exposure and access
- Full CP & UP split in the core network, i.e. UPF supports UP data processing, while all other nodes act as CP functions
  - Different compared to EPC, e.g. PGW
  - This allows for independent **scaling** of CP and UP

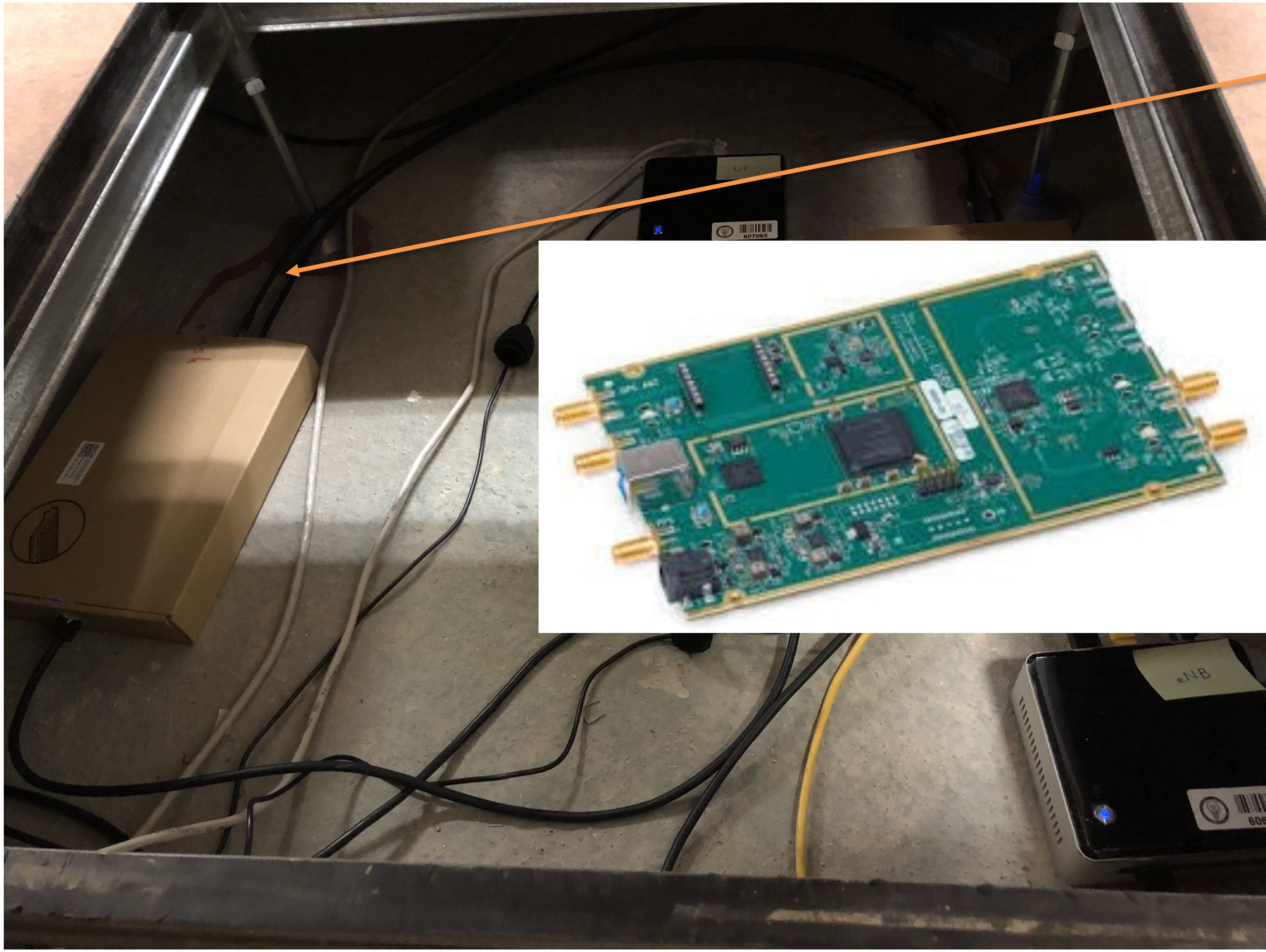
LET'S DO SOME SLICING

From: G. García-Avilés et al., "POSENS: a Practical Open-source Solution for End-to-end Network Slicing," IEEE Wir. Comm. Mag.

Ingredients

- Infrastructure
  - Radio and Core
- Virtual Infrastructure manager (VIM)
  - To run and connect virtual machines
- Network functions
- Orchestrator
  - The ability to instantiate, run, assign resources, monitor performance, etc.

Radio Part - SDR



Wires

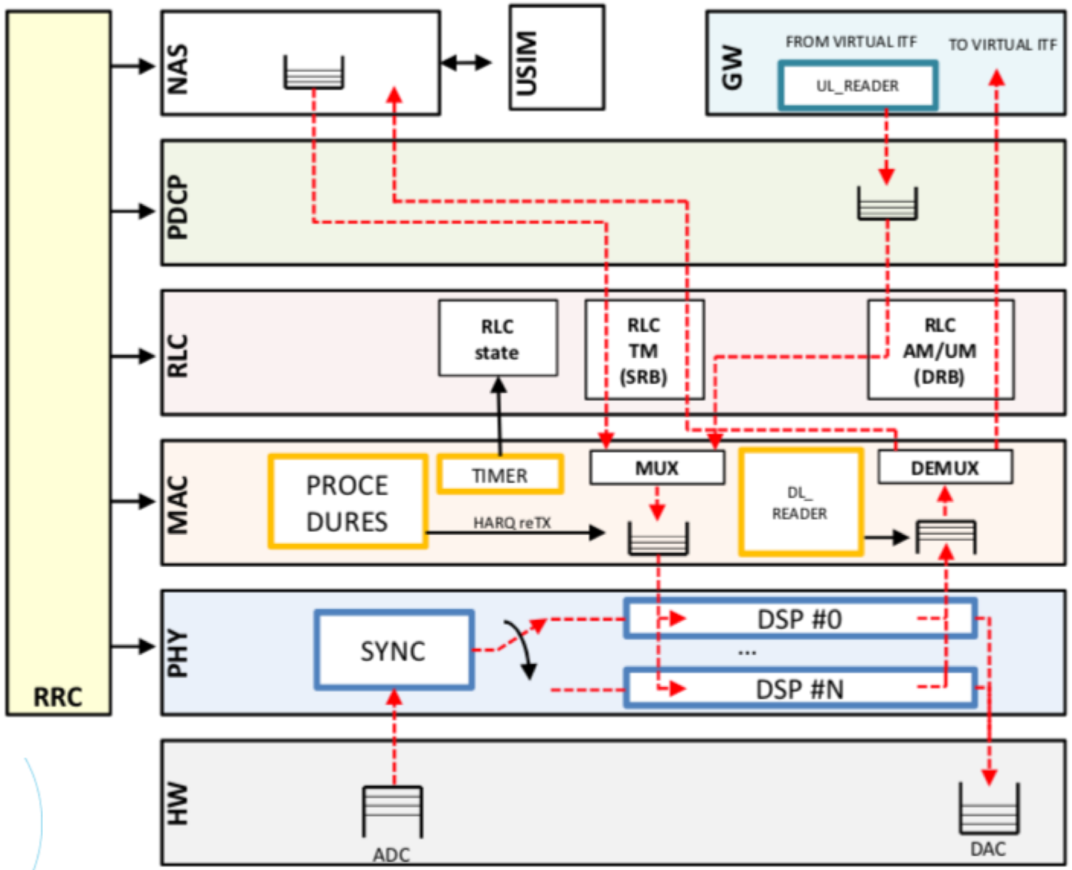
Ettus B210

Intel NUC

Radio part - Functions

- **Software** Defined Radio
  - Many components are implemented by software on a personal computer
  - “Affordable” for Academia
  - “Open” projects

- Two prevalent choices
  - srsLTE
  - OpenAirInterface



From: Gomez-Miguelez et al., “srsLTE: An Open-Source Platform for LTE Evolution and Experimentation,” ACM WiNTECH 2016

Software pitfails

- Bandwidth incompatibilities:
  - SRS supports operation with all the bandwidth settings specified by 3GPP, i.e., 1.4, 3, 5, 10, 15, and 20 MHz, OAI does not work with the 1.4, 3 and 15 MHz configurations
  - srsENB implementation does not work reliably with 20MHz channels
- Interconnection with CN
  - srsENB employs the same subnetwork for both user plane (S1-U interface) and control plane (S1-C interface).
  - OpenAir-CN can be configured to use two different subnetworks
- Problematic queue management

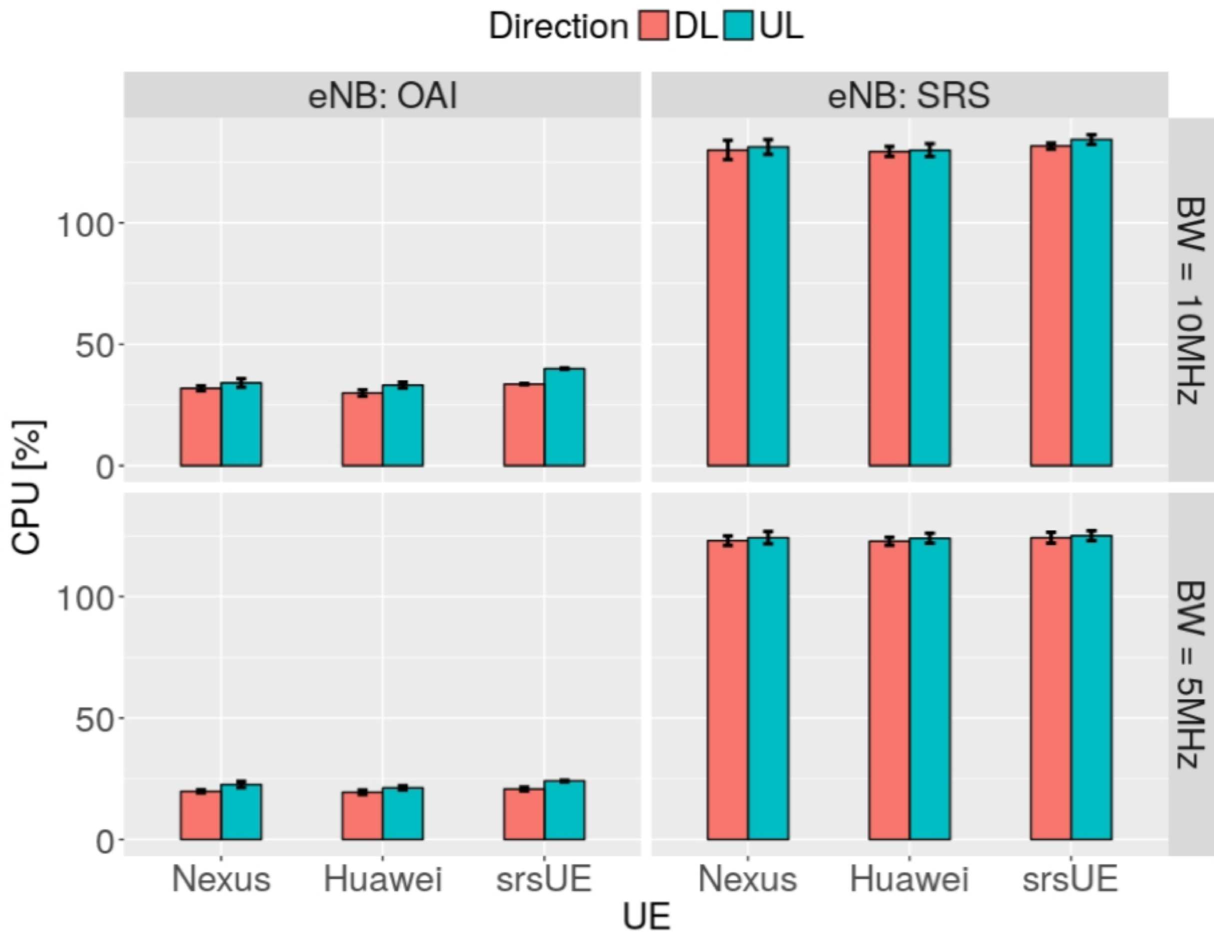
From: F. Gringoli et al., “Performance Assessment of Open Software Platforms for 5G Prototyping,” IEEE Wir. Comm. Mag.

SRS vs OAI

- srsLTE's more modular and easier to customize
- For instance: fix during experiments the MCS assignments which the eNB enforces
  - **srsLTE**: function `sched_ue` in `srsLTE/srsenb/src/mac/scheduler_ue.cc`
  - **OAI**: files `eNB_scheduler_ulsch.c` and `eNB_scheduler_dlsch.c` contain the functions `schedule_ulsch_rnti` and `schedule_dlsch_rnti`
    - We developed a patch
    - We discovered that it was altered by other functions

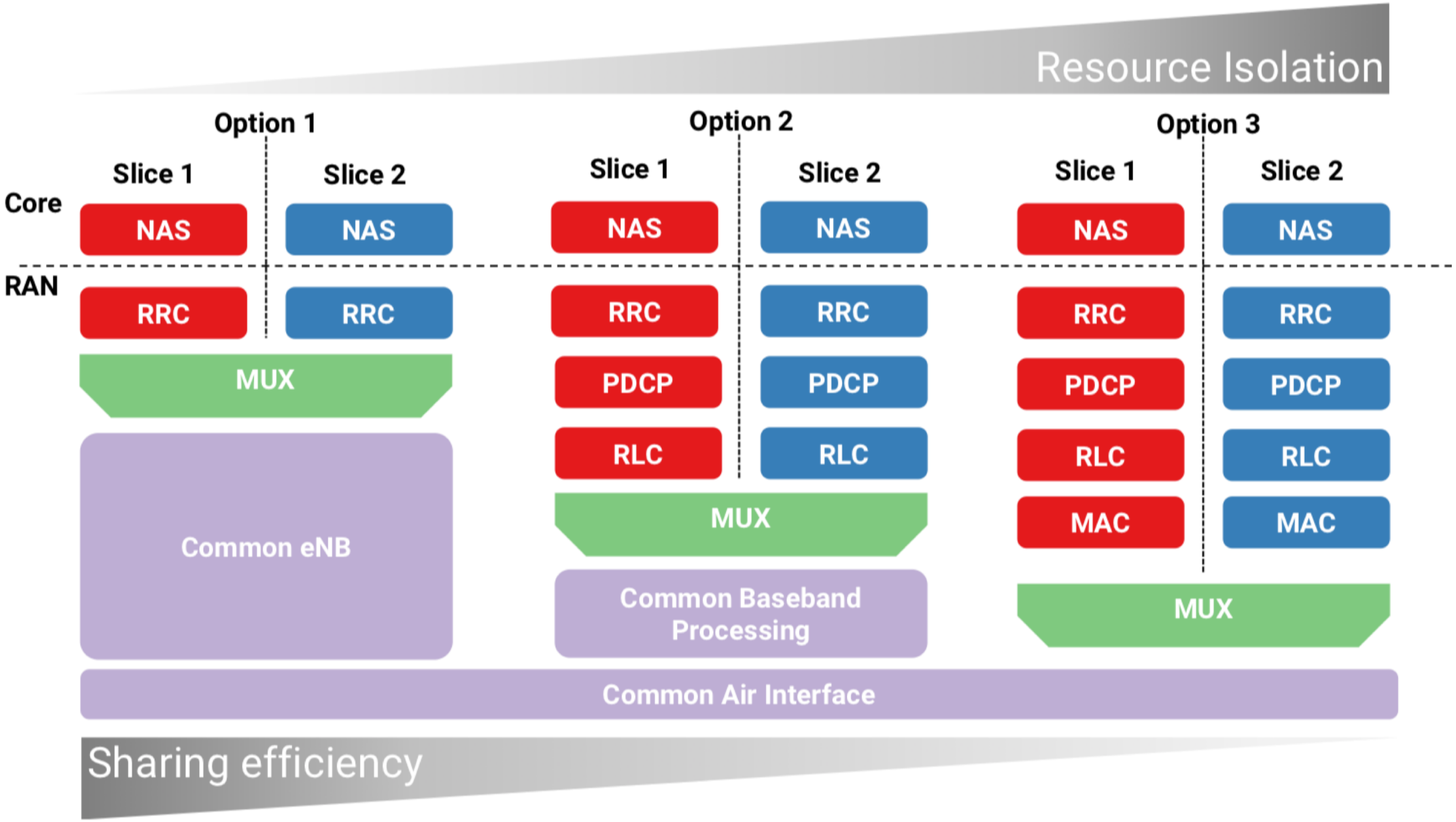
SRS vs OAI

- OAI (way) more CPU efficient



Slicing the RAN

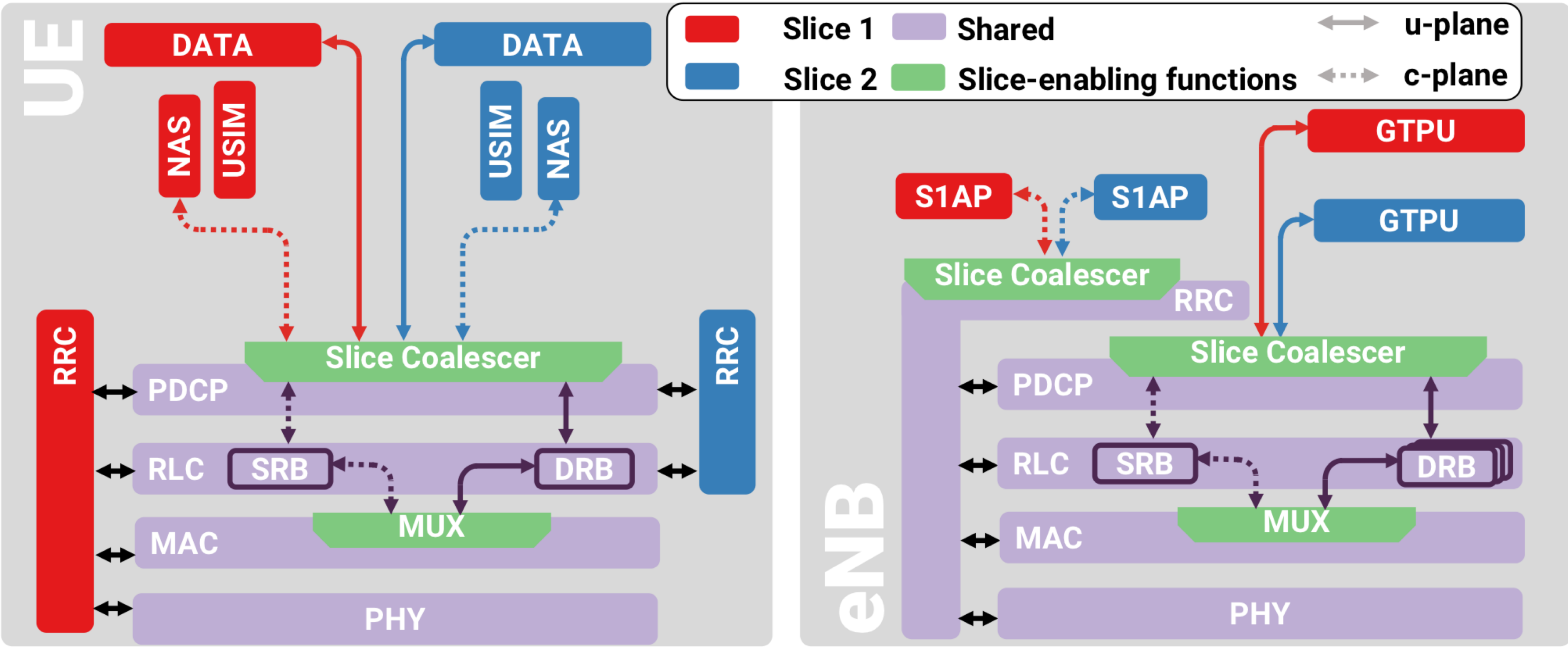
- Traffic from different tenants has to be handled over the same spectrum



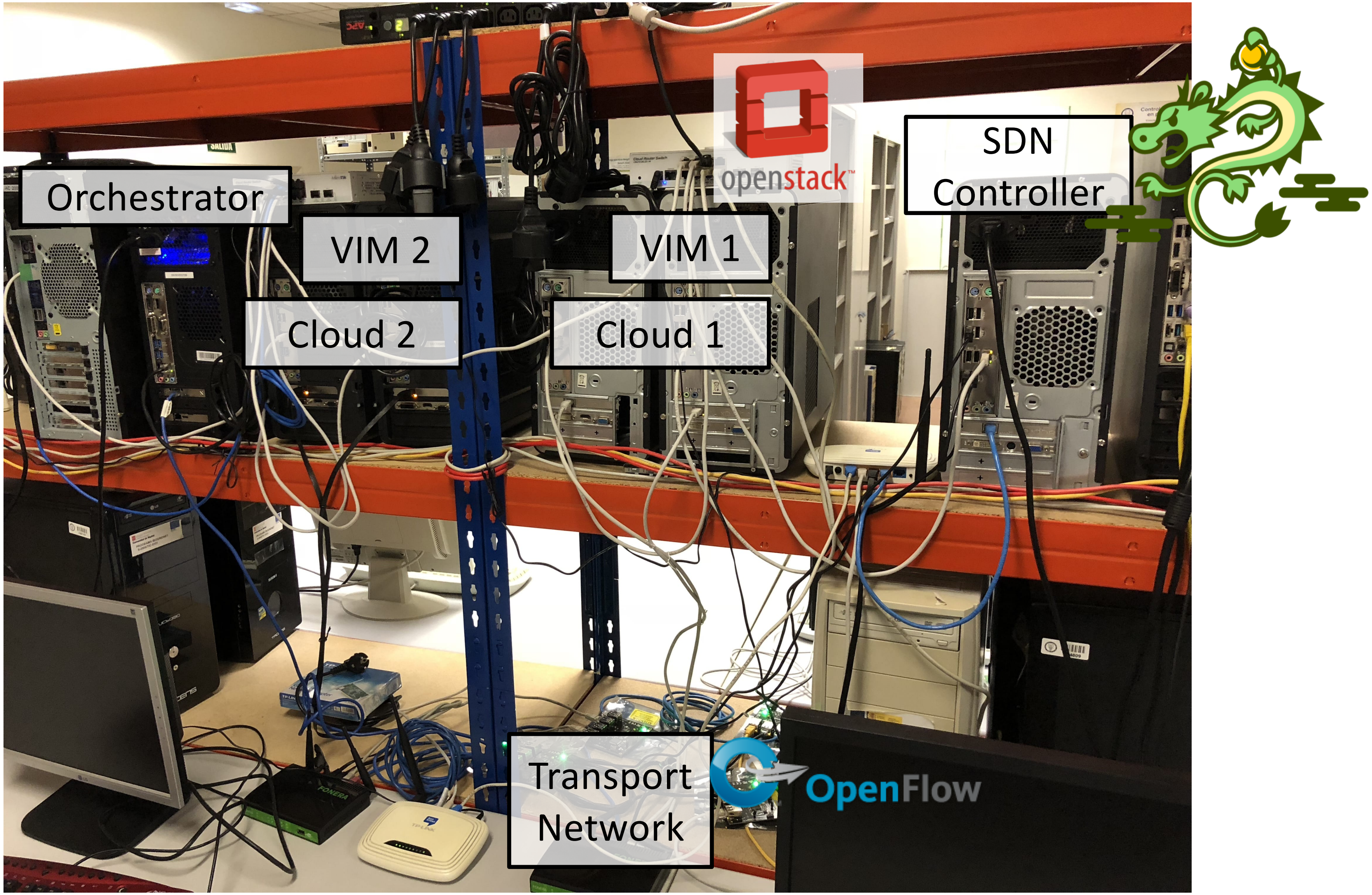
Adapted from: P. Rost et al., "Network Slicing to Enable Scalability and Flexibility in 5G Mobile Networks," in IEEE Comm. Mag.

Changes to the UE and eNB

- No changes to the UE



Core part



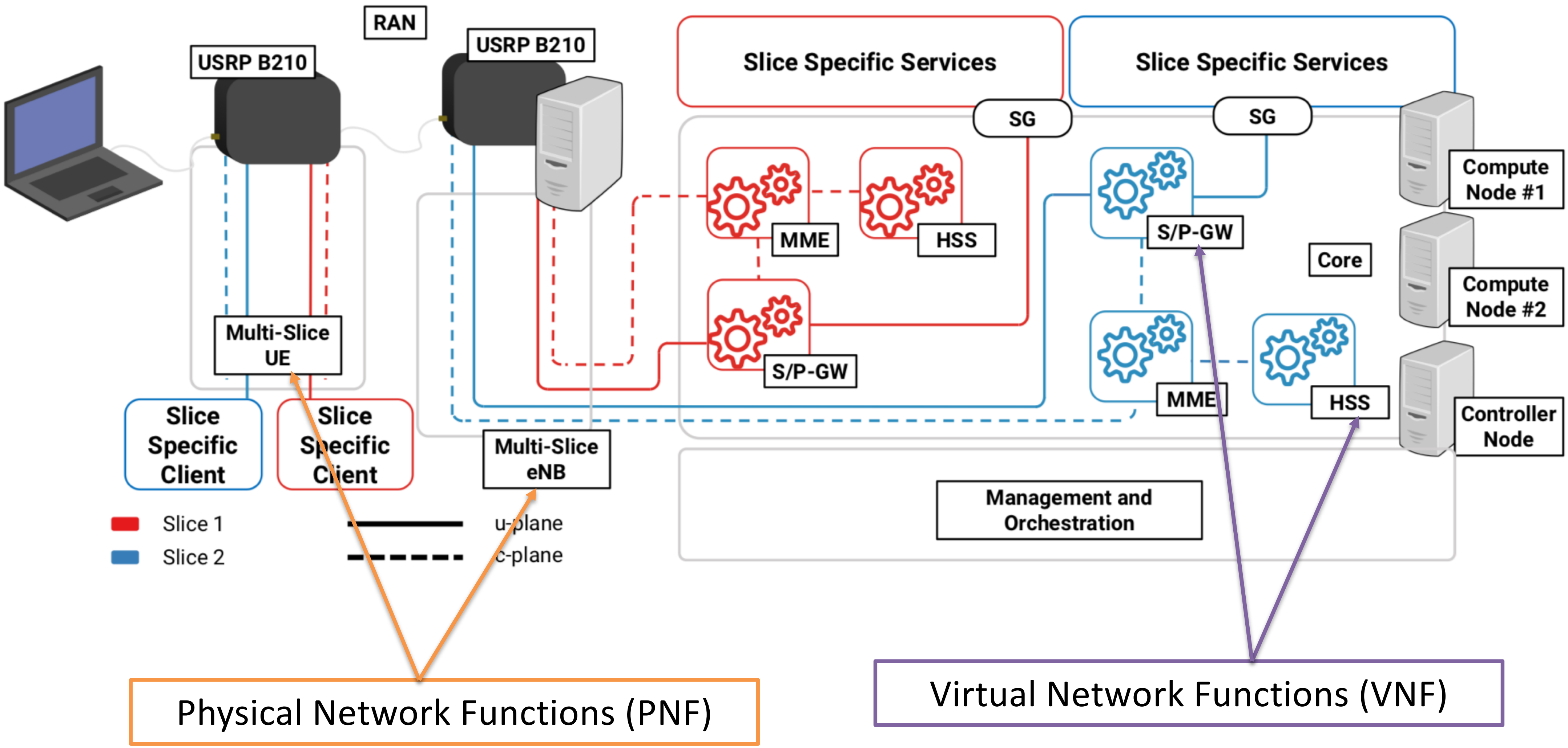
Orchestrator
VIM 2
Cloud 2
VIM 1
Cloud 1
openstack
SDN
Controller
Transport
Network
OpenFlow

Non-radio part - Functions

- Core Network functions
  - Both OAI and SRS provide working software
- Management and orchestration
  - Ad hoc, through Command Line Interface (CLI)

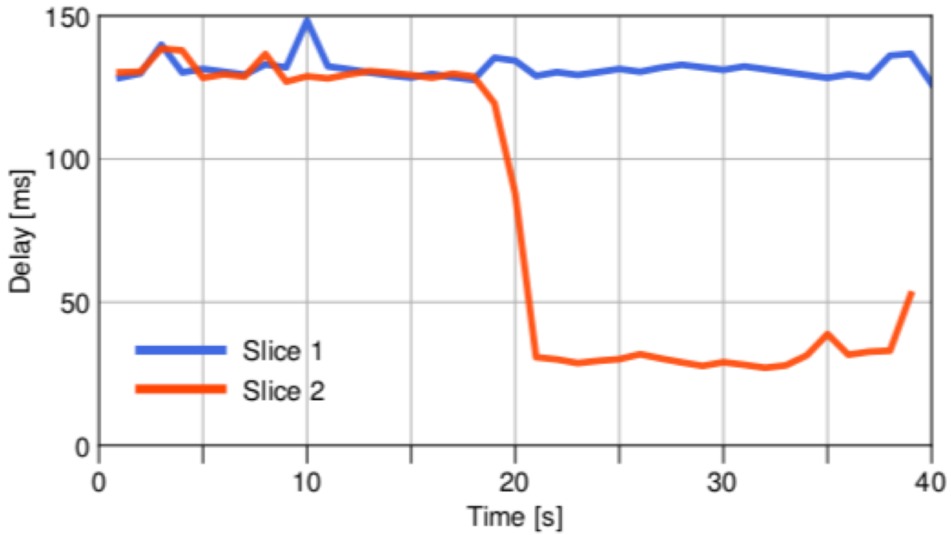
Work	Base Software	Main purpose	Main Feature	Limitations	Open Source
Mendes et al. [5]	srsLTE	RAN Slicing	Multiple, per tenant, eNB virtualization	Implementation only up to MAC layer.	No
Chang et al. [6]	OAI	RAN Slicing	Thorough evaluation of slices utilization	RAN Slicing only.	No
Foukas et al. [7]	OAI	RAN Slicing	SDN-based RAN slicing	It does not include a core.	Download upon request
Foukas et al. [8]	OAI	End-to-end slicing	Core Network handling multiple slices	Single Slice UE only.	No
POSENS	srsLTE	End-to-end slicing	Slice aware shared RAN.	One RAN split available.	Yes

Overview

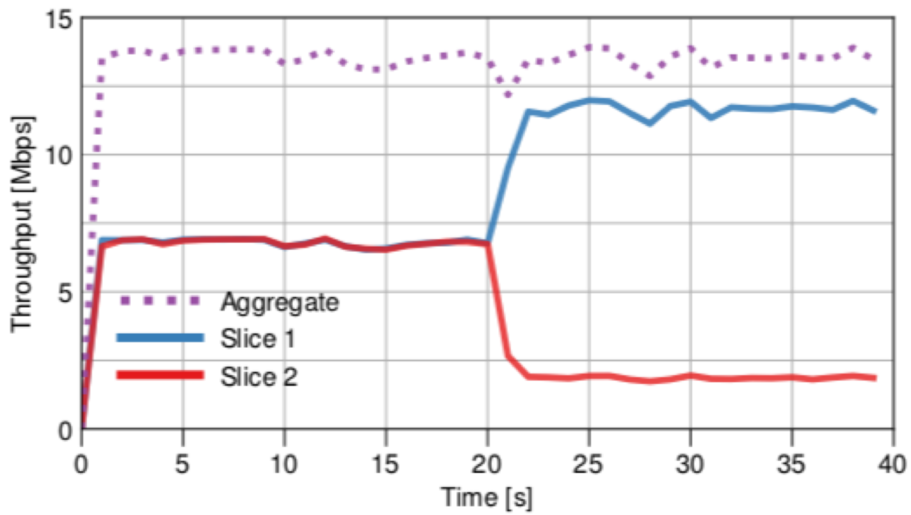

RAN
USRP B210
USRP B210
Slice Specific Services
Slice Specific Services
SG
SG
MME
HSS
S/P-GW
S/P-GW
Core
MME
HSS
Compute Node #1
Compute Node #2
Controller Node
Multi-Slice UE
Slice Specific Client
Slice Specific Client
Multi-Slice eNB
Management and Orchestration
Slice 1
Slice 2
u-plane
c-plane
Physical Network Functions (PNF)
Virtual Network Functions (VNF)

Results

- It works!
  - Function re-allocation
    - GW brought closer
  - Service re-composition
    - Added shaper + fw



- Also with a Nexus-5
  - Equipped with a programmable SIM card



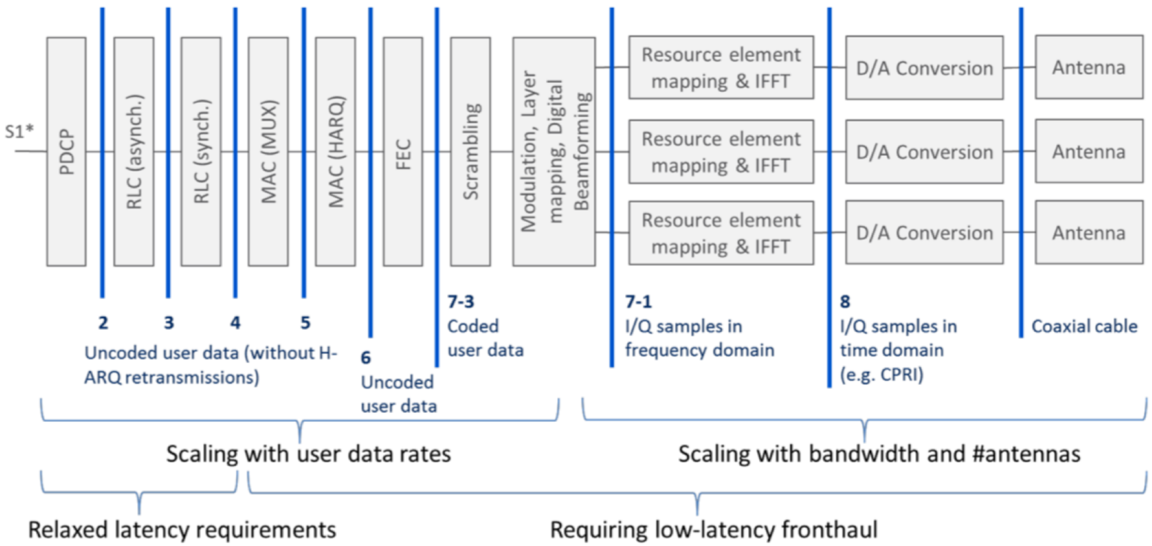
Some Orchestrators

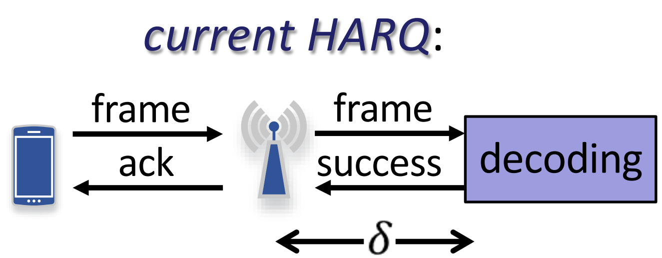
- OSM - “Open Source MANO”
  https://osm.etsi.org/
- ONAP-“Open Network Automation Platform,” https://www.onap.org/
- Cloudify http://cloudify.co/
- Open Baton- Fraunhofer and T. Berlin http://openbaton.github.io/
- XOS -“A Service Abstract Layer for CORD,” http://xos.wpengine.com/
- VLSP “Very Lightweight Network & Service Platform” University College London, http://clayfour.ee.ucl.ac.uk/usr/index.html

RESEARCH CHALLENGES

Protocol stack redesing

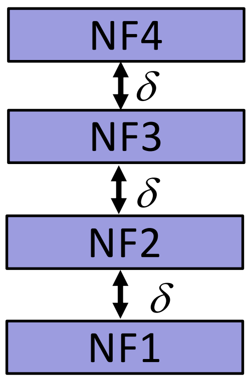
- Current protocol stack
  - Designed considering that certain functions are co-located and can exchange data with no latency
  - This has introduced temporal dependencies that limit the flexibility in the placement





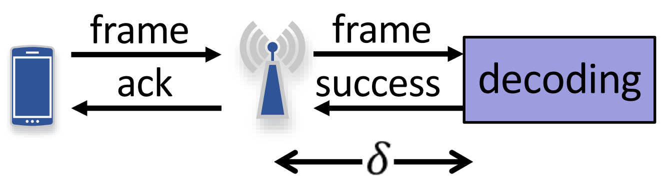
Removing tight constrains

Current stack:

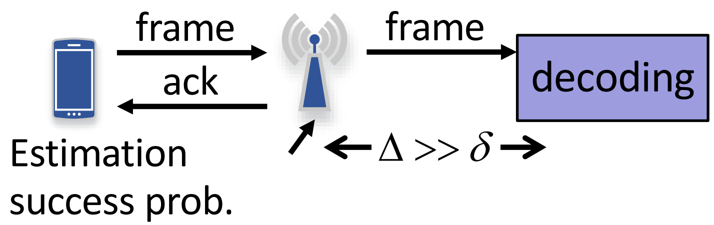


Network functions

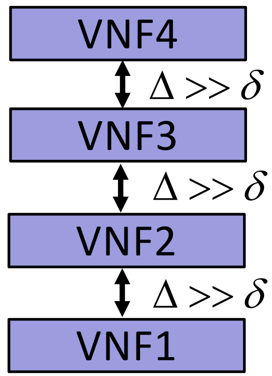
current HARQ:



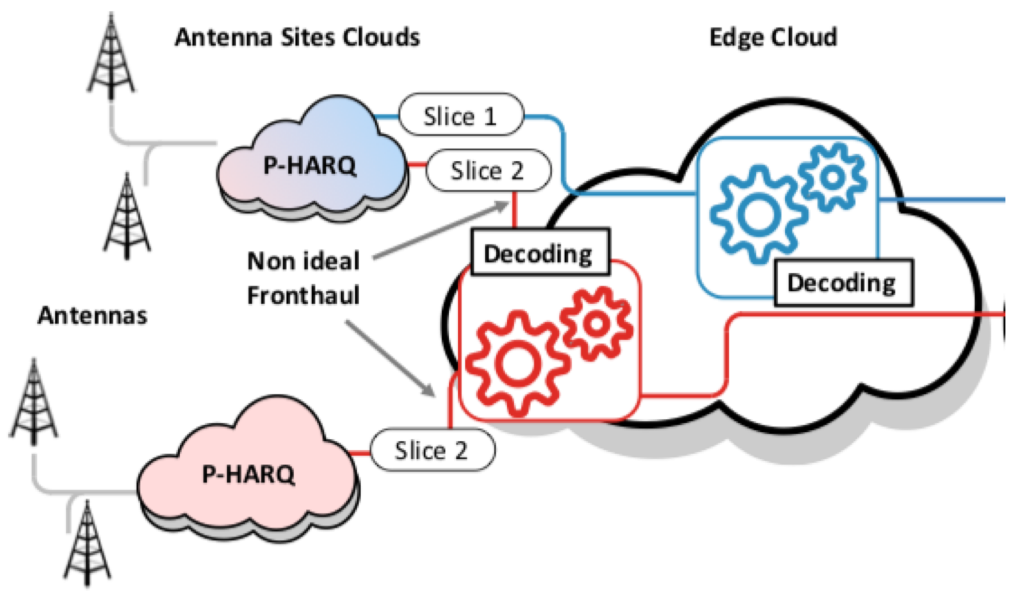
Opportunistic HARQ:



Redesign:



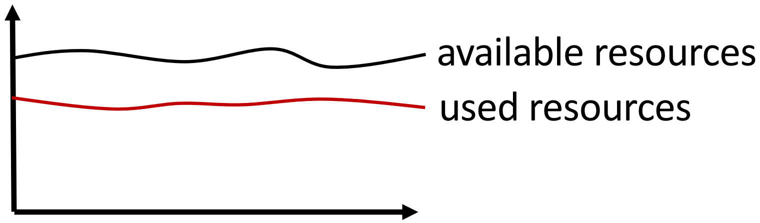
Virtual Network functions



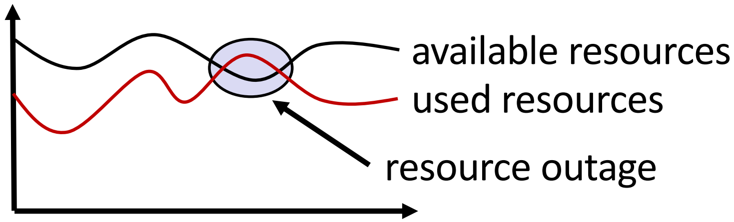
P. Rost and A. Prasad, "Opportunistic Hybrid ARQ – enabler of centralized-ran over non-ideal backhaul," IEEE Wir. Comm. Mag

Bringing resource-awareness

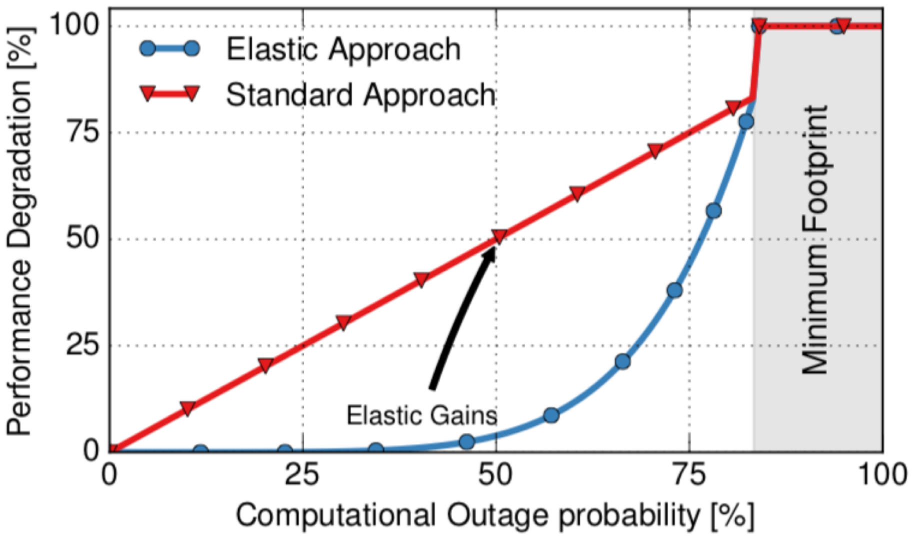
- Traditional environment for wireless functions
  - Capacity and computational load are known: Resources are always available



- Environment with virtualized wireless functions
  - Fluctuations on the capacity and load: resource outage may occur



Graceful performance degradation



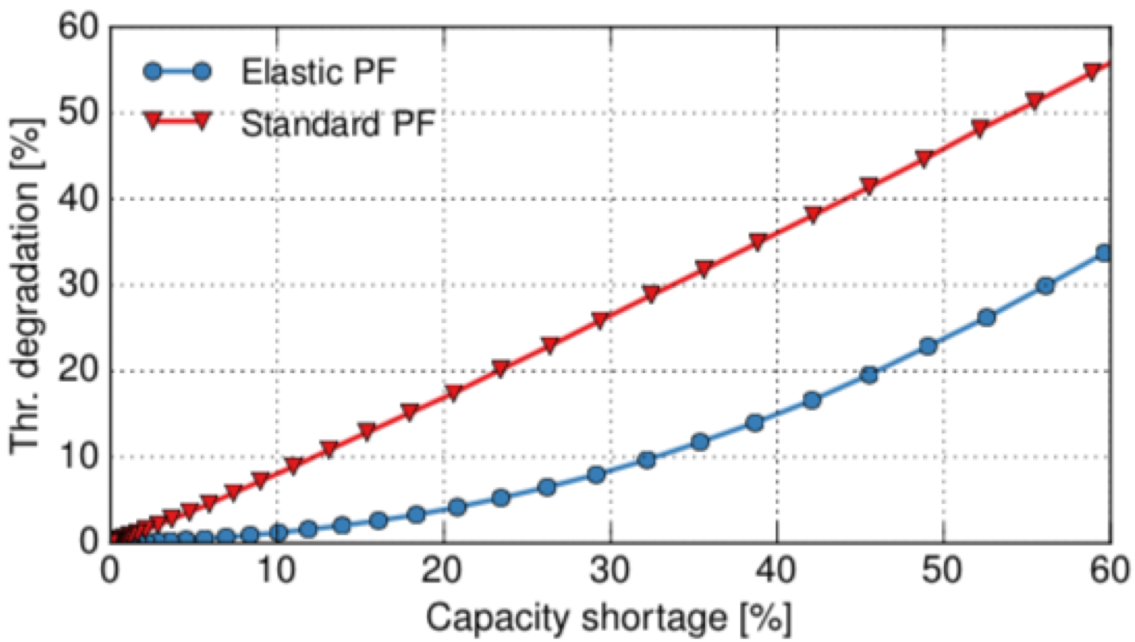
P. Serrano et al., "The path towards a cloud-aware mobile network protocol stack", Trasactions on ETT

Computational-aware Scheduling

- Traditional scheduler: dimension for maximum capacity (i.e., all users using highest MCS)
  - Resource wastage if this rarely happens
- Novel approach: assume limited capacity

Algorithm:

1) Schedule
2) Compute CPU cost. If feasible, done
3) Discard users until feasible
4) Add non-scheduled users (if any)
5) Go-to 2

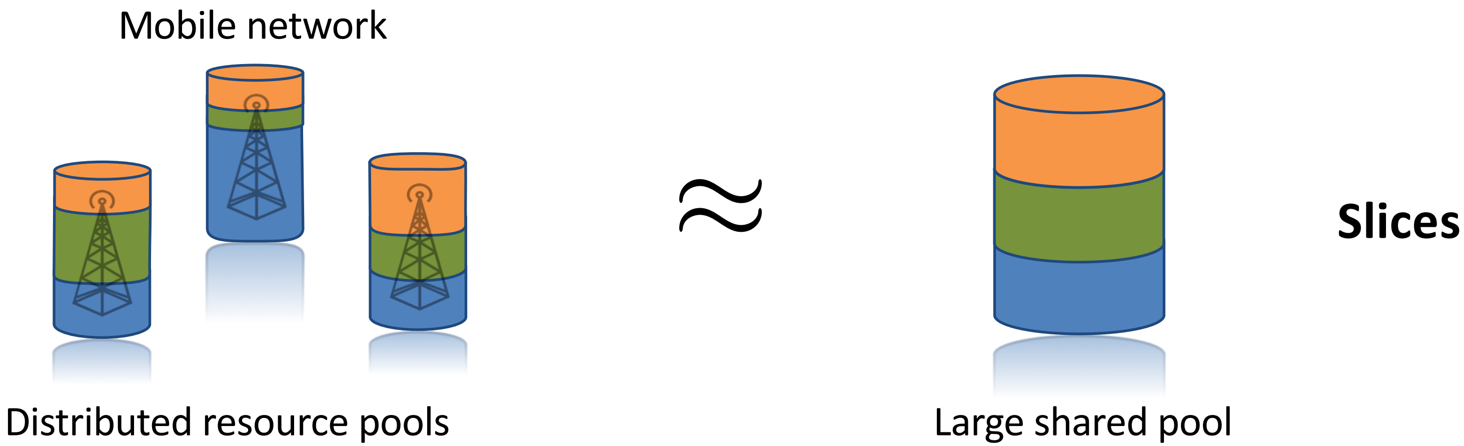


Resource allocation

- How to share resources among slices
  - This includes computational resources as well as radio resources
  - Focus on the (more scarce) radio resources
- In the context of RAN sharing, different strategies are possible [3GPP TR 22.852]
  - Share-based resource allocation: resources are allocated based on **predetermined network shares**
    - **E.g.,** which represent the tenants' level of contribution to the network
  - Reservation-based resource allocation: tenants issue requests for resources to the infrastructure provides

Share-based Slicing

- Resource allocation strategy
  - Network resources are "pooled"
  - Tenants can buy a share (portion) of the pool
  - Resources are allocation proportionally to the shares



Model

- Users move, re-associate with resources, which are given given shares

$\mathcal{U}$
users

$\mathcal{B}$
$\{b_1, b_2, \ldots, b_{|\mathcal{B}|}\}$
resources

$\ldots$

$\mathcal{O}$
$\{o_1, o_2, \ldots, o_{|\mathcal{O}|}\}$
slices/tenants

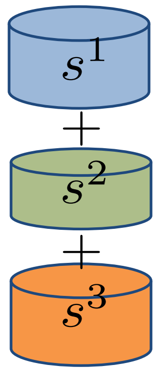
at&t

verizon

$\ldots$

Sprint

Each $o \in \mathcal{O}$ is assigned a share:



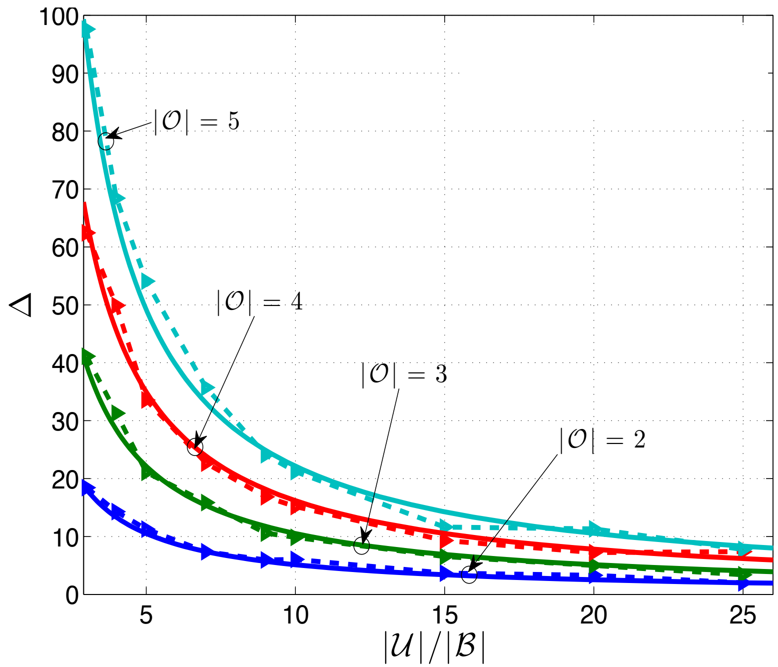
$$s_o \in [0, 1] \text{ s.t. } \sum_{o \in \mathcal{O}} s_o = 1$$

P. Caballero, et al. "Multi-Tenant Radio Access Network Slicing: Statistical Multiplexing of Spatial Loads," IEEE ToN '18

Results: capacity savings

- Capacity savings: extra resources needed by a static approach to achieve the same utility
- **Theorem:** Given homogeneous slices with random loads and user capacities, the capacity savings are:

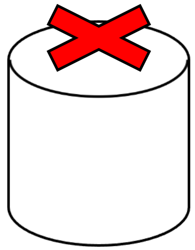
$$\Delta_o \approx \exp\left[\underbrace{\overbrace{\frac{|\mathcal{B}|}{2|\mathcal{U}^o|}}}\underbrace{\overbrace{(1-s_o)}}\right]-1$$



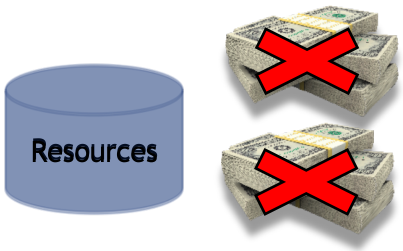
Sharing gains increase with # slices and with low loads.
This is expected to hold in future networks (small cells)

Reservation-based Slicing

When accept/reject to maximize InP's revenue?

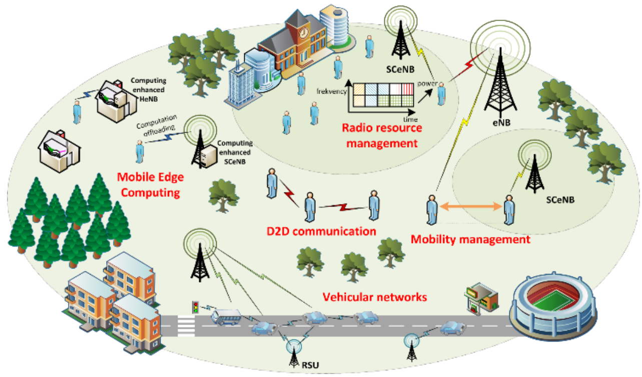


NETFLIX



YouTube

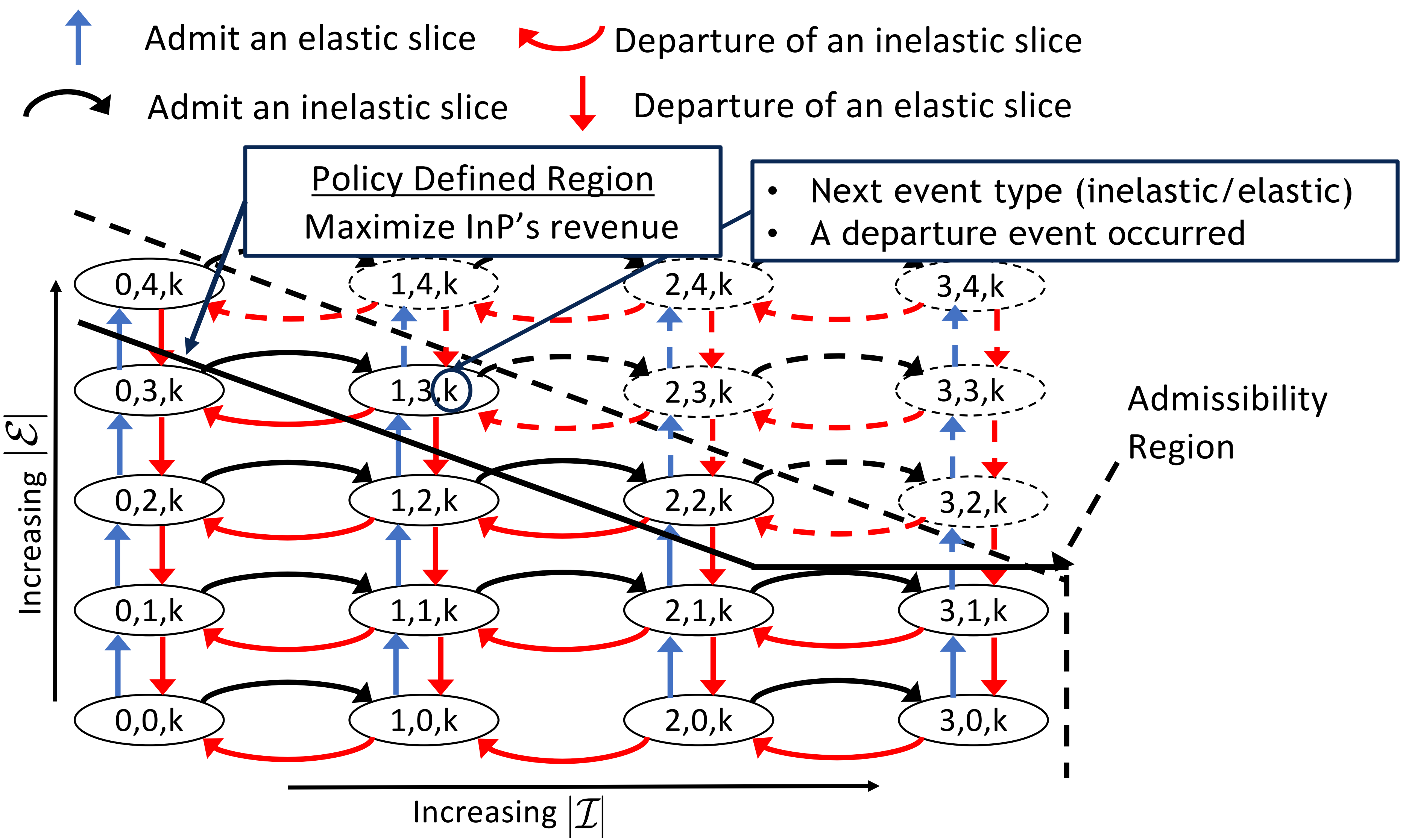
at&t



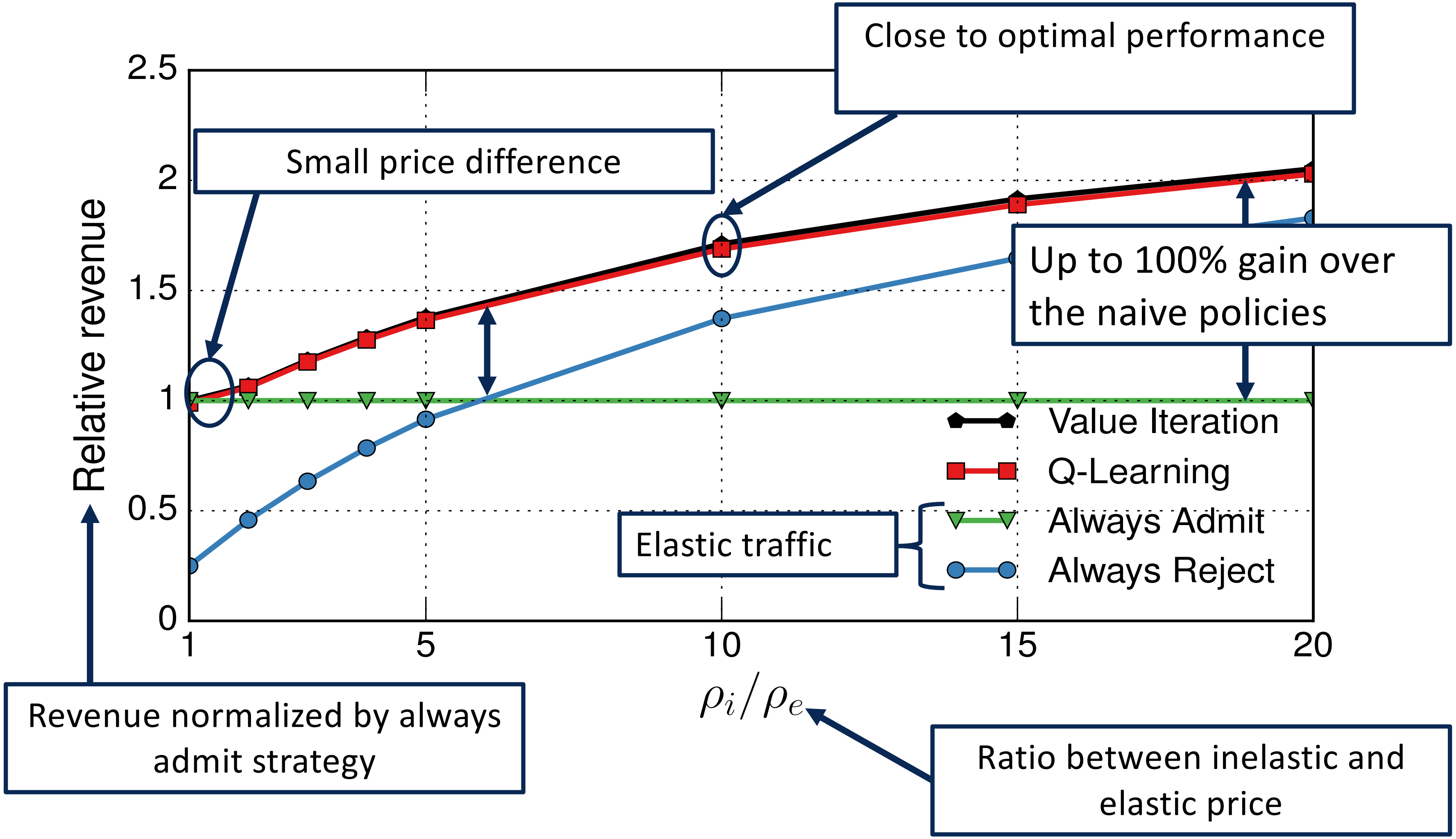
Overprovisioning is not possible

V. Sciancalepore et al., "Mobile Traffic Forecasting for Maximizing 5G Network resource Utilization," IEEE INFOCOM 2017

Semi-Markov Decision Process



Results: Revenue

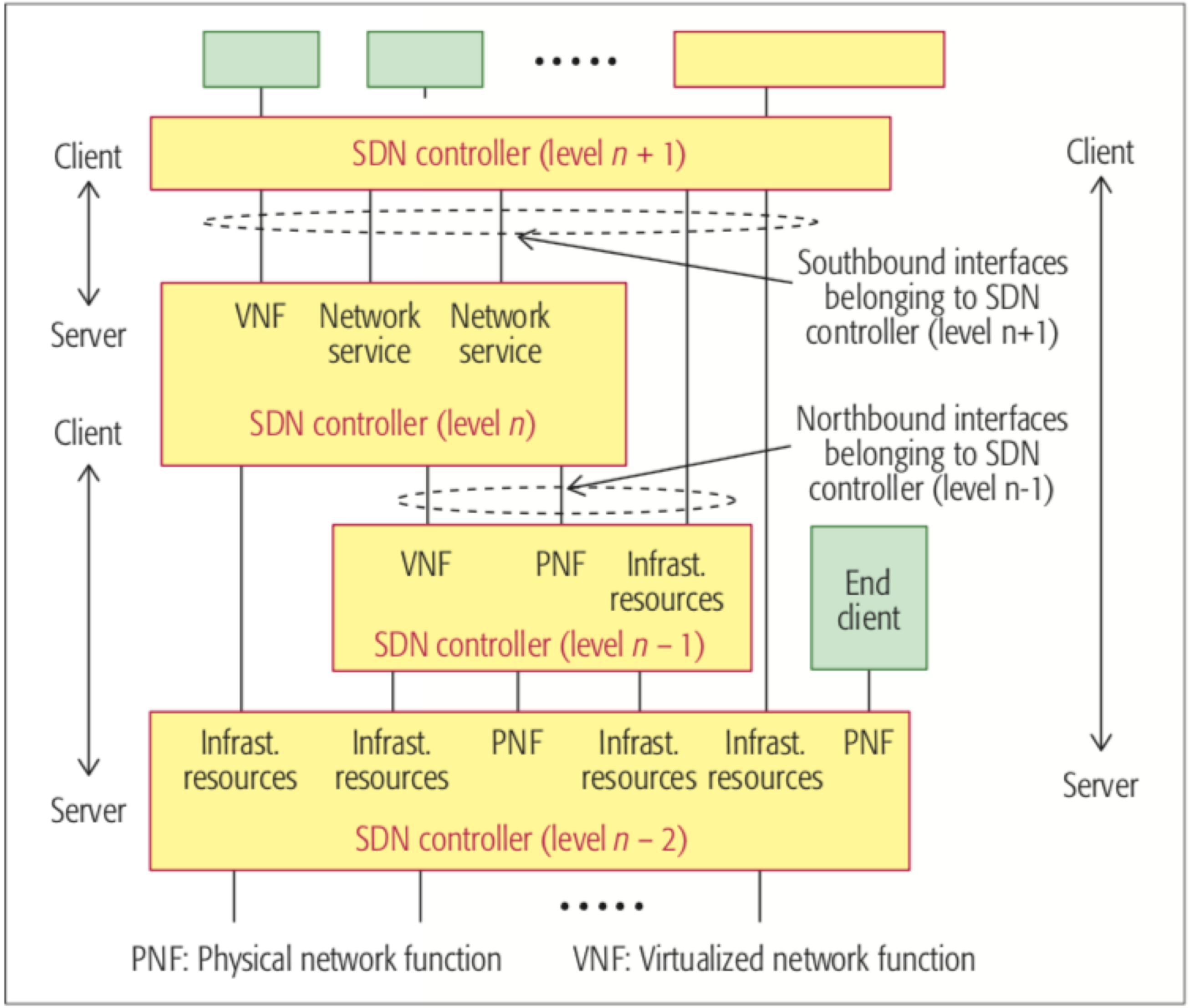


More challenges

- Security and privacy (shared resources)
  - Multi-level security framework
  - Active networking (late 90’s)
  - Resource isolation
- Performance Issues
  - Partial sharing
- Management framework
  - Interfaces, information models, complexity

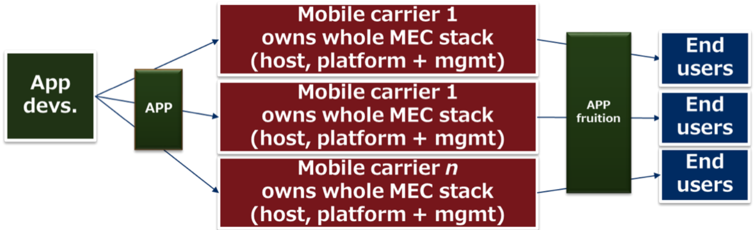
Complex Interactions

- Recursion (virtualization) in the SDN plane

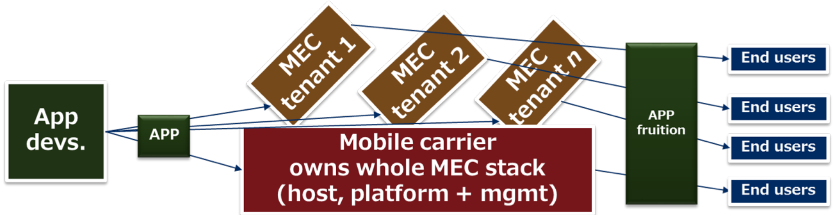


MEC Slicing

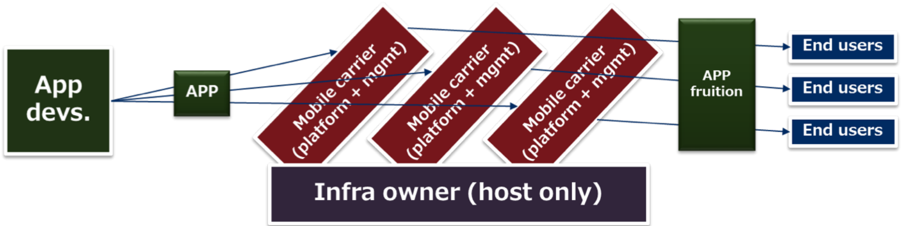
- Legacy: mobile carrier offers MEC in IaaS/PaaS to MEC app devs



- Exploring multi-tenant MEC system
  - MECaaS: telco carrier owns full MEC stack (infra + MANO). Tenants receive credentials and privileges to operate it



  - Multi-tenant IaaS: edge cloud provider owns infra only and no MEC-MANO stack. Customers (mobile carriers) attach their MEC-MANO to the deployed system



Challenges: Std. Status (July 18)

- SA1 (SMARTER - TR 22.891) - TR 22.863, TR 22.861, TR 22.862
  SA2 (NextGen TR 23.799 – Slicing TS 23.501/TS 23.502)
  SA5 (Slicing Orchestration TR 28.801)
- Slicing concepts/definitions, 3GPP use cases, Slice automation, Slicing challenges (Many drafts currently)
- End-to-End network Slicing / FMC Slicing (Study - Huawei)
  Network Sharing – APIs (Vodafone)
- Softwarization – Gap analysis – TD-208
  Network Virtualization Framework - ITU-T Y.3011
- Multi-Access Edge computing
  (Extended to accommodate fixed-access)
- SDN Architecture TR-521 (Service-oriented approach)
  Network Slicing TR-526

Summary of NSO implementations

Solution	Leader	VNF Definition	Resource Domain				MANO			Interface Management			Domain	
			Cloud	SDN	NFV	Legacy	NFVO	VNFM	VIM	CLI	API	GUI	Single	Multiple
ARIA TOSCA	Apache Foundation	TOSCA	✓							✓	✓		✓	
Cloudify	GigaSpace	TOSCA	✓		✓		✓	✓		✓	✓	✓	✓	
ESCAPE	FP7 UNIFY	Unify	✓	✓	✓		✓		✓	✓	✓		✓	✓
Gohan	NTT Data	Own	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	
ONAP	Linux Foundation	HOT, TOSCA, YANG	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Open Baton	Fraunhofer / TU Berlin	TOSCA, Own	✓		✓		✓	✓		✓	✓	✓	✓	
OSM	ETSI	YANG	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	
Tacker	OpenStack Foundation	HOT, TOSCA	✓		✓		✓	✓		✓	✓	✓	✓	
TeNOR	FP7 T-NOVA	ETSI	✓	✓	✓		✓				✓	✓	✓	
X-MANO	H2020 VITAL	TOSCA			✓		✓				✓	✓		✓
XOS	ON.Lab	-	✓	✓	✓			✓			✓	✓	✓	✓

From: de Sousa, N. F. S., Perez, D. A. L., Rosa, R. V., Santos, M. A., & Rothenberg, C. E. “Network Service Orchestration: A Survey,” arXiv preprint

SUMMARY & CONCLUSIONS

Network Slicing

- Architectural model for efficient network sharing, via slices.
- Slice: set of connected resources that has been designed to specifically meet the needs of service
- Central technology for 5G networks
  - Efficient support for very diverse services
  - Builds on SDN and NFV
  - Better support for Multi-Tenancy
    - New mobile network ecosystem

The future of Network Slicing

- Network slicing will be adopted, it enables a more efficient use of the infrastructure.

- Several efforts aligned
  - Evolution of SDN, NFV
  - Adopted by 5G, IETF, etc.
  - Many ongoing initiatives (incl. Open source)

Challenges

- Cost/benefit analysis

- Technical challenges
  - Orchestration of efforts
  - Security, Privacy, Resiliency
  - Trials (e.g., 5G-VINNI, 5TONIC)
  - Use of Machine Learning / AI techniques

Some Recent References

Surveys

- J. Ordonez-Lucena et al., “Network Slicing for 5G with SDN/NFV: Concepts, Architectures, and Challenges,” IEEE Comm. Magazine, May 2017
- de Sousa, N. F. S., Perez, D. A. L., Rosa, R. V., Santos, M. A., & Rothenberg, C. E. “Network Service Orchestration: A Survey,” arXiv preprint

Research

- P. Serrano et al., “The path towards a cloud-aware mobile network protocol stack,” Transactions on Emerging Telecommunications Technologies
- C. Marquez et al., "How Should I Slice My Network? A Multi-Service Empirical Evaluation of Resource Sharing Efficiency," ACM MobiCom ’18
- J.X. Salvat et al, "Overbooking Network Slices through Yield-driven End-to-End Orchestration," ACM CoNEXT '18
- P. Caballero, et al. “Multi-Tenant Radio Access Network Slicing: Statistical Multiplexing of Spatial Loads,” IEEE Trans. on Networking ‘18

Practice

- G. García-Avilés et al., “POSENS: a Practical Open-source Solution for End-to-end Network Slicing,” IEEE Wireless Comm. Magazine, Special Issue on 5G Testing and Field Trials (soure code: https://github.com/wnluc3m)
- F. Gringoli et al., “Performance Assessment of Open Software Platforms for 5G Prototyping,” IEEE Wireless Comm. Magazine, Special Issue on 5G Testing and Field Trials

A Primer on 5G Network Slicing:
Concepts, Algorithms, and Practice

Pablo Serrano
twitter.com/pbluc3m
www.it.uc3m.es/pablo/
pablo@it.uc3m.es

Thanks! Questions?

Many thanks to: Albert Banchs, Carlos J. Bernardos, Andrés García-Saavedra, Marco Gramaglia, Yan Grunenberger, Diego López, Vincenzo Sciancalepore