Real-time wireless connectivity challenges in Open RAN

Blas Romero-Garcia, system developer @Ericsson

Agenda

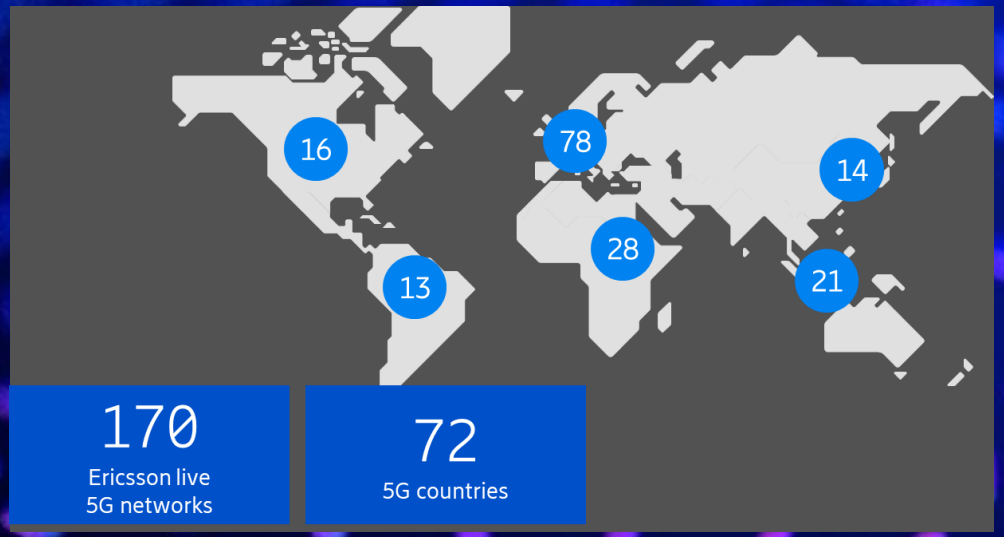
- Background: R&D and 5G @ Ericsson
- Understanding the complexity of a mobile communication system
- Open RAN: fundamentals and challenges

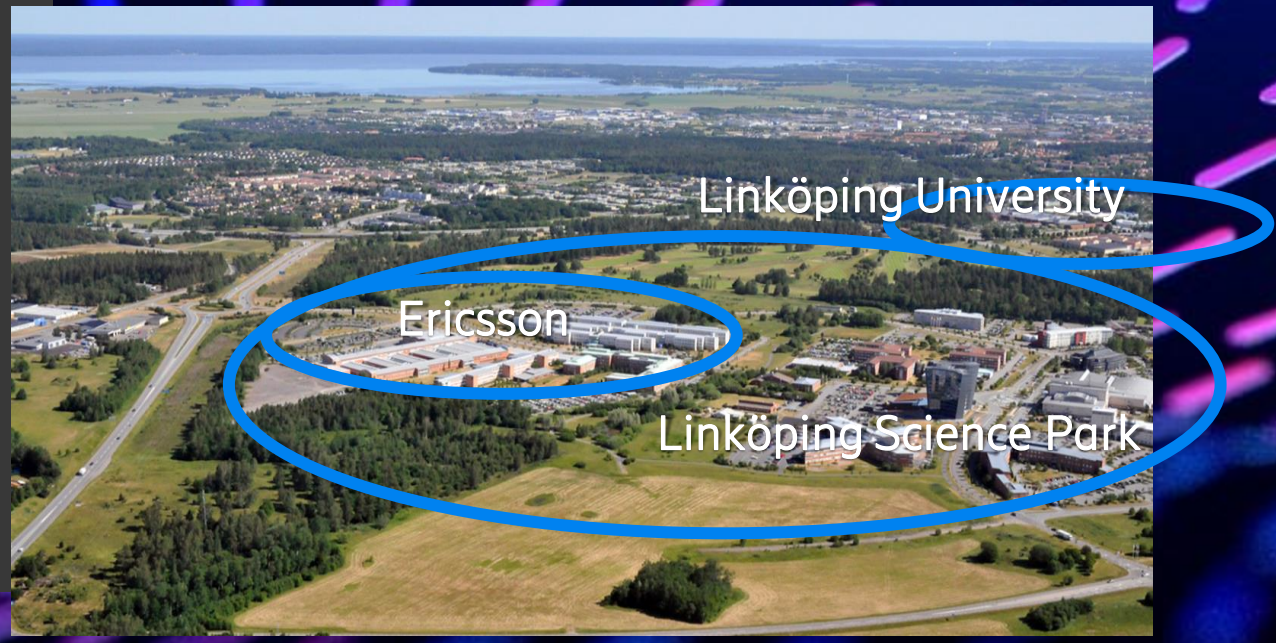
5G in scale & Ericsson R&D

- Many of our external interfaces are controlled by international standardization organizations (3GPP)
  - In these organizations we are represented, but so are most of our competitors and customers.

- Ericsson networks: 170 live 5G networks in 72 countries
- 150,000 employees worldwide

- Linköping site
  - 5G product development, Research and a big lab
  - Aprox. 1000 employees in total
  - +200 patents yearly

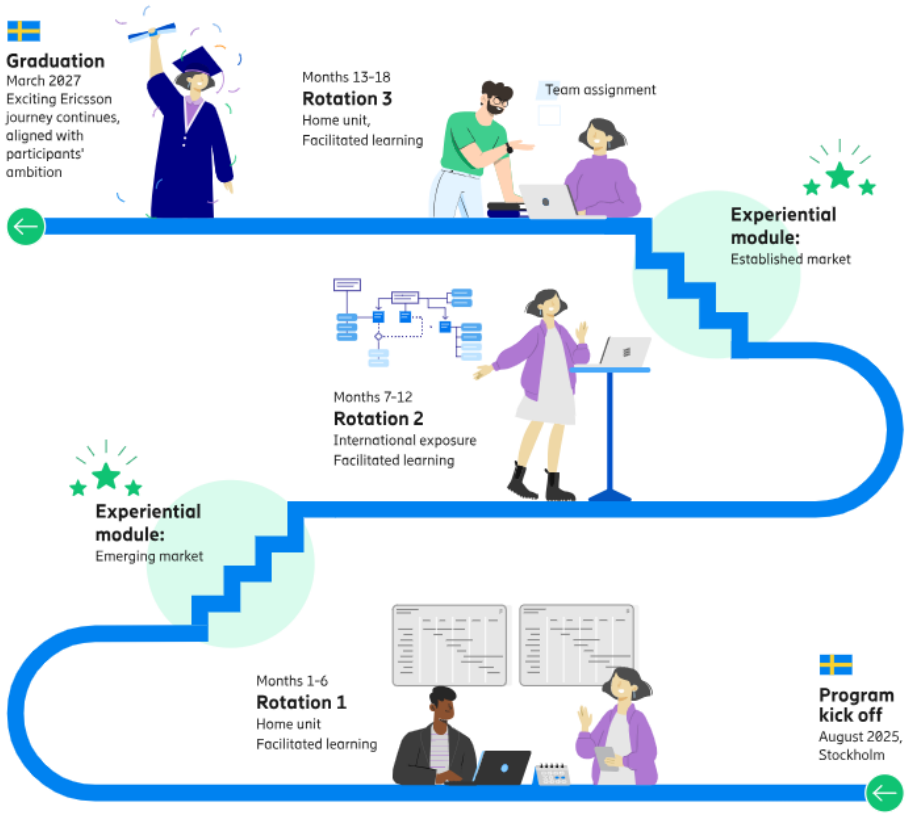




Looking for a candidate to Early Career Program

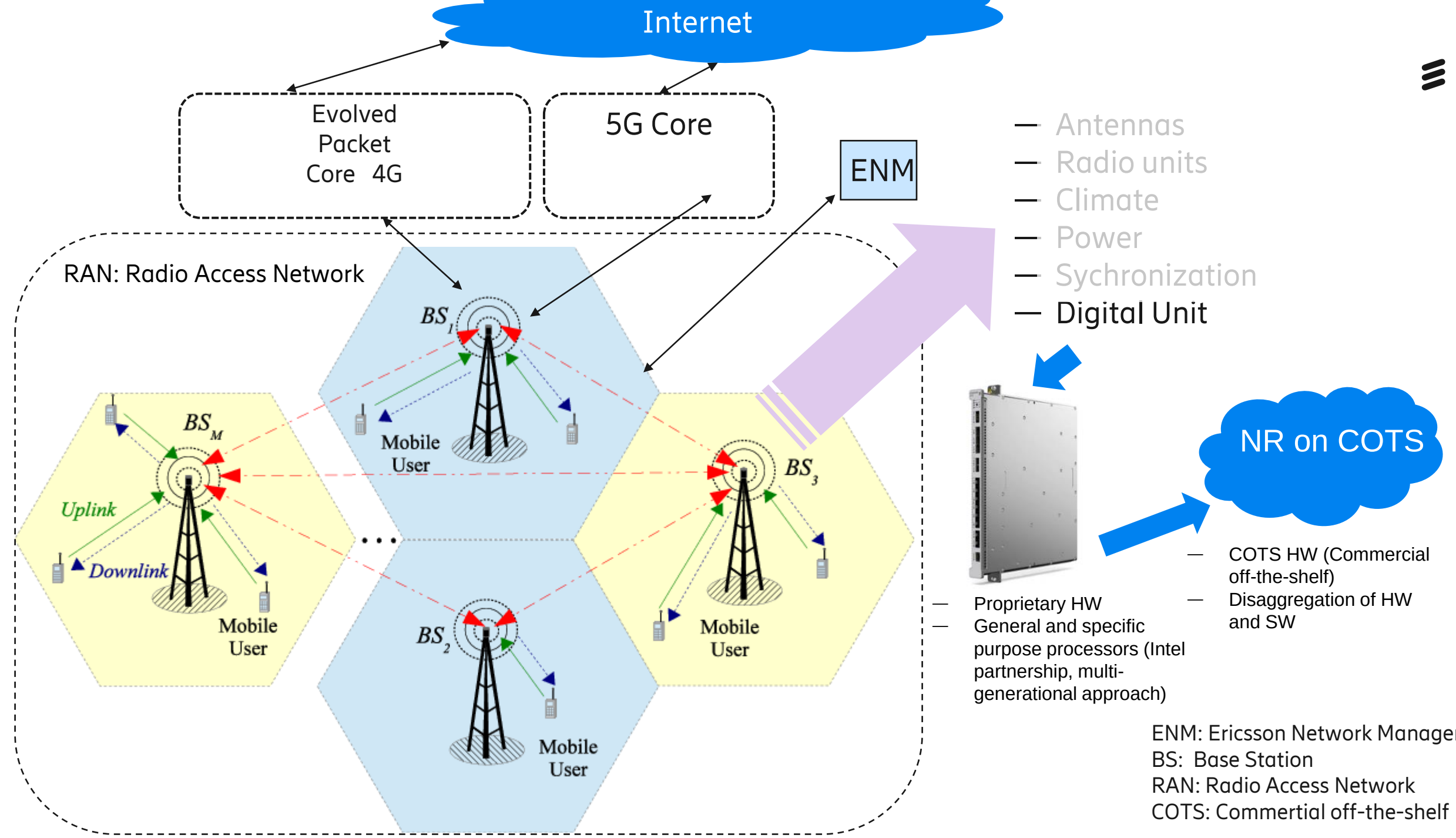
- Our Early Career Program (ECP) is designed to bring out the best in you and help you achieve your Imagine Possible. The 18-month global programs are intended for recent graduates or professionals early in their career, encouraging and empowering the leaders in them.

- Careers at Ericsson
- ECP position in Linköping
- https://www.ericsson.com/en/careers/student-young-professionals/early-career-program



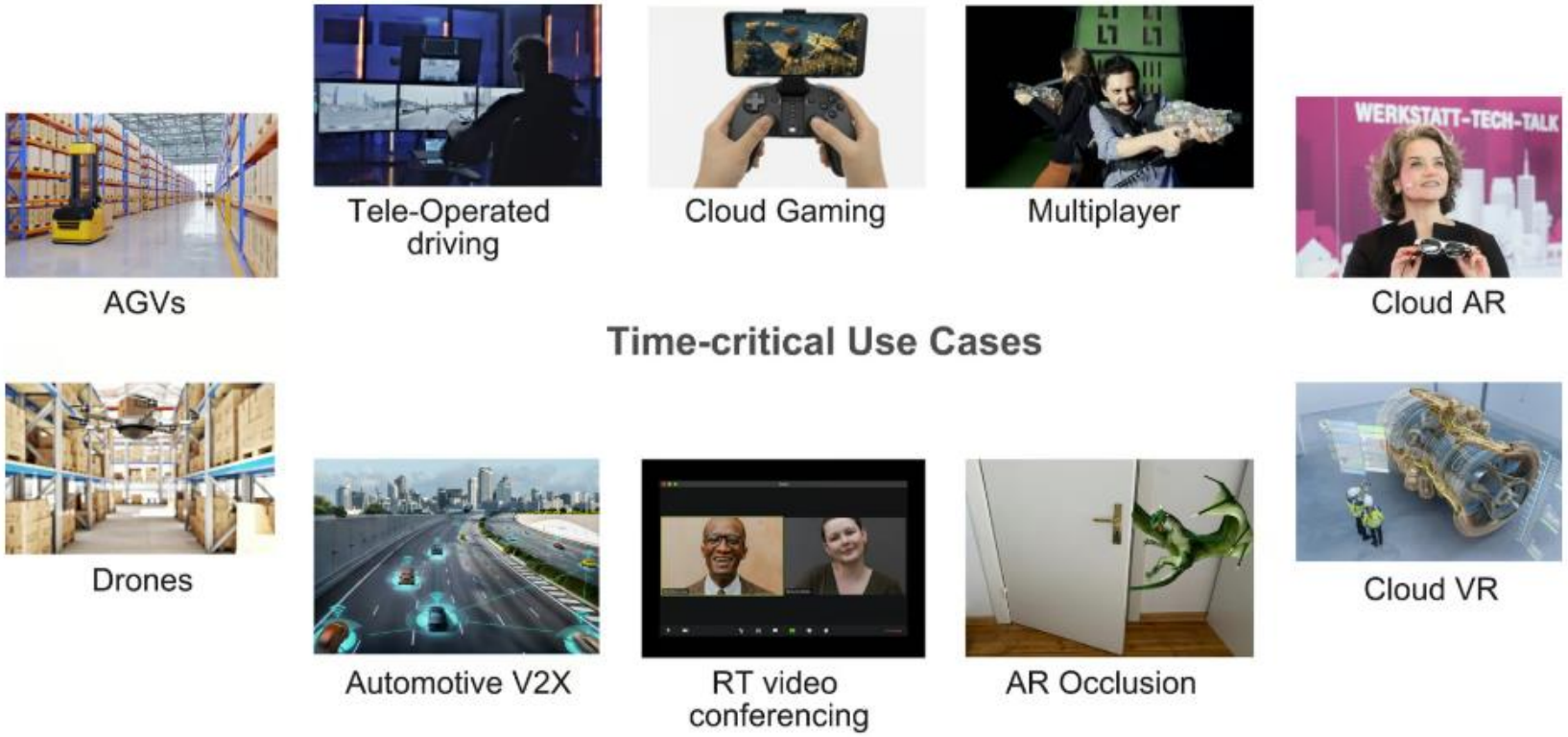
18 months of accelerated learning: Functional deep dives, mentorship from experts and access to global alumni network.

5G network: fundamentals and time-critical use cases


Internet
Evolved Packet Core 4G
5G Core
ENM
RAN: Radio Access Network
BS1
BSM
BS2
BS3
Mobile User
Mobile User
Mobile User
Mobile User
Uplink
Downlink
Antennas
Radio units
Climate
Power
Sychronization
Digital Unit
Proprietary HW
General and specific purpose processors (Intel partnership, multi-generational approach)
NR on COTS
COTS HW (Commercial off-the-shelf)
Disaggregation of HW and SW
ENM: Ericsson Network Manager
BS: Base Station
RAN: Radio Access Network
COTS: Commertial off-the-shelf

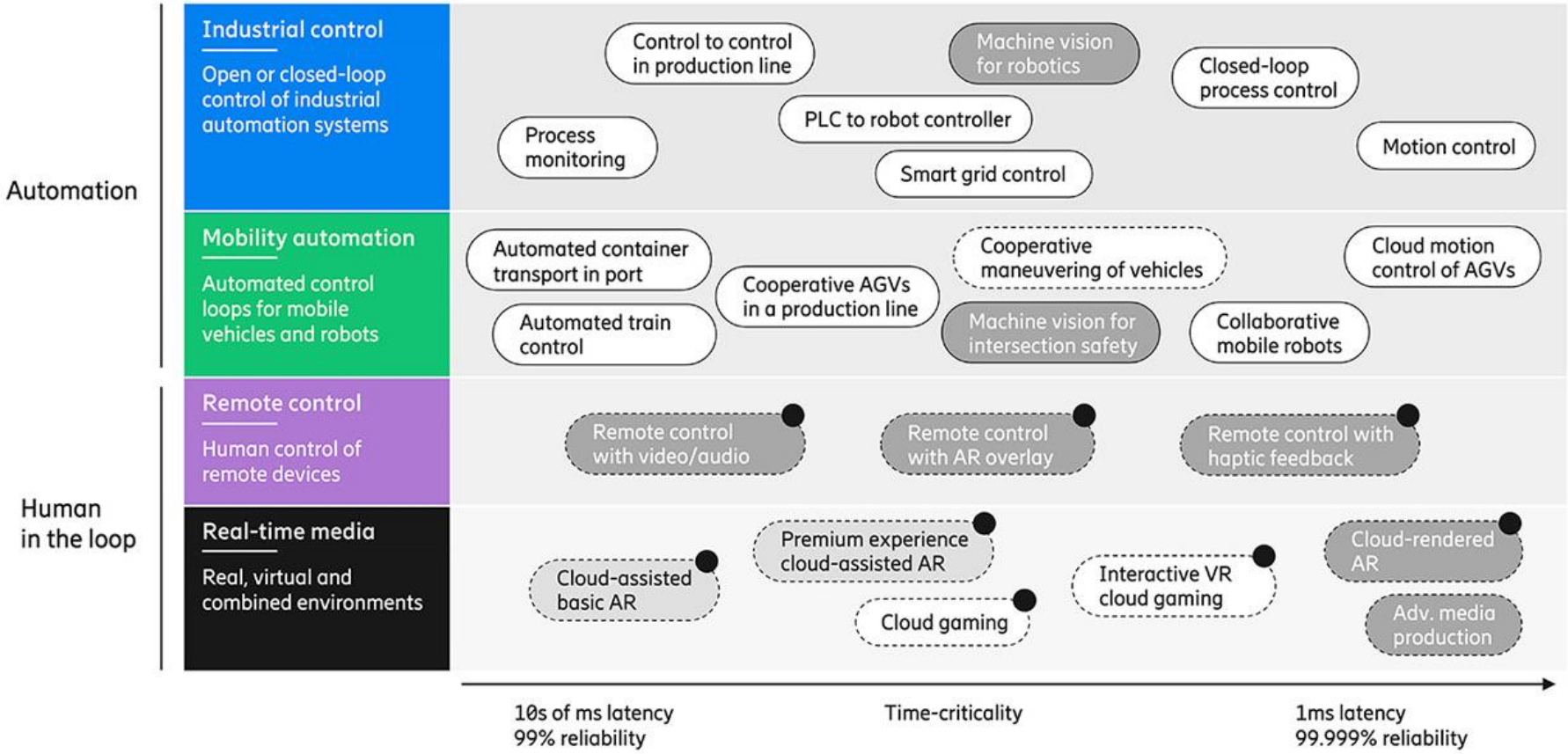
5G: Time-critical use cases

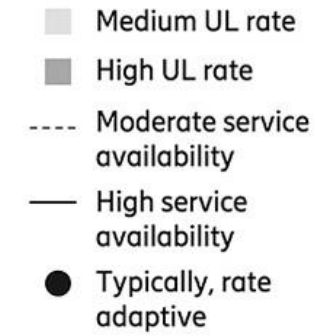
— 4 time-critical use case families: real-time media, remote control, industrial control, mobility automation

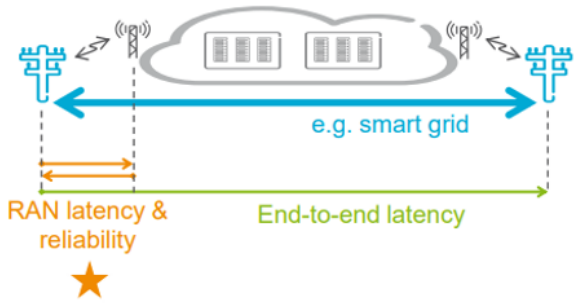


Source: Deutsche Telekom (see detailed image sources: Enabling time-critical applications over 5G – Ericsson)

5G applications: fast or critical?







3GPP/ITU target on RAN latency and reliability.

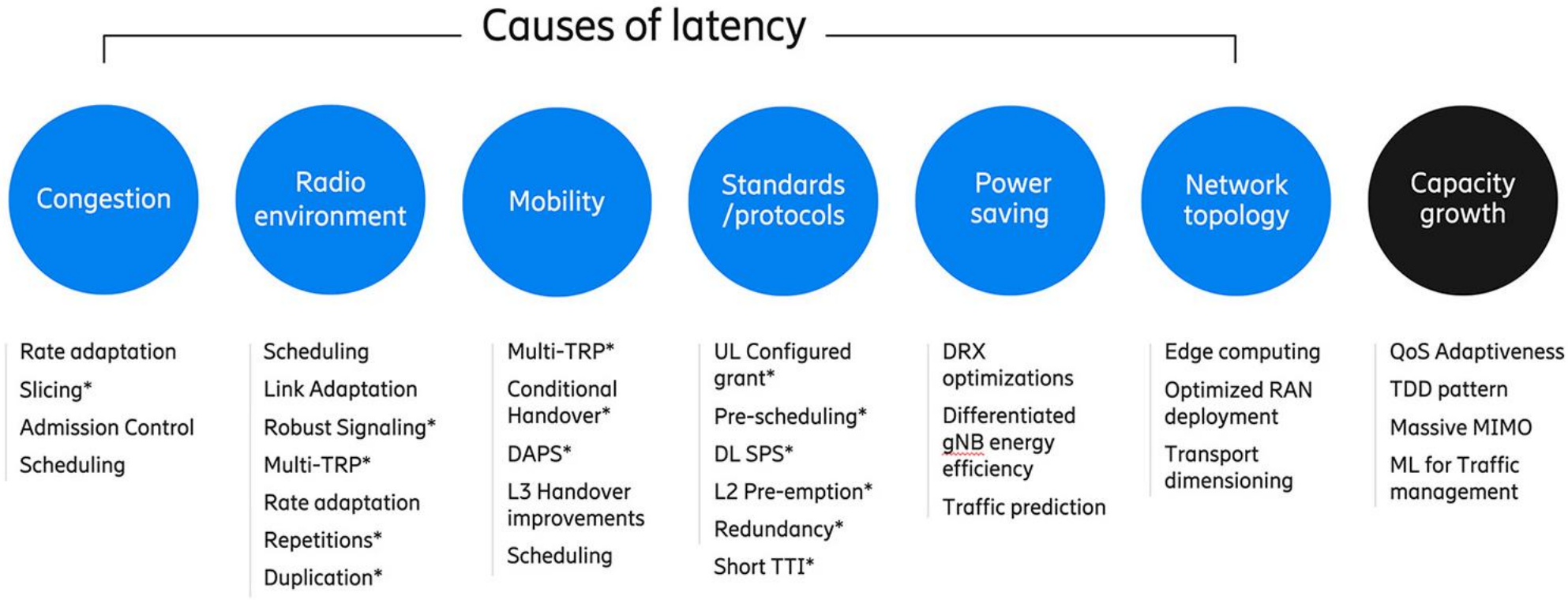
Traditional mobile broadband: High peak rates & best-effort low latency

≠

Real-time critical: High reliability & consistent low latency

(URLLC: Ultra-reliable and Low Latency Communication)

Technical challenges preventing time-critical applications

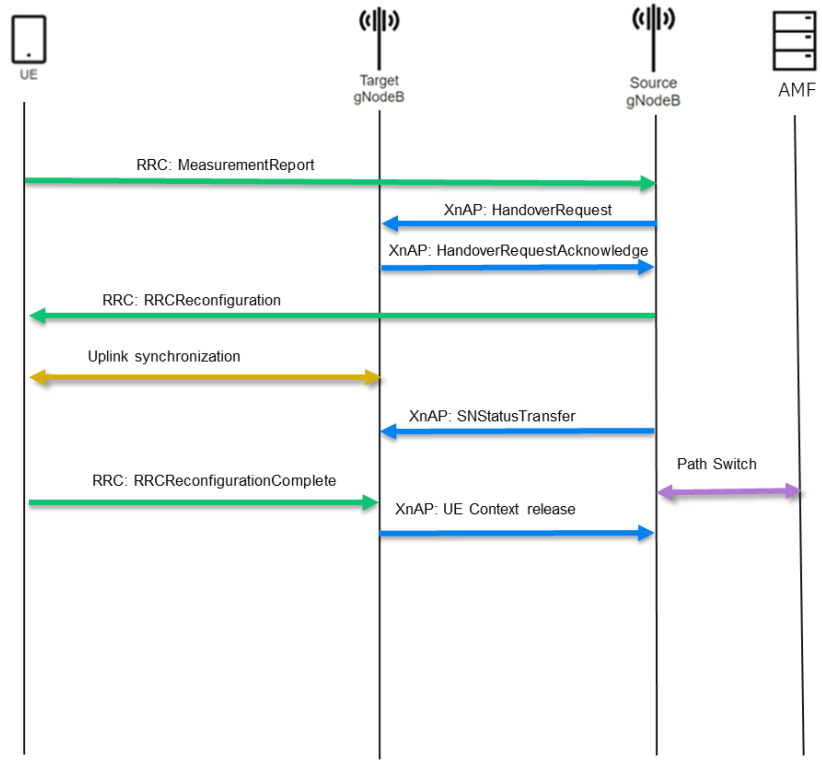


Timing aspects in 5G: Mobility

- Communication between end user equipment (smartphone) and base stations regulated by standard protocols (3GPP)

- Delays in the base station can cause:
  - End user experience degradation (jitter and disconnections)
  - Accessibility issues (signals not reaching the UE on time, timing out)

- Processing resources shared in the computing nodes among all the connected users: SW dimensioned to support thousands of requests per second



UE: User Equipment
XnAP: Protocol between two gNodeBs
gNodeB: 5G node
AMF: Access Management Function
RRC: Radio Resource Controller

Open RAN: fundamentals and challenges

What is Open RAN?

— Transforming Radio Access Networks towards open, intelligent, virtualized and fully interoperable RAN

O-RAN key elements

NR (New Radio, 5G radio access technology) on COTS (Commercial off-the-shelf)

Open interfaces

Standardized interfaces (O-RAN) combined with 3GPP-defined interfaces to facilitate interoperability

Disaggregation of hardware and software

RAN applications software running on general-purpose hardware.

Intelligent open management

Automated management systems with capabilities for AI and ML for efficient lifecycle management of network functions

Why O-RAN?

- Vendor independence
- Simplified network management
- Flexibility and scalability

How do customers perceive quality in our products?

Quick delivery of new features	KPI Key Performance Indicators	Flexibility/escalability	Security

Availability

The quality of being able to be used (99,999% system up)

- Specially important for critical services
- Crashes: few and short

Accessibility

The quality of being able to be reached or entered

- Challenge in Mobile systems: Number of users accessing the network not constant

Latency

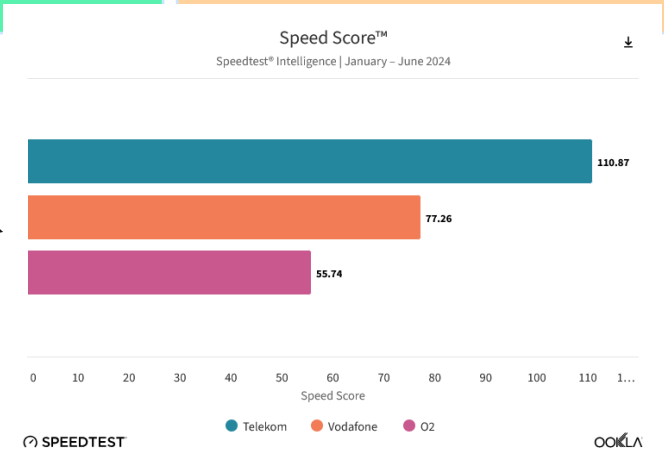
Round-trip-time, time to get the response from the remote end. How much time does it take for an user to get access to the system?

Speed

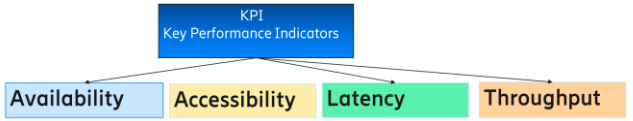
How many bits/sec can the UE send/transmit?

— **Benchmarking between vendors:**

- Traditionally performed on the user equipment (mobile) side
- With O-RAN it will be possible to benchmark the software
  - Same HW, SW from different vendors
  - Energy consumption? Capacity levels?



NR on COTS challenges: Availability



- Current availability: 99,999% (5 minutes downtime per year)
- Future applications might require even higher availability

How to improve availability?

Traditional approach in purpose-build HW

COTS: Commercial off-the-shelf

purpose-build HW

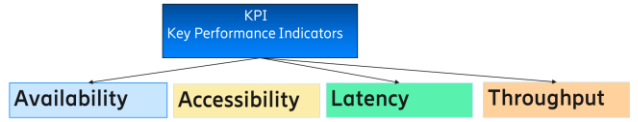
- High-availability underlying HW, better than 99,999% (infrastructure)
- Improve **SW reliability** (fewer SW errors leading to node crash)
- **Shorter restart times** (if the node crashes, service can be reestablished quickly)

general purpose (COTS)

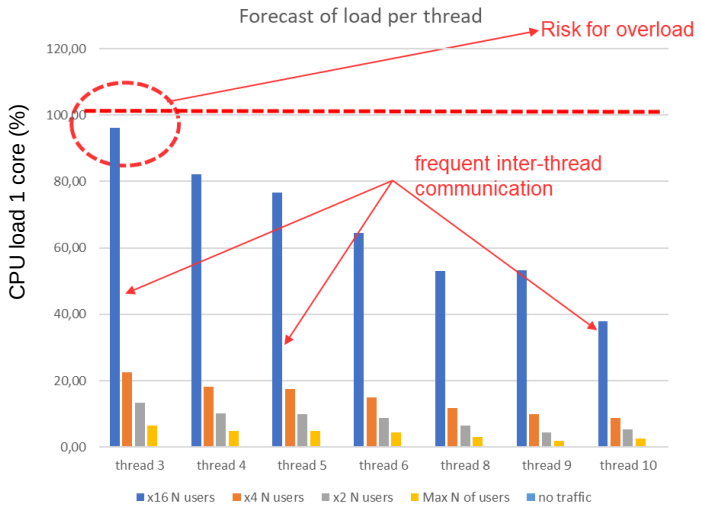
- Other applications running in the server (in addition to 3PP customer applications and CaaS layer) layer that need to be upgraded and could affect availability
- If there is only one server, it is difficult to reach better availability than in purpose-build HW

https://www.ericsson.com/en/news/2023/2/ericsson-dt-and-vay-demo-5g-teledriving

NR on COTS challenges: Accessibility



- System dimensioned for a maximum number of users
- Under certain situations, the number of users can reach the maximum. Traffic will be rejected to prevent CPU overload (which would cause longer latency/ timeouts)
  - Way forward:



Traditional approach in purpose-build HW

How to improve availability?

Commercial off-the-shelf

purpose-build HW

- **Optimizations** to reduce load
- Introduce **multi-threading** (threads handling users in parallel, running in different cores) **(Drawback**: increase in memory, also limited in embedded systems)

general purpose (COTS)

- HW not designed for best performance with the specific SW, customized configuration may be required
- Core isolation is needed (processes with different latency requirements competing for the cores in the server)

Summary

- A solution in which the software of the radio access network can be deployed on COTS using a single server is very interesting for some customers
  - Achieve vendor independence
  - Facilitate introduction of new vendors

- A COTS solution is also expected to simplify network management (same hardware independently of the vendor)

- Different parts of the application with different real-time response requirements can require core isolation in the COTS server.

- Performance of specific purpose processors is expected to be better than performance on COTS (HW designed for the software) and higher capacity per core.

Questions?