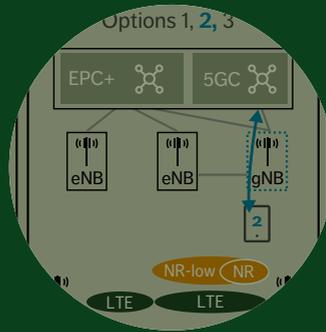


Review

ERICSSON
TECHNOLOGY



5G ECOSYSTEM: ARCHITECTURE OPTIONS



ERICSSON

Simplifying the 5G ecosystem.

BY REDUCING ARCHITECTURE OPTIONS

Previous mobile generations have taught us that industry efforts to reduce fragmentation yield massive benefits. In the case of 5G, an industry effort to focus deployment on a limited set of key connectivity options will be critical to bringing it to market in a timely and cost-efficient way.

**TORBJÖRN CAGENIUS,
ANDERS RYDE,
JARI VIKBERG,
PER WILLARS**

The multiple connectivity options in the 3GPP architecture for 5G have created several possible deployment alternatives. Initial deployments focus on options 3 (non-standalone New Radio) and 2 (standalone New Radio). However, the deployment of several additional options would create a level of complexity that impacts the whole 5G ecosystem – across operator network operations, equipment vendors and user equipment (UE) chipset vendors as well as spectrum assets. To avoid ecosystem fragmentation, we believe that the best approach is to limit the number of options that are deployed.

■ There is much more to introducing 5G than simply deploying New Radio (NR) technology. For a successful 5G launch, the operator needs to secure a

network that includes end-to-end (E2E) capabilities aligned across devices, RAN, core and management systems. 5G is also a technology transformation for operators striving for more flexibility and speed in network deployment – and with an expectation of being able to address new business opportunities with use cases beyond mobile broadband (MBB). One of the key strategic topics that operators need to decide on is which connectivity options to support in the network to address the targeted use cases.

5G connectivity options

In Release 15, the 3GPP [1] has defined multiple architectural options for a UE to connect to the network, using LTE/eLTE and/or NR access to connect to Evolved Packet Core (EPC) or 5G Core (5GC) networks. A new use of dual connectivity has also been applied to use LTE/eLTE and NR as the master or secondary radio access technology (RAT)

Connectivity option	Core network	Master RAT	Secondary RAT	3GPP term	3GPP release
Option 1	EPC	LTE	-	LTE	Rel. 8
Option 3	EPC	LTE	NR	EN-DC	Rel. 15, Dec 2017
Option 2	5GC	NR	-	NR	Rel. 15, June 2018
Option 4	5GC	NR	eLTE	NE-DC	Rel. 15, March 2019
Option 5	5GC	eLTE	-	eLTE	Rel. 15, June 2018
Option 7	5GC	eLTE	NR	NGEN-DC	Rel. 15, March 2019

Figure 1 UE connectivity options

in different combinations. This has resulted in six connectivity options for a UE, as shown in *Figure 1*. Note that while the option terminology is not explicitly used in the 3GPP standards specifications, it originates from the 5G study phase of 3GPP Release 15 and is widely used in the industry.

The six connectivity options shown in *Figure 1* define how any single UE is connected to the network at a given time. In most cases, a network will support a set of such options simultaneously. One base station may have different UEs connected via different connectivity options, as well as moving a UE connection between the options depending on factors such as radio conditions. Legacy LTE/EPC (option 1) is the baseline, and the industry has an aligned view that the initial 5G deployments are based on options 3 and 2. The next step, therefore, is to establish industry alignment on the potential use of options 4, 5 and 7.

The need for industry alignment

Mobile network operators that deploy 5G must be able to support UE, radio network, core network and management products that are manufactured by a multitude of device and network equipment vendors. With multiple connectivity options, and even more possible combinations of options, there is a high risk that different operators will deploy different options, in a different order. If that happens, chipset, device and network equipment vendors are likely to get contradictory requirements from different operators or markets. This would cause significant product and integration complexity, as well as creating interoperability issues that prolong the time it takes to establish a complete ecosystem that supports the deployed options.

The complexity caused by a multitude of deployed connectivity options would also have an impact on the E2E testing of services in the operator network,

including both existing services like voice as well as new ones. Further, the higher the number of options deployed, the more complex and time consuming it will be for the operator community to establish 5G roaming in the industry.

Network deployments based on options 3 and 2

Option 3 is the best short-term alternative for 5G deployment, as it relies on existing LTE/EPC (option 1). Option 3 will provide good performance in several aspects, allowing optimized transmission on NR when NR coverage is good, extending NR downlink (DL) usage on a higher band by combining with a lower-band LTE for uplink (UL) data, and, if needed, aggregating throughput over both NR and LTE spectrum. It also provides reliable and smooth mobility based on anchoring in LTE/EPC, even if the NR coverage is spotty. The use of dual connectivity has, however, introduced some challenges on the UE side with dual transmitters, which, in some

cases, will limit performance and coverage.

One of the main drivers for going beyond option 3 is to provide 5GC-enabled capabilities like enhanced network slicing, edge computing support and operational benefits, even though EPC can also support these services to some extent (slicing based on DECOR, for example). Another main driver for going beyond option 3 is to be able to deploy standalone NR and get the radio performance benefits of an NR-only based radio interface. Option 2 (standalone NR) is the first 5GC-based option available in UEs and networks.

Even if general NR coverage is limited, option 2 can initially be deployed for specific use cases in local areas, where devices stay within good NR coverage on a mid or high band. Examples include industrial deployments with ultra-reliable low latency communication requirements, and fixed wireless access (FWA), even if the latter is also well served via option 3.

Key enablers

» LTE-NR spectrum sharing

3GPP specifications allow efficient sharing of operator spectrum, so that one carrier appears as an NR carrier to NR UEs, and an LTE carrier to LTE UEs. Resources are pooled and distributed dynamically between the two RATs, according to instant needs. There is no impact on legacy LTE UEs, and the impact on LTE capacity is very small. Compared with classic refarming, this provides a smooth migration of spectrum from LTE to NR as NR-capable UE penetration increases, enabling NR to be rolled out on new and legacy bands.

» Spectrum regulation

Spectrum is becoming technology neutral in most of the world except for a few markets and frequency bands where the spectrum license is currently tied to a specific RAT, prohibiting NR to operate in existing frequency bands. It is important that regulators acknowledge the need for NR deployment in all bands. This is a key enabler for migration to wide area coverage of services like MBB/voice and cMTC over 5G, depending on the possibility to deploy NR in lower frequency bands.

» Dual-mode core network

4G devices will be the major device type and traffic consumer for a long time [2]. In addition, operators are introducing new 5G devices depending on both EPC (option 3) and 5GC (option 2). A “dual-mode” core network with both EPC and 5GC functionality will support the evolving device fleet in the network and enable a smooth network transformation. To ensure service coverage during the migration period, the dual-mode core network will provide tight interworking between EPC and 5GC for seamless 4G-5G mobility.

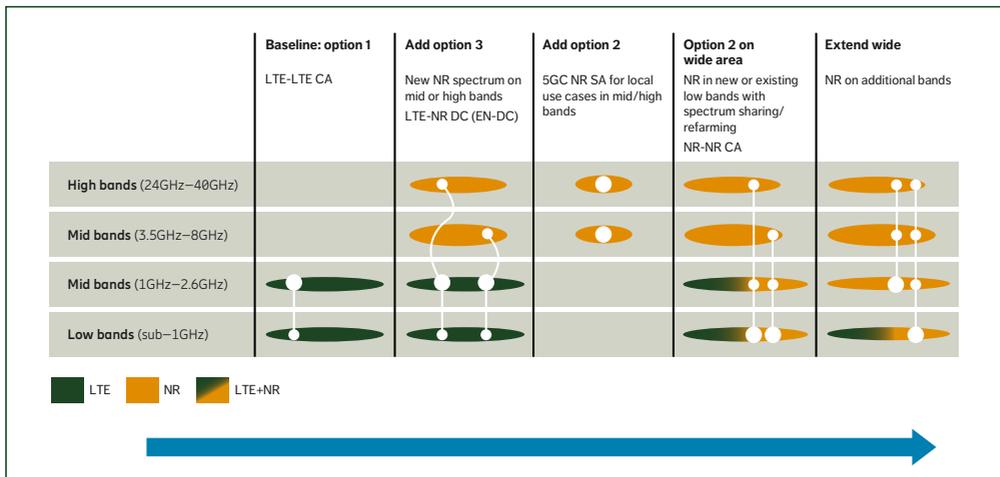


Figure 2 Spectrum migration steps for the 5G network

Figure 2 illustrates the evolution of spectrum usage in a network, starting with LTE deployed on sub-1GHz and 1–3GHz bands. First, NR is deployed on 3.5GHz and/or mmW and with LTE bands using option 3. The next step is to deploy option 2 for specific use cases in local areas – such as for FWA and industrial deployments.

Expanding standalone NR coverage and capacity

When deploying option 2 for wide-area use cases like MBB, it is important to ensure continuous NR coverage within the targeted area (initially urban for example). Spotty NR coverage would result in frequent mobility events between NR and LTE for wide-area use cases, even though intersystem mobility between option 2 and LTE/EPC will be well supported. For these use cases, option 2 requires a sufficiently low NR band in relation to the site grid. In many cases, the site grid for a 3.5GHz deployment will give good DL coverage both outdoors and indoors, but not enough UL coverage. NR on 3.5GHz should therefore typically be combined with NR on low

band to provide continuous coverage in both the UL and DL [3]. The low NR band can be new, reformed or an existing LTE band that is shared between NR and LTE. With reformatting or sharing, a key enabler is that the spectrum license allows NR deployment (see fact box on page 4, spectrum regulation).

To support option 2 for MBB in an area, it is also advisable to deploy NR in one or more legacy LTE bands using LTE-NR spectrum sharing (see fact box on page 4, LTE-NR spectrum sharing). Together with NR on low and mid/high bands, this maximizes the throughput via NR carrier aggregation (CA). This is essential to provide good MBB performance, especially in areas without DL coverage from new NR bands. While NR deployment is limited, mobility to option 2 should only be triggered when the UE

WITH REFORMING OR SHARING, A KEY ENABLER IS THAT THE SPECTRUM LICENSE ALLOWS NR DEPLOYMENT

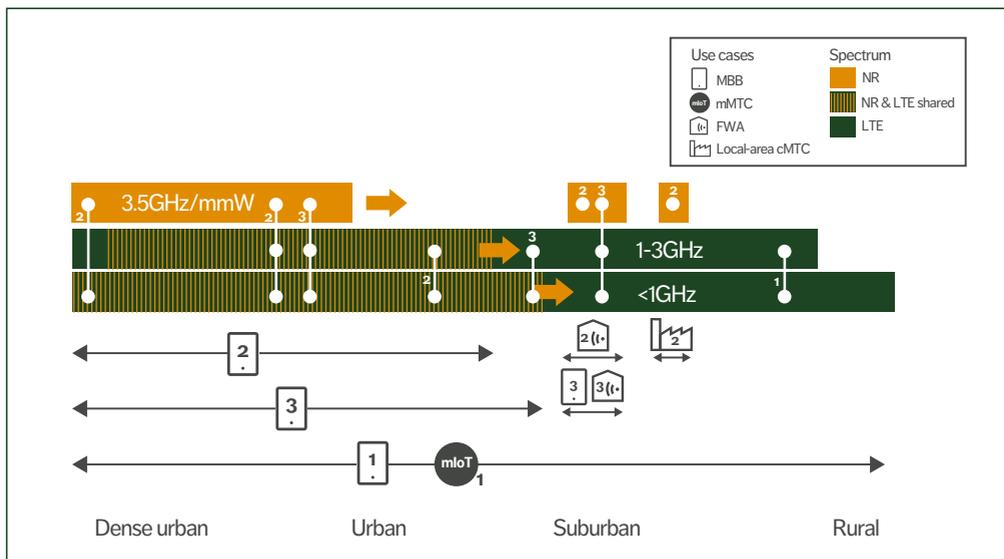


Figure 3 Network deployment during migration, including supported use cases

has enough coverage of sufficient NR spectrum, which can be handled with thresholds and offsets. The possibility of aggregating bands using CA, with a single UL transmitter in the UE, is an important benefit of option 2, compared with the dual connectivity used in options 3, 4 and 7. The third step of Figure 2 shows the use of NR on multiple legacy LTE bands using LTE-NR spectrum sharing.

Options 1 and 3 provide good support for smartphones and MBB. Moving MBB traffic to option 2 requires support for voice telephony. This means that NR must be able to support voice natively, as well as supporting seamless mobility via handover to LTE/EPC when leaving the option 2 coverage area. As an intermediate step before NR supports (and is dimensioned for) voice, the voice service can rely on EPS fallback to LTE/EPC. Tight 5GC-EPC interworking is needed for both voice solutions, and this will also provide good intersystem mobility for other services (see fact box on page 4, dual-mode core network).

When a UE leaves an area where the NR coverage is not good enough for option 2, the network can trigger intersystem mobility to EPC, either to option 3 or 1. The support of option 2 can thus be extended gradually in ever-larger areas in an operator's network, starting with dense urban areas. By deploying option 2 in high-traffic areas first, a significant amount of traffic can be migrated from EPC to 5GC, even if the geographic coverage is initially more limited.

Many LTE sites will be modernized with more advanced radios for improved performance (such as 4T4R) or by adding modern baseband hardware, and will then typically be prepared to support NR on the LTE bands. The deployment of option 2 in a RAN capable of option 3 is then done with a software upgrade. The same gNB will serve some UEs in the same NR cell with option 3 and others with option 2.

Figure 3 illustrates network deployment during the migration from LTE to NR. In selected urban

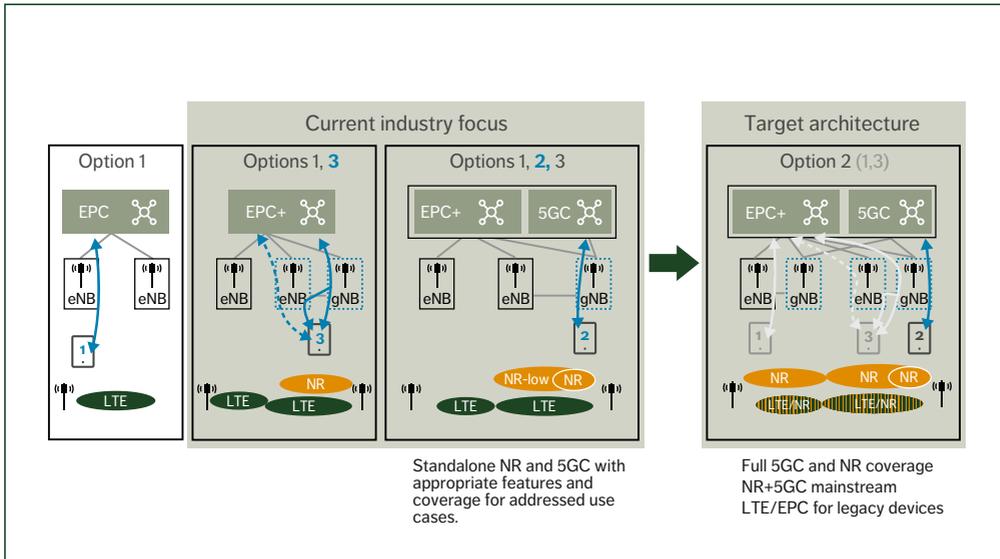


Figure 4 Migration steps toward target architecture

areas, NR (in orange) is deployed on 3.5GHz or mmW to add capacity to the network. NR is also deployed, on a sub-1GHz band (to complement UL coverage), and in legacy LTE bands with LTE-NR spectrum sharing.

The horizontal black lines in Figure 3 represent the coverage of options 1, 2 and 3. Option 1 is used in a large part of the network to support MBB and act as the main solution for Massive Machine Type Communication (mMTC) – specifically, Narrowband Internet of Things (NB-IoT) and LTE-MTC standard (LTE-M). Option 3 can be used anywhere there is NR coverage. Early option 2 deployments in local areas include FWA and industrial deployments. Option 2 for general MBB is supported where there is low-band NR and sufficient NR bandwidth (mid/high band and/or on 1-3GHz).

The orange arrows in Figure 3 indicate that areas of good NR coverage are expanded geographically, covering more urban areas, and in time also extending into suburban areas and beyond.

With the support of options 1, 2 and 3, key use cases such as mMTC, MBB and industrial critical-MTC (cMTC) will be supported in the near- and mid-term with good performance.

Target architecture

The industry has specified a new radio access technology – NR – and a new core network – 5GC – as the foundation for the evolution of 3GPP networks, which, in our view, makes option 2 the long-term target architecture for the industry. In the long-term target architecture, option 2 is deployed with wide coverage, used broadly in most devices, and should be the basis for future investments and feature growth.

Figure 4 illustrates the migration steps to the 5G target architecture for the mobile industry, recognizing option 2 as the long-term target. The first step is to add option 3, followed by option 2 in selected areas. By gradually expanding the areas where option 2 is deployed, the operator and the

●● THE INDUSTRY HAS DECIDED TO BASE THE INITIAL DEPLOYMENT OF 5G ON OPTIONS 3 AND 2 ●●

industry will always invest in steps leading to the long-term target architecture. Eventually, the option 2 coverage will be sufficient to also support wide-area mMTC use cases that will benefit from both NR and 5GC.

At some point in the future there will also be mMTC solutions based on NR/5GC. However, many mMTC services are already adequately served by the existing mMTC solutions NB-IoT and LTE-M. The mMTC services in the low-end Low Power Wide Area (LPWA) segment are just one example. To avoid fragmentation, the best alternative for these use cases is continued use of NB-IoT and LTE-M for a long time.

The timing to reach this long-term target may vary between markets. It should be noted that even when the target is reached, networks will need to continue to support a set of legacy devices (LTE/EPC-based), in particular in the area of mMTC. When the UE penetration for NR support is high enough, selected bands can be fully refarmed to NR-only, as shown in the last step of Figure 2.

Analysis of options 5, 7 and 4

The industry has decided to base the initial deployment of 5G on options 3 and 2. While options 5, 7 and 4 may initially seem beneficial for specific operators' deployment cases, it is important to recognize that none of them are direct steps leading toward the long-term target architecture. Further, the use of options 5, 7 and 4 would add unnecessary complexity in the target architecture, in the interaction with other network functions, and in the

evolution of new features, which needs to take the combination of all existing options into account.

After a thorough analysis of options 5, 7 and 4 that encompassed the main drivers, potential benefits and drawbacks, we have come to the conclusion that all three can and should be avoided. We have also identified preferred alternative solutions for each option.

Option 5

The main driver for deploying option 5 is to allow devices that move outside the area covered by option 2 to remain connected to 5GC, which would also increase the 5GC coverage to eLTE areas.

A key question to consider is: which use cases require nationwide 5GC coverage? Traditional MBB/voice obviously requires wide-area support, but this is well supported with intersystem mobility during the build-out of NR coverage, as it was in previous generation shifts. 5GC provides a range of new values but the need for other wide-area 5GC-based services in the near term is undefined. In the longer term, we expect wide-area option 2 to enable the new use cases that emerge.

Option 5 could be used to increase wide-area 5GC coverage, but reaching full wide-area 5GC coverage would take time and investment, as it would require new UEs, new RAN functionality and retesting the system. Option 5 would have a major impact on the UEs in terms of supporting the 5GC non-access stratum and the new parts of the eLTE radio interface, as legacy LTE devices are not supported. In addition, substantial interoperability retesting between networks and UEs would be required to ensure the operation of legacy features and services, including VoLTE. Further, option 5 requires substantial upgrades of the eNB software and, in many cases, the eNB baseband hardware as well.

In summary, option 5 is unlikely to provide a faster route to 5GC wide-area coverage than the wide deployment of option 2. Wide deployment of option 2 is the better alternative, particularly since option 5 would mean investing in technology that does not capitalize on the benefits of the latest radio technology (NR).

Option 7

Option 7 builds on option 5 and cannot exist without it. If option 5 were to be used, it is very likely that option 7 would also be supported in areas with NR. The driver for option 7 is the same as for option 3; that is, to use dual connectivity to aggregate NR and LTE bands to enhance capacity, but in this case for a UE connected via eLTE to 5GC. According to the same logic explained in the Option 5 section above, we recommend using option 2 instead.

THE MOBILE INDUSTRY HAS AN OPPORTUNITY TO SIMPLIFY THE 5G ECOSYSTEM

Option 4

Option 4 is an addition to option 2, using dual connectivity to add eLTE to an NR anchor. It is primarily relevant when serving MBB traffic via 5GC. The driver for option 4 is to maximize throughput when the amount of NR spectrum is limited. An example of this type of situation would be if NR is deployed on 700MHz, 3.5GHz and mmW, but the UE is outside coverage of the two higher bands.

In terms of drawbacks, option 4 would require new software support in eNB, gNB and UE, with related interoperability testing. Further, the future evolution of features would need to consider option 4,

and its use would require continued investments in eLTE deployments for a long time.

Option 4 is not necessary, and performs worse than option 2 with NR-NR CA and enough NR spectrum, in areas serving MBB via 5GC. Using option 2 instead of option 4 also focuses investments on the rollout of the long-term target architecture.

Conclusion

Our analysis shows that the mobile industry has an opportunity to simplify the 5G ecosystem by focusing network deployments on connectivity options 3 and 2, which are capable of delivering all the 5G benefits without adding unnecessary complexity and cost (as in options 5, 7 and 4). The flexible design of radio and core networks supports a smooth migration with LTE-NR spectrum sharing and dual-mode core technologies. The regulation of frequency bands should allow NR deployment in existing LTE bands that are in sync with the required spectrum migration.

Operators have the opportunity to avoid connectivity options 5, 7 and 4 by implementing a proactive spectrum migration strategy that considers NR for new low bands, and by refarming or introducing LTE-NR spectrum sharing in existing low/mid bands. This approach will reduce network upgrade cost and time, simplify interoperability between networks and devices, and enable a faster scaling of the 5G ecosystem.

A 5G deployment approach based exclusively on options 3 and 2 ensures that investment is focused on the long-term target architecture, leveraging full 5GS capabilities. Early key use cases for wide-area, like MBB including voice services, are fully supported during the migration period, along with services to existing devices.

Terms and abbreviations

4T4R – 4-Branch Transmit/Receive Antenna and Radio Arrangement | **5GC** – 5G Core | **5GS** – 5G System | **CA** – Carrier Aggregation | **cMTC** – Critical Machine Type Communication | **CN** – Core Network | **DC** – Dual Connectivity | **DECOR** – Dedicated Core Network | **DL** – Downlink | **E2E** – End-to-end | **eLTE** – Evolved LTE | **eNB** – Evolved Node B | **EN-DC** – E-UTRA – NR Dual Connectivity | **EPC** – Evolved Packet Core | **FWA** – Fixed Wireless Access | **gNB** – Next Generation Node B | **IoT** – Internet of Things | **LPWA** – Low Power Wide Area | **LTE-M** – LTE-MTC Standard | **MBB** – Mobile Broadband | **mMTC** – Massive Machine Type Communication | **mmW** – Millimeter Wave | **NB-IoT** – Narrowband Internet of Things | **NR** – New Radio | **RAT** – Radio Access Technology | **UE** – User Equipment | **UL** – Uplink

References

1. **3GPP Release 15 specifications, e.g. TS 23.501, TS 38.401, available at:** <http://www.3gpp.org/release-15>
2. **Ericsson Mobility Report, available at:** <https://www.ericsson.com/en/mobility-report>
3. **Ericsson Technology Review, November 2018, The advantages of combining 5G NR with LTE, available at:** <https://www.ericsson.com/en/ericsson-technology-review/archive/2018/the-advantages-of-combining-5g-nr-with-lte>

Further reading

- » **Ericsson, 5G deployment options, 2018, available at:** <https://www.ericsson.com/assets/local/narratives/networks/documents/5g-deployment-considerations.pdf>
- » **Ericsson, Core evolution from EPC to 5G Core, download available from:** <https://pages.digitalservices.ericsson.com/core-evolution-to-5g>

THE AUTHORS



Torbjörn Cagenius

◆ is a senior expert in network architecture at Business Area Digital Services. He joined Ericsson in 1990 and has worked in a variety of technology areas such as fiber-to-the-home, main-remote RBS, fixed-mobile convergence, IPTV, network architecture evolution, software-defined networking and Network Functions Virtualization. In his current role, he focuses on 5G and associated network architecture evolution. He holds an M.Sc. in electrical engineering from KTH Royal Institute of Technology in Stockholm, Sweden.



Anders Ryde

◆ is a senior expert in network and service

architecture at Business Area Digital Services, based in Sweden. He joined Ericsson in 1982 and has worked in a variety of technology areas in network and service architecture development for multimedia-enabled telecommunication, targeting both enterprise and residential users. This includes the evolution of mobile telephony to IMS and VoLTE. In his current role, he focuses on bringing voice and other communication services into 5G, general 5G evolution and associated network architecture evolution. He holds an M.Sc. in electrical engineering from KTH Royal Institute of Technology in Stockholm, Sweden.



Jari Vikberg

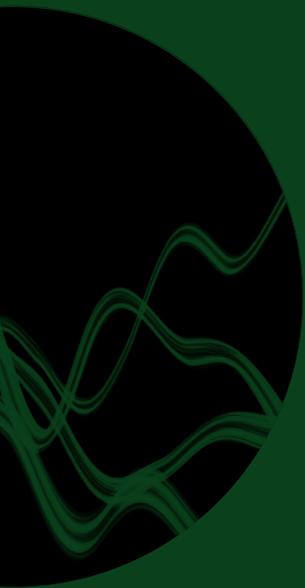
◆ is a senior expert in network architecture and the chief network architect at CTO office. He joined Ericsson in 1993 and has both wide and deep technology competence covering network architectures for

all generations of RANs and CNs. He is also skilled in the application layer and other domains, and the impact and relation these have to mobile networks. He holds an M.Sc. in computer science from the University of Helsinki, Finland.



Per Willars

◆ is an expert in network architecture and radio network functionality at Business Area Networks. He joined Ericsson in 1991 and has worked intensively with RAN issues ever since. This includes leading the definition of 3G RAN, before and within the 3GPP, and more lately indoor solutions. He has also worked with service layer research and explored new business models. In his current role, he analyzes the requirements for 5G RAN (architecture and functionality) with the aim of simplifying 5G. He holds an M.Sc. in electrical engineering from KTH Royal Institute of Technology.



ISSN 0014-0171
284 23-3320 | Uen

© Ericsson AB 2018
Ericsson
SE-164 83 Stockholm, Sweden
Phone: +46 10 719 0000