NOKIA

RAN slicing as enabler for low latency services

Presented by A. Maeder, NOKIA Bell Labs
Contributions by Z. Li, P. Rost, C. Sartori, A. Prasad, C. Mannweiler

ITG 5.2.4 Fachgruppentreffen

Dresden, June 10th, 2016

Outline

Low latency-services

- Which services are we talking of?
- Requirements

What is “RAN slicing” and why do we need it?

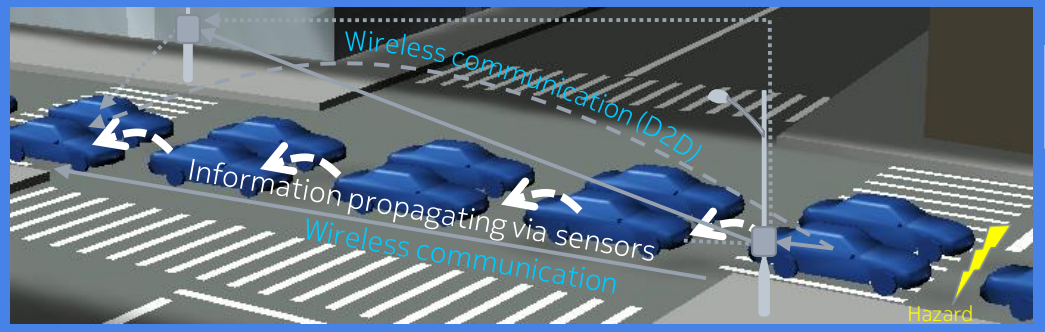
- Principles of RAN slicing
- Architecture enablers

Enabling low latency services in a common infrastructure

Outlook

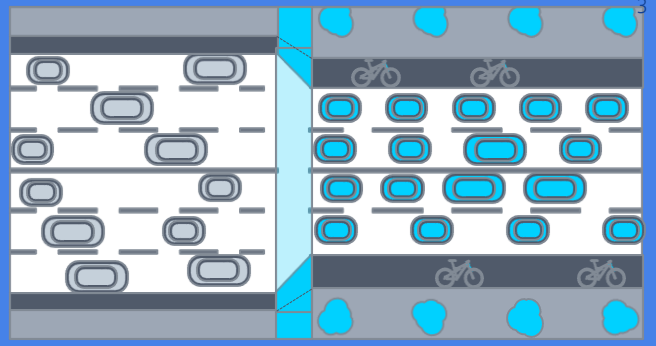
Low-latency use cases

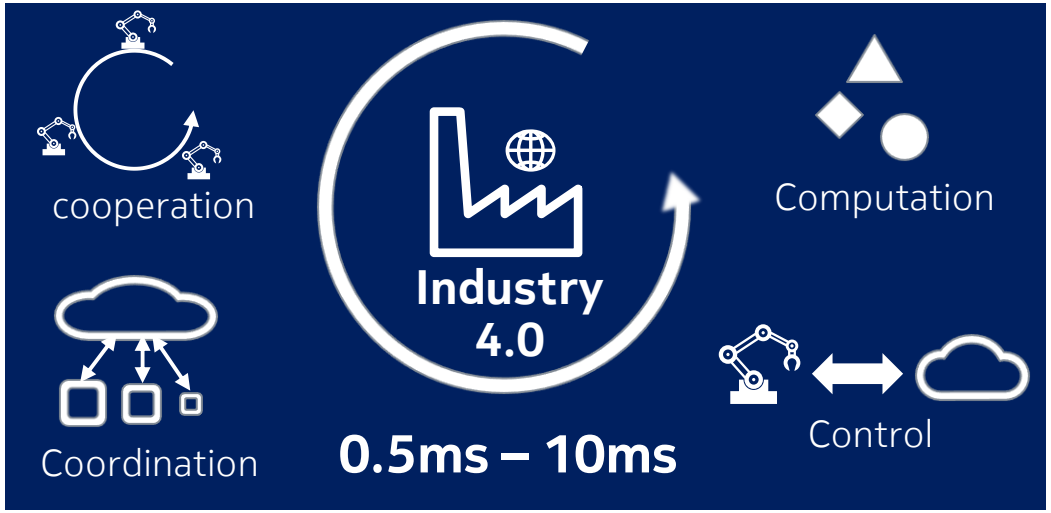
Automotive



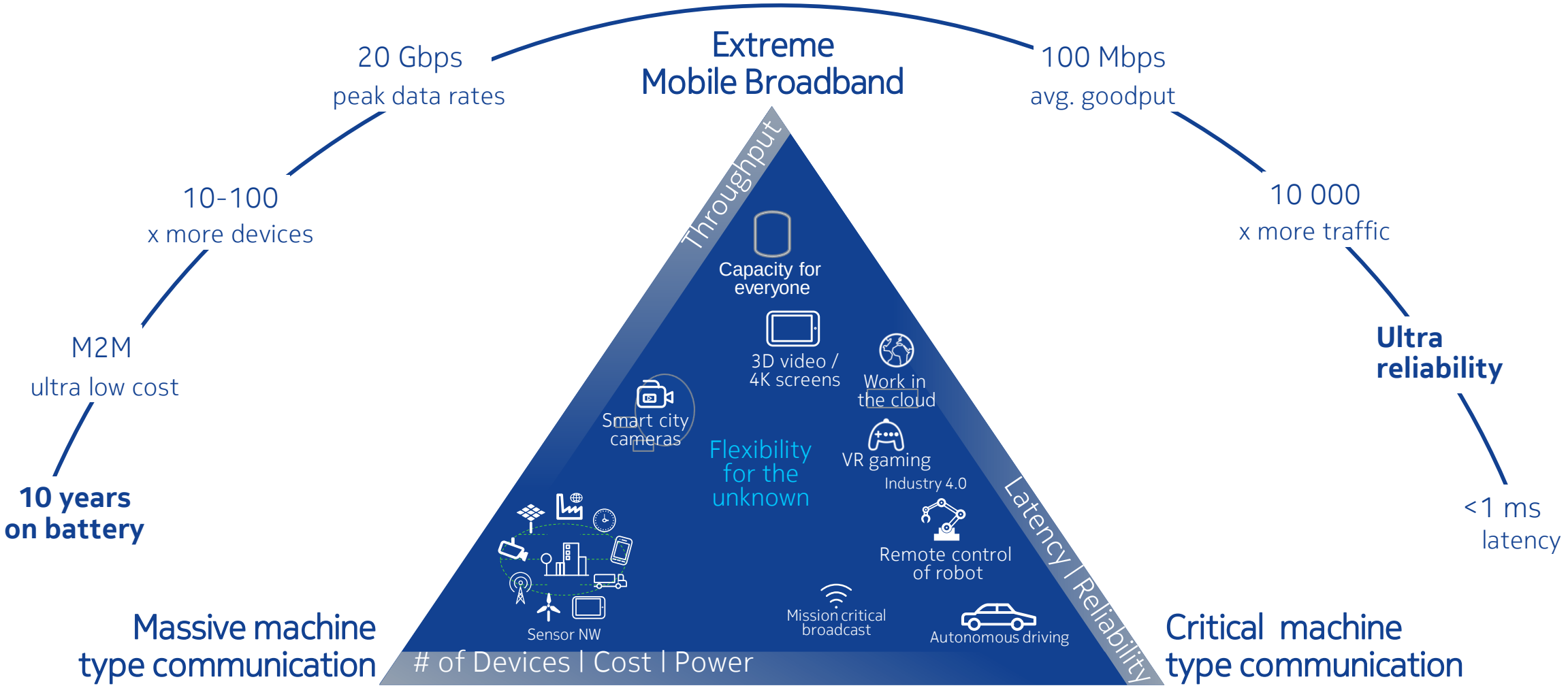
<3ms

4x more cars possible
Safer and faster individual transport



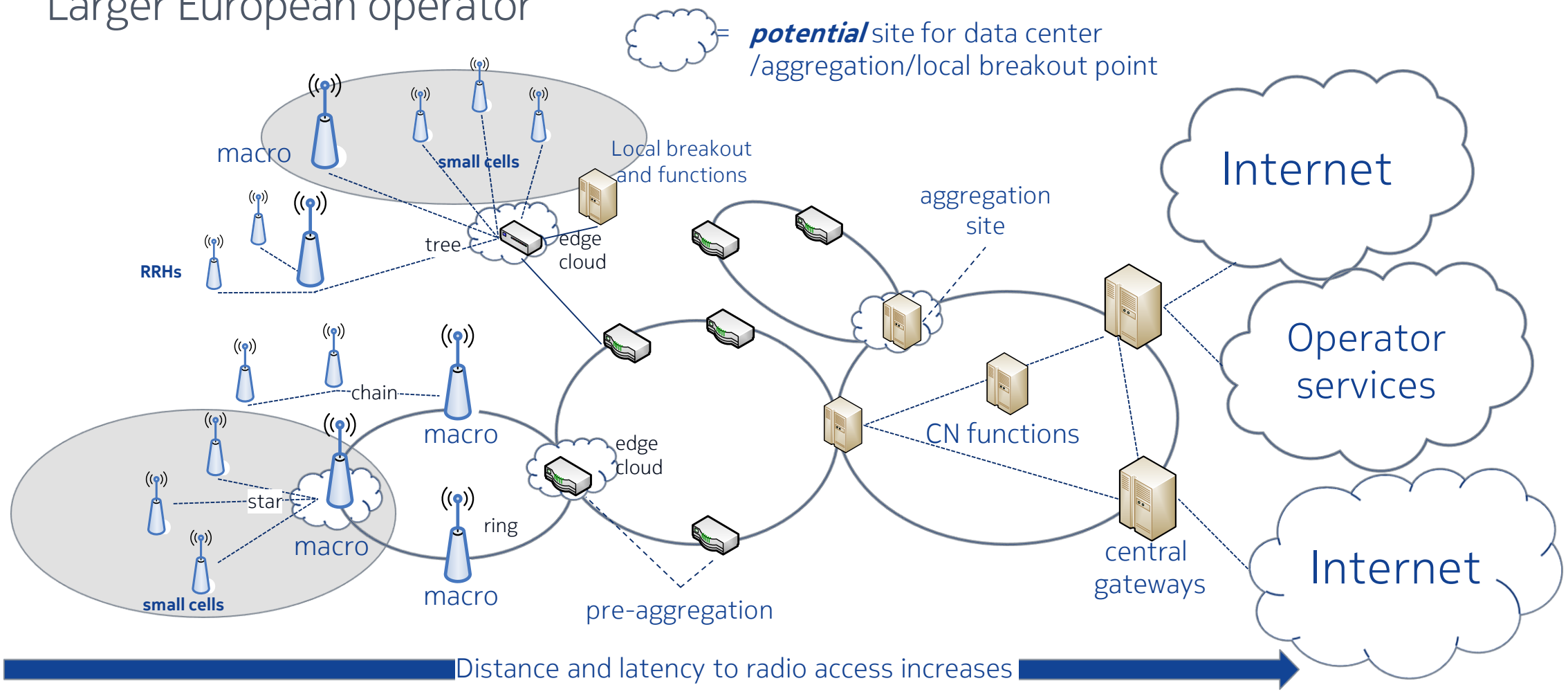


Diversity of use cases



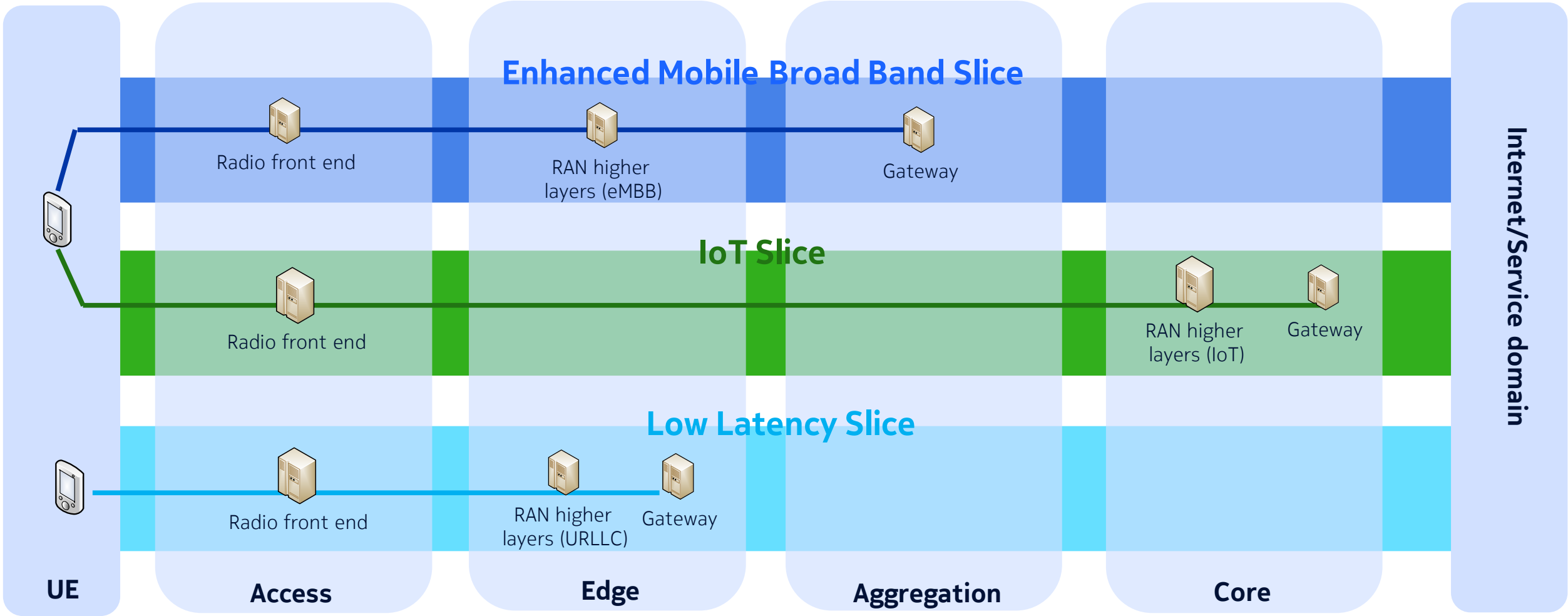
Mobile Network Topology
Larger European operator

➔Diverse x-haul, network topologies, aggregation scenarios



Network and RAN slicing

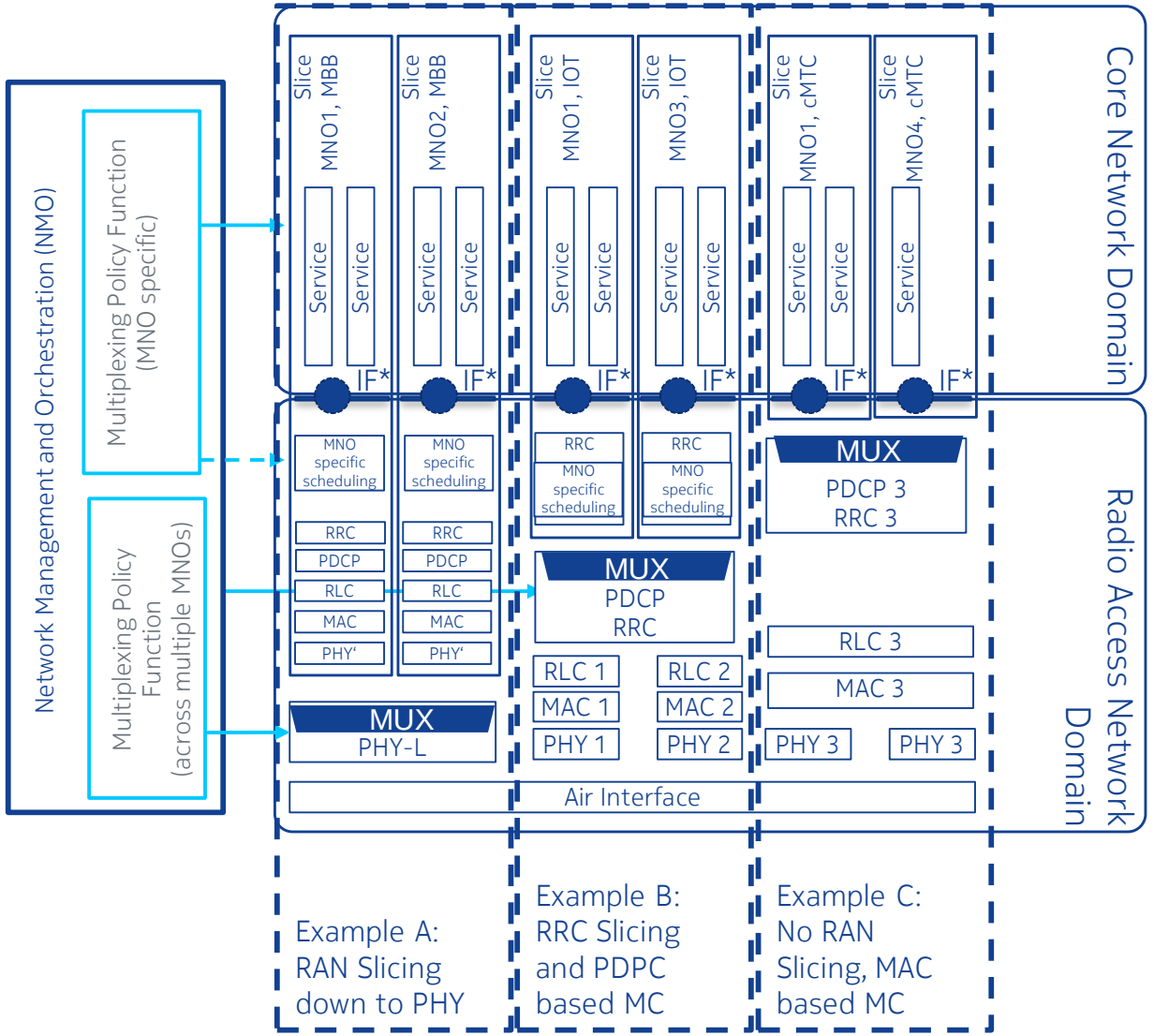
Flexibility to support different use cases in the same physical network



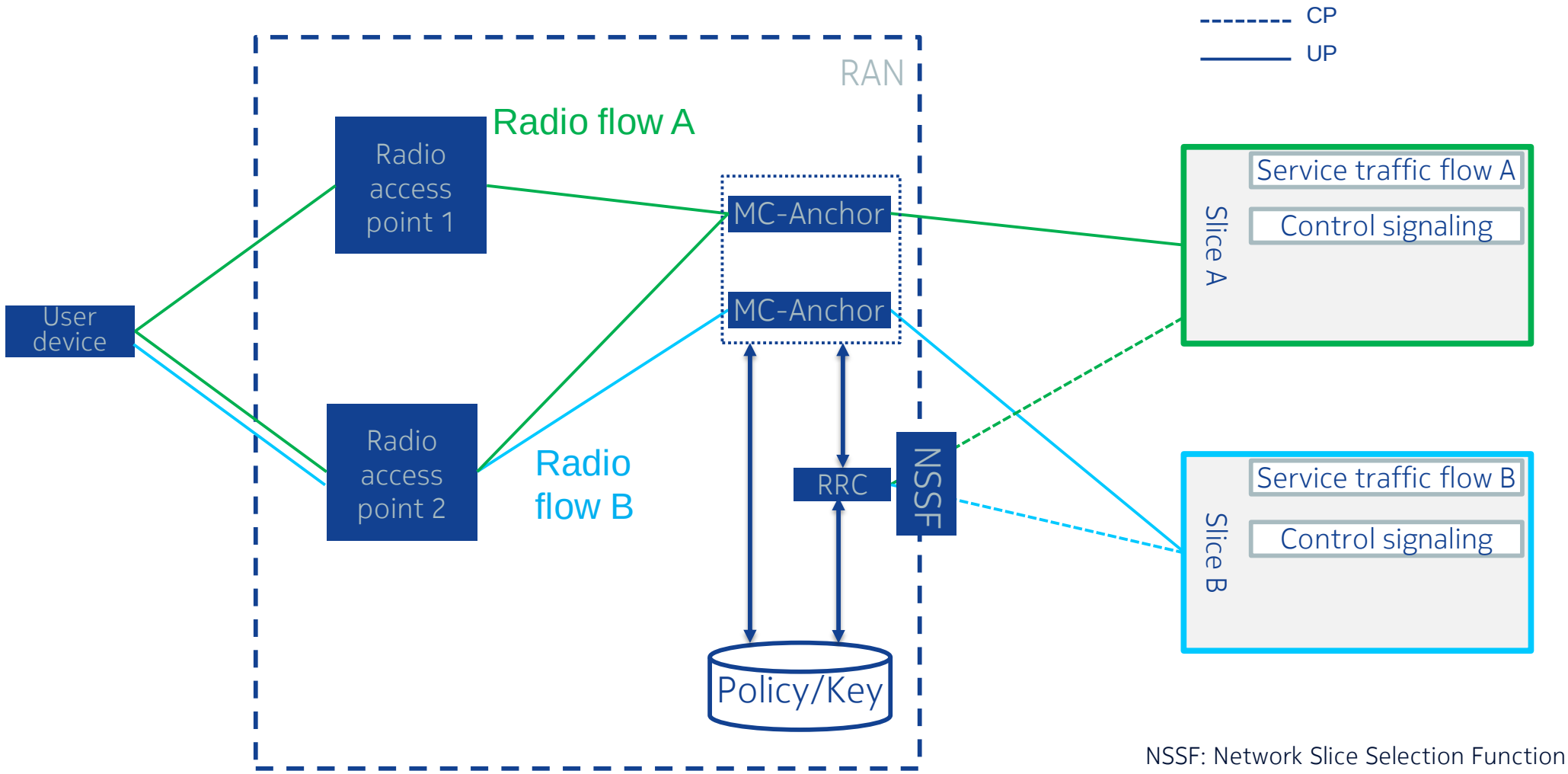
Isolation requirements

- Isolation of radio resources:
  - in order to fulfil slice-specific service level agreements (SLAs).
  - Function of the QoS framework and can be enforced by resource scheduling on MAC or on higher layers.
- Isolation of processing resources for network functions, e.g. if network functions are operating in a virtualized environment.
- Cryptographic isolation between slices, by providing cryptographic keys to the corresponding encryption function in RAN.

RAN slicing – basic concept

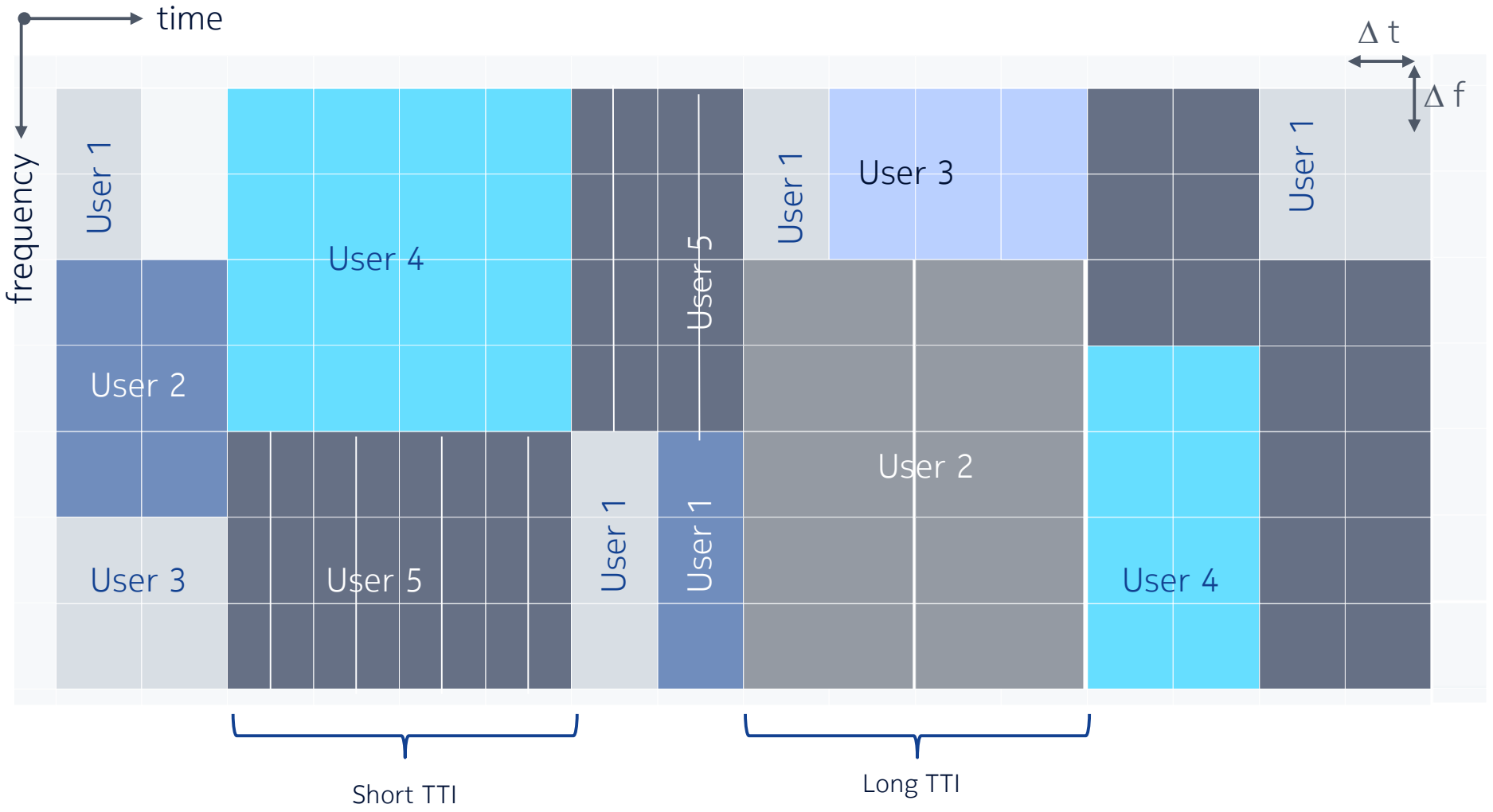


Basic architecture concept in 5G



NSSF: Network Slice Selection Function

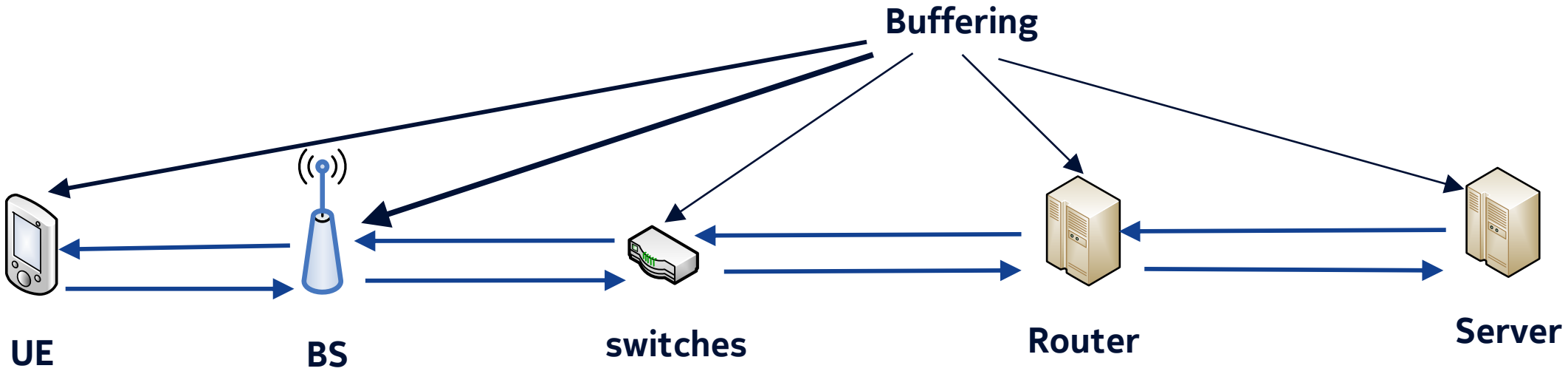
Mapping to lower layers
Exploiting air interface flexibility



Different TTI length per tile as enabler for slices with low-latency requirements

However, trade-off with cell capacity (multi-user diversity gains) need to be considered

Where does delay come from?



Air interface:
- Frame alignment/TTI
- Resource request (UL)
- RACH (UL)
- HARQ

On all nodes:
- Processing delay
- Packet buffering
- Multiplexing (e.g., TDM)

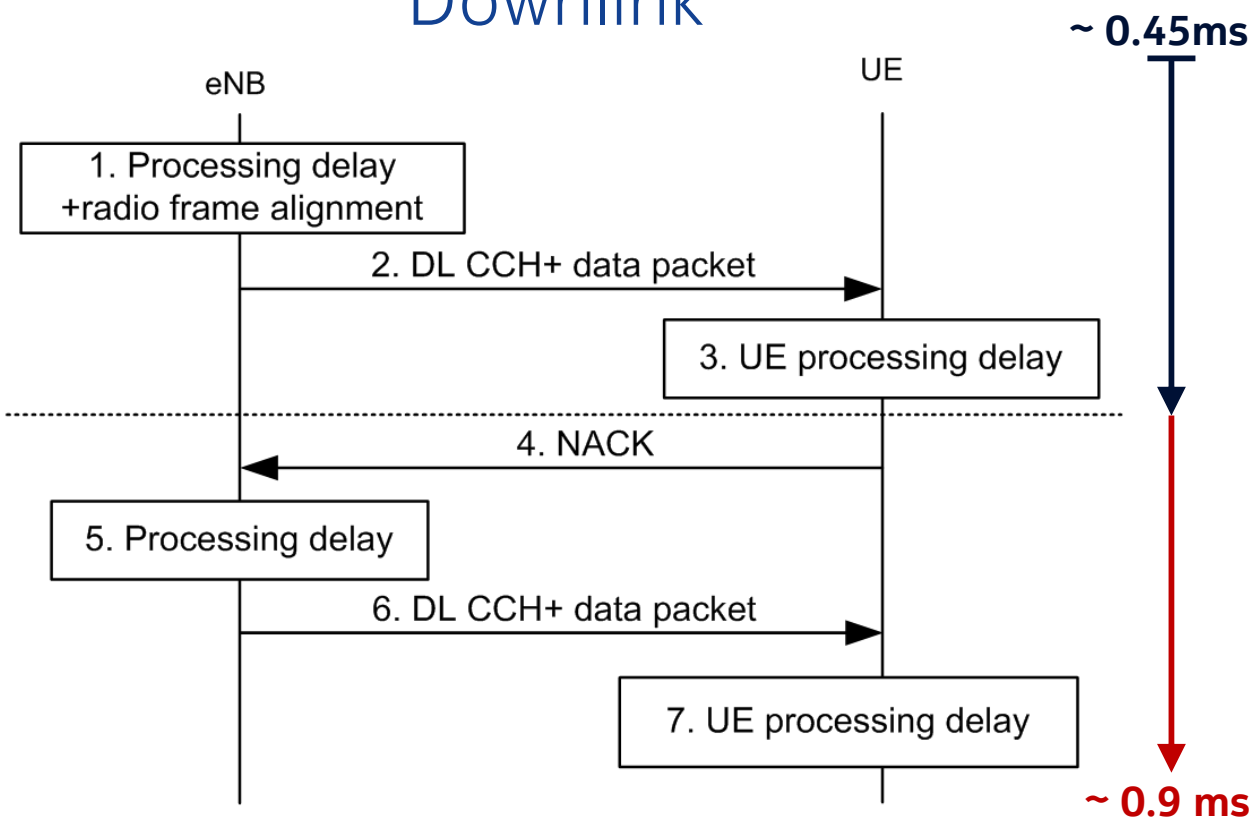
On all links:
- Transmission delay
- Duplex (e.g., TDD)

Air interface latency in 5G

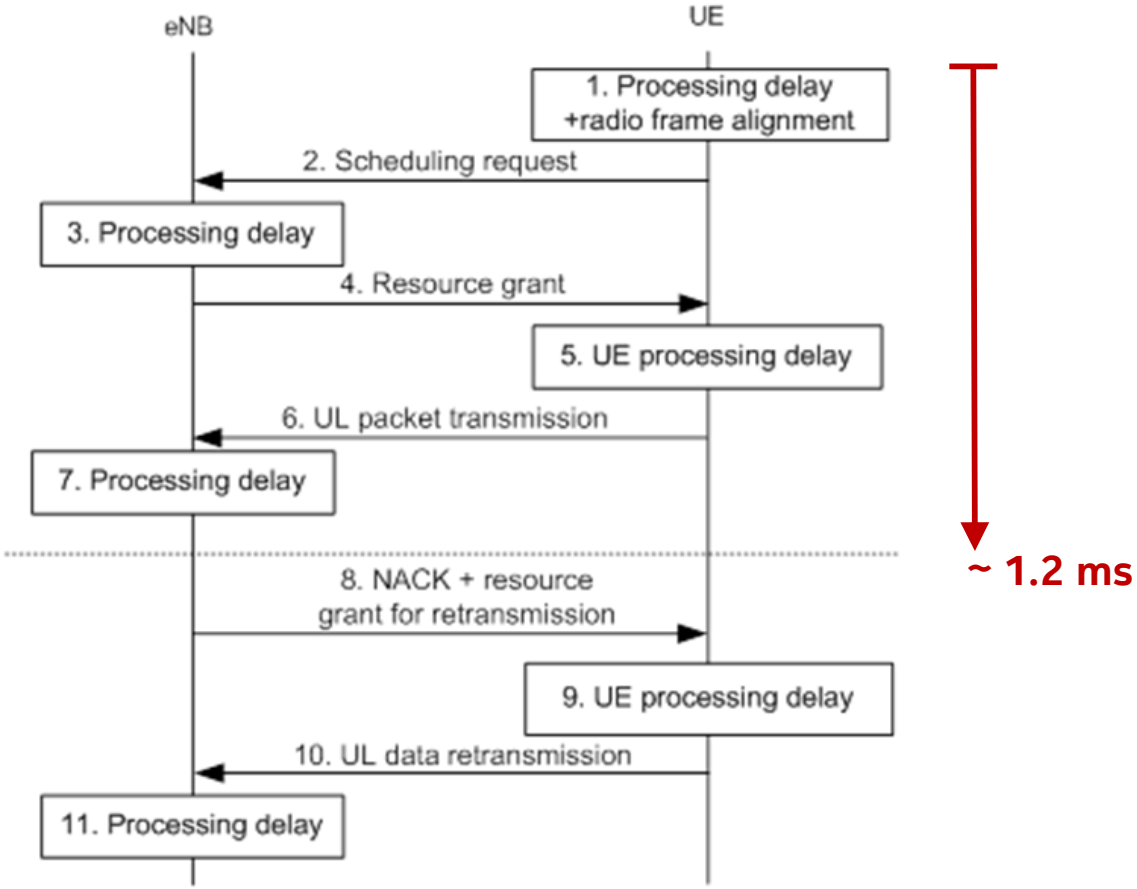
FDD, 0.125ms subframe length, 0.1ms processing in DL, 0.15ms in UL

Target: 0.5ms in UL and DL

Downlink



Uplink



Source: NOKIA Networks, “R1-165028; URLLC U-Plane Latency Analysis”

What to do in 5G?

Air interface:

- Shorter TTI/subframe (0.125 ms)
- Bi-directional subframes
- Shorter processing times (10x reduction)
- Robust coding/no or low #HARQ retransmissions
- Resource pre-allocation (e.g., semi-persistent scheduling – avoid resource request)
- Enhanced random access (e.g. 1 or 3 step instead of 5 step)
- Service-aware RRM
- Enhanced mobility mechanisms
- Device-to-device communication
- Multi-connectivity schemes

On all nodes:

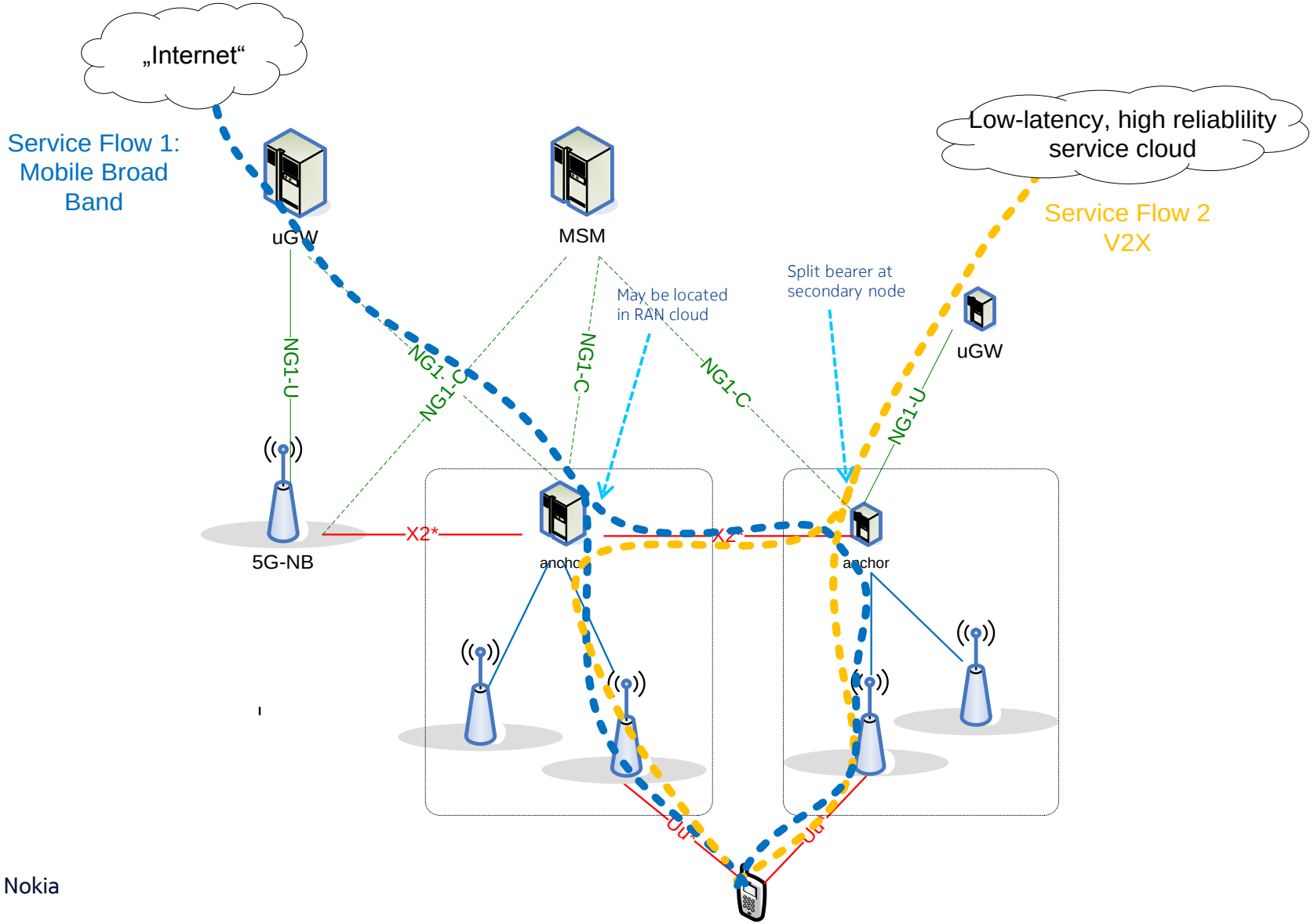
- Reduced processing delay (10x reduction)
- Priority queuing very small buffers
- Low number of hops to reduce instances of processing and buffering
- Path switching instead of packet switching

E2E architecture view

- Functions close at network edge
- Mobile edge computing optimized for low-latency
- Local area routing
- End-to-end QoS enforcement
- Resource isolation and/or dedicated infrastructure for low-latency
- Centralized resource coordination

Support for low-latency in unified architecture

Scenario: support of mobile broadband and ultra low latency



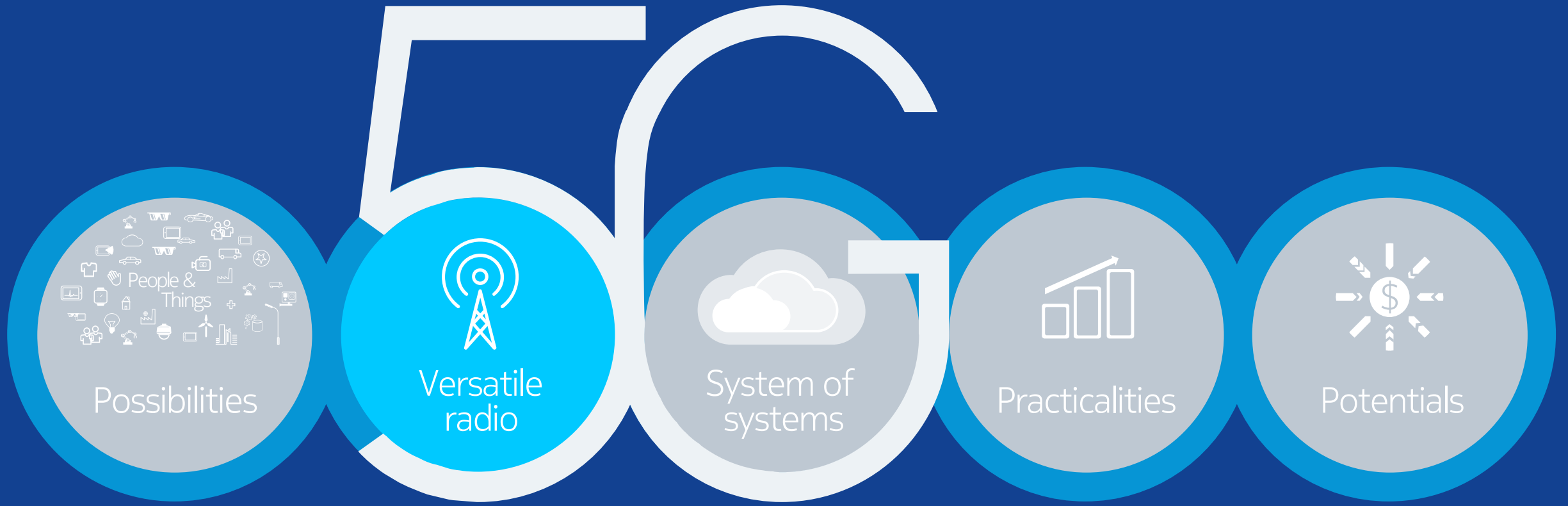
Summary and Outlook

Network and RAN slicing are key concepts for 5G

- Enables flexibility for different use cases in the same network
- Reduces management and maintenance efforts

Low latency services – enabled by function optimization and deployment flexibility

- Configuration and optimization across all RAN functions
- New resource management schemes needed
- Flexible architecture concepts to support different use cases



Key to the programmable world

RAN slicing as enabler for low latency service

NOKIA

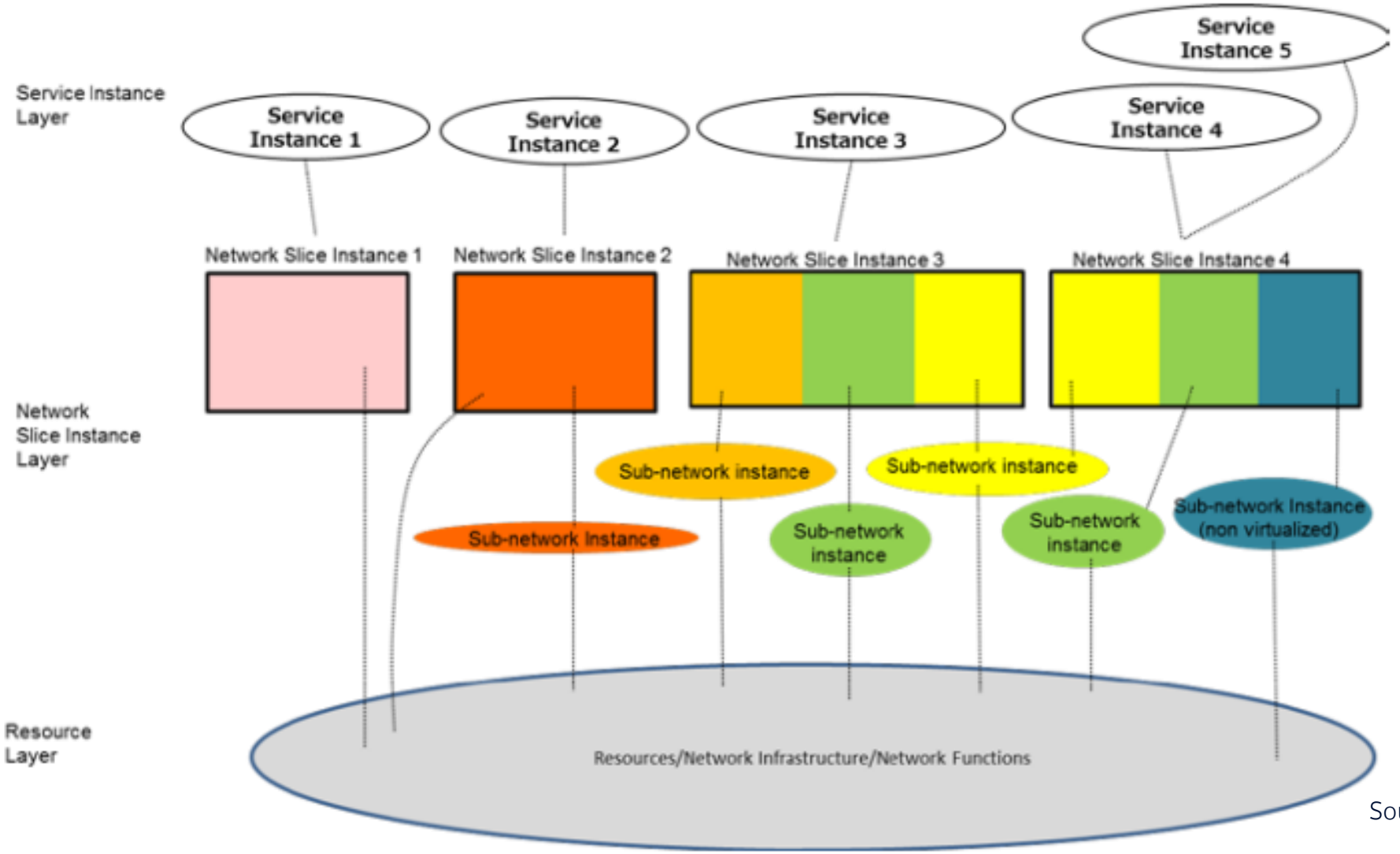
References

[1] NGMN Alliance, “Description of Network Slicing Concept”, January 2016

[2] NOKIA Networks, “R1-165028; URLLC U-Plane Latency Analysis”, 3GPP RAN1#85, May 2016

[3] 3GPP, “TR 38.913 V.0.3.0; Study on Scenarios and Requirements for Next Generation Access Technologies”, May 2016

Definitions



Source: NGMN


NOKIA