# Ericsson Technology Review

#5, May 2024

# Hybrid 5G Core network architecture with full 4G and 5G device support

Charting the future of innovation

Ericsson Technology Review · 1924 · 2024
L.M. Ericsson Review
Celebrating 100 Years

Ericsson Technology Review – Charting the future of innovation | #5, May 2024

# Hybrid 5G Core network architecture with full 4G and 5G device support

Authors:
Torbjörn Cagenius, Ralf Keller, Lasse Olsson

In the future, when most devices are served by the 5G Core network, communication service providers will need an efficient way to manage legacy devices. Ericsson's approach simplifies the network architecture and harmonizes the service offering by creating a hybrid core network that can support legacy 4G and 5G non-standalone devices alongside 5G standalone devices.

ISSN 0014-0171 284 23-3410 | Uen
© Ericsson AB 2024
Ericsson, SE-164 83 Stockholm, Sweden
Phone: +46 10 719 0000



**Mobile network architecture evolution in the 2020s has been characterized by the initial 5G buildout and the ongoing transition to the 5G System (5GS) based on 5G New Radio (NR) standalone (SA) architecture and the 5G Core (5GC) network.**

By 2028, it is projected that the global population coverage of 5G will reach 85 percent and the majority of traffic will be handled by 5G networks [1]. Network support for 2G and 3G will soon be phased out, as those communication service providers (CSPs) that have not already done so are currently in the process of shutting these networks down.

To accelerate the adoption of 5G services, new 5GC-enabled capabilities and Service-Based Interfaces (SBIs) will be needed. The SBIs will replace some of the legacy Diameter interfaces and support new charging models defined in the 5GS. This migration presents an opportunity to harmonize policy and charging across 5G non-standalone (NSA) and 5G SA devices and thereby reduce operational cost.

CSPs will also need the capability to handle legacy 4G devices efficiently from both a core network architecture perspective and an operational perspective all the way to the end of the 4G era when most spectrum will be allocated to 5G and capacity shifts to handle 5G – and eventually 6G – traffic. Examples of 4G legacy devices that will require support include low-cost feature phones and long-lived Massive Internet of Things (M-IoT) devices such as the low-power wide-area (LPWA) devices used in utility companies' power meters and in cars with 4G access. Many of these devices have lifetimes that are measured in decades. The most efficient way to prepare the core network architecture for the eventual sunsetting of 4G is to use the network functionality provided by the 5GC.

The hybrid core architecture we propose is designed to provide future-proof support for the long tail of legacy 4G and 5G NSA devices in the network. Based on the 5GC, with common subscription management, policy and charging, our hybrid core architecture proposal minimizes the need for system integration, optimizes operations and aligns with 5GC ways of working, including the use of SBIs. It is even possible to decommission some legacy Evolved Packet Core (EPC) nodes with no impact on existing 4G radio access networks (RANs) and devices.

## Terms and abbreviations

**3GPP** – 3rd Generation Partnership Project | **5GC** – 5G Core | **5GS** – 5G System | **AMF** – Access and Mobility Management Function | **APN** – Access Point Name | **AS** – Access Stratum | **CCS** – Converged Charging System | **CHF** – Charging Function | **CSP** - Communications Service Provider | **DNS** – Domain Name System | **EPC** – Evolved Packet Core | **E-UTRAN** – Evolved Universal Terrestrial Radio Access Network | **FWA** – Fixed Wireless Access | **gNB** – gNodeB | **HSS** – Home Subscriber Server | **IMS** – IP Multimedia Subsystem | **LPWA** – Low Power Wide Area | **LTE** – Long Term Evolution | **M-IoT** – Massive Internet of Things | **MME** – Mobility Management Entity | **NAS** – Non-Access Stratum | **ng-eNB** – Next-Generation eNodeB | **NG-RAN** – Next-Generation Radio Access Network | **NNI** – Network-to-Network Interface | **NR** – New Radio | **NSA** – Non-Standalone | **OCS** – Online Charging System | **OFCS** – Offline Charging System | **PCF** – Policy Control Function | **PCRF** – Policy and Charging Rules Function | **PDN** – Packet Data Network | **PDU** – Protocol Data Unit | **PGW** – Packet Data Network Gateway | **PGW-C** – PGW Control Plane Function | **PGW-U** – PGW User Plane Function | **RAN** – Radio Access Network | **SA** – Standalone | **SBI** – Service-Based Interface | **SGW** – Serving Gateway | **SMF** – Session Management Function | **UDICOM** - User Data Interworking, Coexistence and Migration | **UDM** – Unified Data Management | **UPF** – User Plane Function

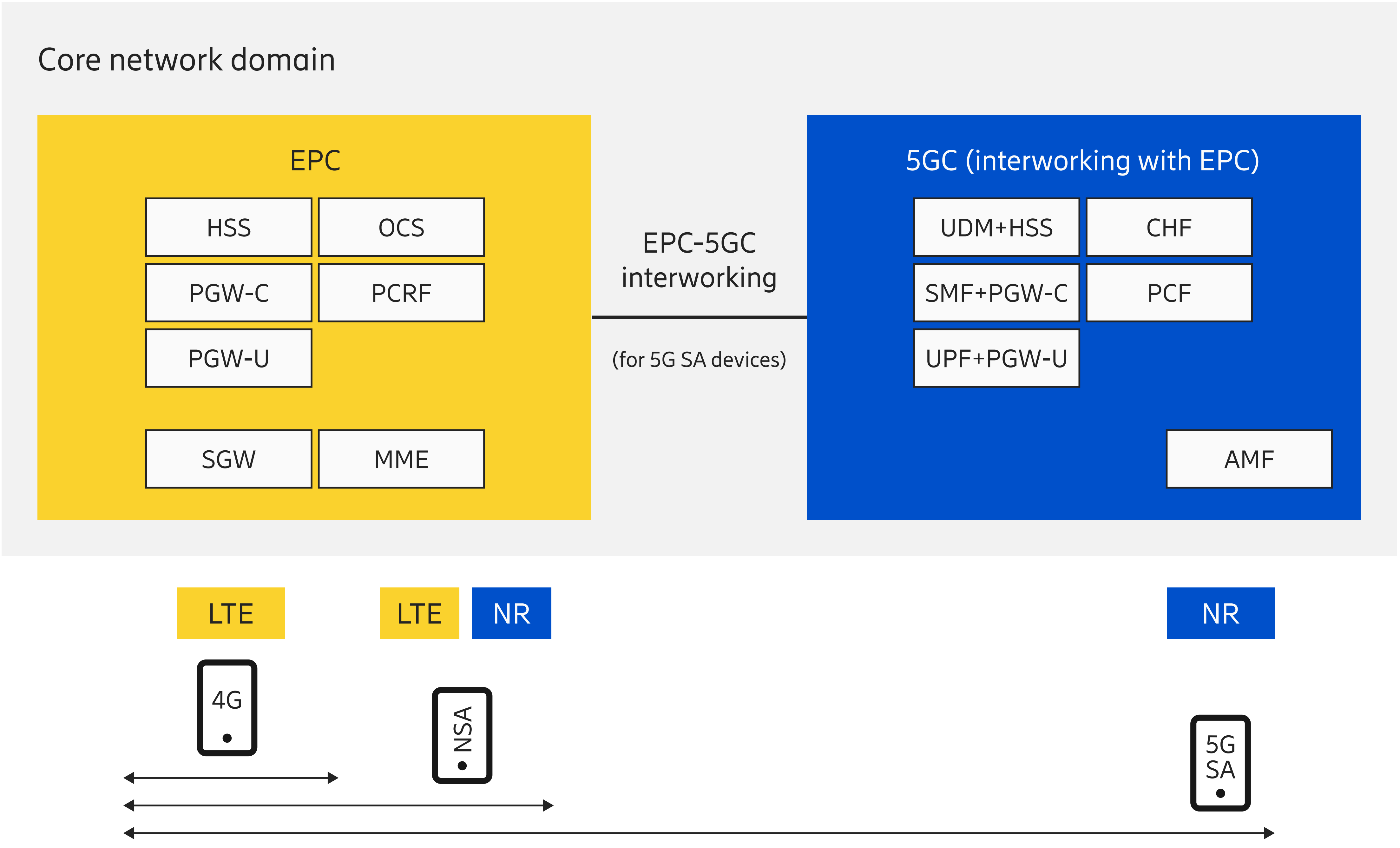


**Figure 1**: EPC-5GC interworking architecture

## 5GC-capable devices versus legacy 4G devices

The 5GS was initially specified for devices that have the corresponding support – that is, devices that support all required functionalities and in particular interfaces to interact with the next-generation RAN (NG-RAN) and the 5GC. These devices interact directly with the gNodeB (gNB) in the NG-RAN and with the Access and Mobility Management Function (AMF) in the 5GC. The communication layer for wireless communication is between the device and the gNB in the NG-RAN known as the access stratum (AS), which includes the Radio Resource Control.

The communication layer between the device and the AMF is called the non-access stratum (NAS). All signaling related to registration in the 5GC, protocol data unit (PDU) session establishment and configuration updates from the AMF are communicated through NAS. We refer to devices that support this type of communication with the AMF as "5GC-capable devices." The terms "NR SA devices" and "5G SA devices" are used synonymously.

Legacy 4G devices use EPC NAS to communicate with the Mobility Management Entity (MME) and thereby to connect to EPC over LTE (Long Term Evolution). 5G NSA devices are also considered to be legacy 4G devices from a core-network perspective because they use EPC NAS and not 5GC NAS.

## The EPC-5G Core interworking architecture

To enable a smooth introduction of the 5GS, the 3GPP (3rd Generation Partnership Project) has defined an interworking architecture [2] between the EPC and 5GC to support 5GC-capable devices that need to move to 4G in cases where the 5G SA coverage is not fully built out, as shown in **Figure 1**.

The interworking architecture enables the device to retain the IP address during mobility between the Evolved Universal Terrestrial Radio Access Network (E-UTRAN) (4G/5G NSA) and NG-RAN (5G SA) as a prerequisite for uninterrupted service, in particular for, but not limited to, voice. The connections for data transmission (PDU sessions if established through the NG-RAN and a packet data network (PDN) connection if established via the E-UTRAN) are always anchored in the SMF+PGW-C (session management function + packet data network gateway control plane) of the 5GC. The 5GC-capable devices – marked as 5G SA devices in Figure 1 – are always anchored in the 5GC, regardless of whether they are connecting over 4G or 5G access. Legacy 4G and NSA devices are anchored in the EPC, which includes using the Home Subscriber Server (HSS).

When a 5GC-capable device is accessing the network over E-UTRAN, the MME is responsible for selecting an SMF+PGW-C as the session anchor point. This occurs during PDN connection establishment if the HSS subscription data allows the session to be anchored in the 5GC and the device supports communication with the 5GC. When all of these criteria are fulfilled, the MME performs a special query toward the Domain Name System (DNS) to find the address of an SMF+PGW-C. In the message to create the session, an indication is set for the 5GC, which thereby dictates the use of the 5GC SBIs – in this case, N7 and N40 for policy and charging respectively.

## Extending the anchoring possibilities to legacy 4G and 5G NSA devices

In Release 16, the 3GPP enhanced the EPC-5GC interworking architecture by extending the anchoring possibilities beyond 5GC-capable devices to also include legacy 4G devices (not supporting 5GC NAS) in a 5GC gateway (SMF+PGW-C). This improves the network evolution flexibility. We call this the hybrid core architecture since it supports the migration of legacy 4G and 5G NSA device types, enabling them to connect to the 5GC.

The new capability for the SMF+PGW-C is to support parameter conversion between the EPC and 5GC through the assignment of a device-unique PDU Session ID per PDN connection on behalf of the device. This PDU Session ID is used internally inside the 5GC network – for example, toward Unified Data Management (UDM), but the PDU Session ID is not sent to the legacy 4G and 5G NSA devices (these devices would be unable to use it).

The MME is enhanced to facilitate granular selection of an SMF+PGW-C for legacy 4G and 5G NSA devices in case legacy PGWs are still present in the network. If not, there is no need for the MME to differentiate. The selection logic may also support forcing a device on a specific Access Point

Name (APN) to SMF+PGW-C, while using a legacy PGW for other APNs. This enables a stepwise migration to 5GC per APN. The resulting network architecture is presented in **Figure 2**, which shows the main core network functions and interfaces supporting the 4G and 5G NSA devices.

Assuming that the network is configured correctly and that the subscription data for the user exists in the UDM of the 5GC, implementation of the hybrid core architecture will result in the following sequence for PDN connection establishment.

## The hybrid core architecture supports the migration of legacy 4G and 5G NSA devices to connect to the 5GC.

The procedure to establish a PDN connection is initiated by the device through the RAN to the core network in the same way a 5GC-capable device would connect over LTE. During this procedure an IP address is assigned to the device for that specific connection. Once the procedure is successfully concluded, the device can send and receive data from external sources. For the core network to correctly establish the connection, the device will provide an APN to the core network. This APN will be crosschecked toward the subscription data in the HSS, assuring that the user is authorized to use the service.

The MME then uses the APN in an extended query sent to the DNS together with an indicator that a hybrid core solution is wanted (an enhancement of the MME). The DNS responds with the address to the session anchor point, SMF+PGW-C. The MME uses the geographical location of the device to find a serving gateway (SGW).

A message is then sent from the MME through the SGW to the SMF+PGW-C, to allocate resources and establish a connection for this specific device. In this message, the MME informs the SMF+PGW-C, that anchoring the connection in the SMF+PGW-C is allowed by the subscription data.

The SMF+PGW-C selects a UPF+PGW-U instance, and assigns an IP address and a PDU Session ID. The latter is used by the UDM, policy control function (PCF) and charging function (CHF) to identify the user and the session. The SMF+PGW-C uses a network repository function (NRF), which is a fundamental part of the Service-Based Architecture, to discover the UDM, PCF and CHF. Once all resources have been successfully allocated by the core network, the assigned IP address will be returned to the device in the response to the initial request to establish a PDN connection.

### Impact on adjacent architecture domains

The hybrid core architecture includes changes to multiple architecture domains, allowing the management of legacy 4G and 5G NSA devices in the same way as 5G SA devices. As shown in Figure 2, the 5GC SBIs N7, N10, N28 and N40 are used to manage user profiles, policy and charging.

Subscription and policy information in the 5GC user data repository must also allow legacy 4G and 5G NSA devices to use the UDM and PCF. This solution makes it possible

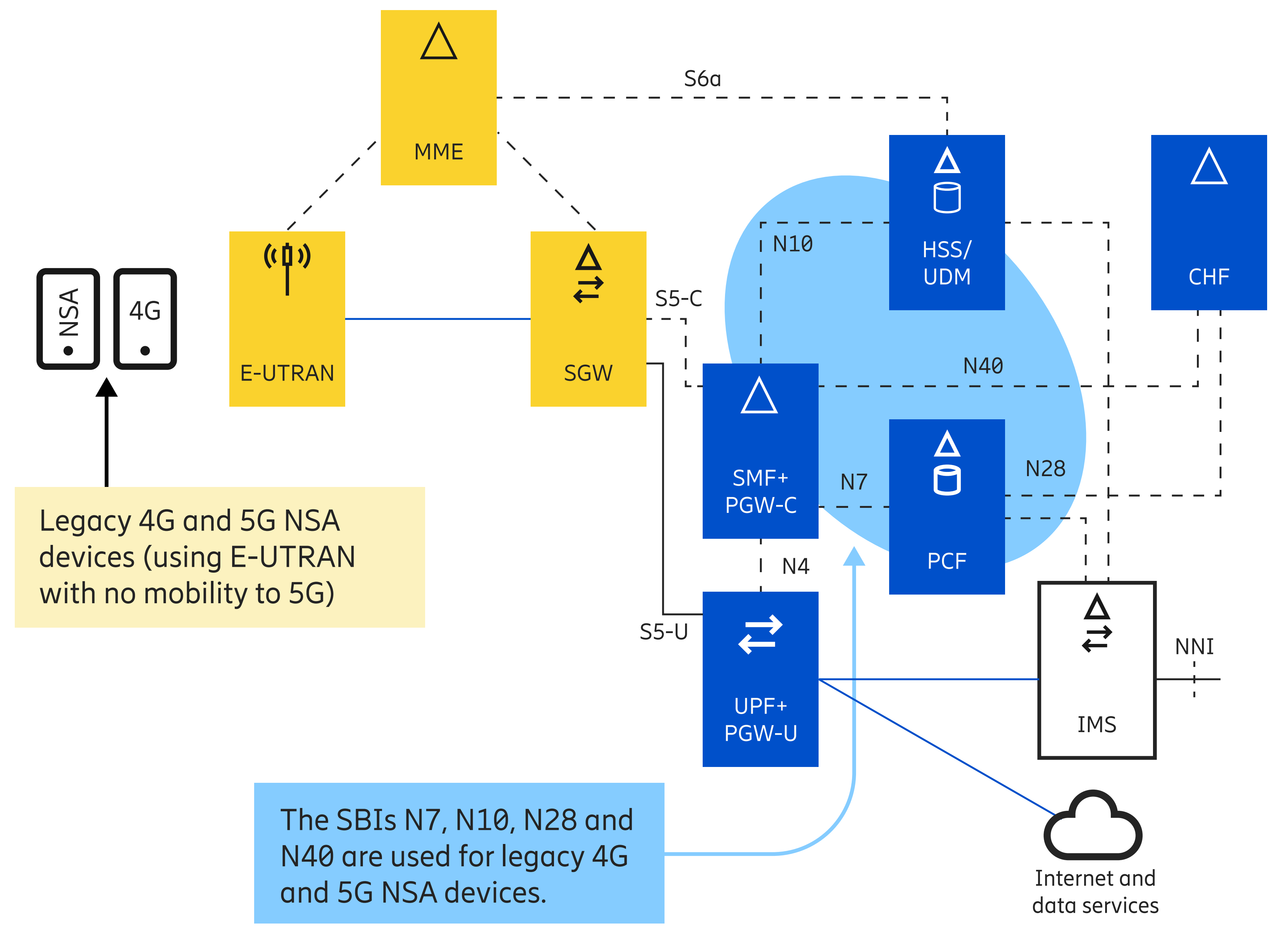


**Figure 2:** Key concept of the hybrid core architecture

to manage both new 5GC-enabled subscribers regardless of the access technology used and legacy 4G users who connect only through the E-UTRAN in a common solution. The subscription data update includes information for the MME to select the SMF+PGW-C in the 5GC as the anchor point. This information is available to the MME through the HSS part of the solution (over the legacy Diameter interface S6a).

Deployment of the 5GC implies that the online charging system (OCS) and offline charging system (OFCS) are modernized to a 5GC converged charging system (CCS). This makes it possible to remove several legacy Diameter interfaces and replace them with SBI-based 5GC interfaces such as N28 (PCF to CHF) and N40 (SMF+PGW-C to CHF). Adding legacy 4G and 5G NSA devices to the CCS implies that the CHF part of the CCS is responsible for all charging data records generation in the 5GC, including all traffic for the legacy 4G and 5G NSA devices. Online rating and balance management of the 4G and 5G NSA devices' traffic works the same way as it did in the OCS.

## Additional architecture considerations

The hybrid core architecture is the first step toward fully embracing the new 5GC architecture. However, the existing EPC network functions PGW and PCRF (policy and charging rules function) as well as OCS/OFCS may still be needed for purposes such as outbound roaming, supporting other business verticals and/or M-IoT that depend on 2G/3G radio network usage. For parts of the legacy network, it will also be necessary to keep existing Diameter interfaces. For example, Diameter is used for IMS (IP Multimedia Subsystem) as well as in the packet core between an MME and HSS (S6a) for attach and mobility procedures.

One aspect of the 3GPP-specified EPC-5GC interworking architecture that is also applicable to the hybrid core is that there is no IP session continuity if mobility from 4G to 2G/3G occurs, which depends on the 4G coverage. The signaling load could increase if this mobility happens too frequently. If the CSP has shut down 2G/3G or has full coverage in suitable bands for 4G, this will not be a problem.

## Key benefits

The hybrid core architecture delivers significant value compared with the device connectivity option of using 5GC NAS over LTE that was specified by the 3GPP but not adopted in the industry (known as "option 5 with ng-eNB"). That option would require a major network upgrade including new devices and LTE RAN, and even then it would not support legacy 4G devices as they are not 5GC NAS capable.

A key benefit of the proposed hybrid core architecture is that it makes it possible to modernize user profiles, policy and charging by using single policy and charging functionalities in the 5GC for legacy 4G devices, 5G NSA devices and 5GC-capable devices. This provides opportunities for savings in system integration and in total cost of ownership, enabling continued network and service evolution.

Another benefit of the proposed architecture is that the migration of each use case to the hybrid core offloads the legacy nodes. Once all use cases have been migrated to the hybrid core and any additional architecture considerations described above have been taken into account, network deployment can be simplified by decommissioning the legacy PGW, charging and policy functionality. As a result, the associated Diameter-based interfaces such as Gx (policy control), Gy (online charging) and Sy (spending limit reporting) will no longer be needed.

The migration of 5G NSA devices to the hybrid core architecture could be regarded as a step toward 5G SA, as the devices will be transferred to a common subscription management, policy and charging solution based on the 5GC. This step will facilitate a transition from Diameter to SBI-based charging and spending limit control using the 5GC charging architecture. That will help familiarize CSPs with the operational methods of SBI-based charging architecture and prepare for the addition of new charging trigger points that the 5GC architecture brings, enabling more advanced charging schemes. Examples of such trigger points are related to service exposure, network slice admission control, network slice management and data analytics.

Once a CSP has successfully launched 5G SA to handle the majority of its own mobile broadband subscribers, or perhaps even started the transition to 6G, there will be a long tail of 4G devices that still require support. These can be the CSP's own subscribers with feature phones, inbound roaming devices or 4G M-IoT devices (NB-IoT and Cat-M). The M-IoT business is characterized by long usage cycles, where CSPs are offering contracts to automotive companies, utilities and others with up to 20 years of device support. Another key aspect to consider is the need for low-cost radio modems in M-IoT devices, driving the requirement on simple architecture and high volumes. The hybrid core network architecture is a cost-efficient solution for these types of devices.

## The migration process

A CSP that wants to implement the hybrid core architecture must start by taking a use-case perspective and deciding which device categories and APNs to move for that particular use case. Second, it must consider a list of prerequisites for the selected use case prior to migration. Third, it must perform the actual migration for the device categories and APNs. Finally, it needs to perform post-migration actions.

A CSP typically supports a large variety of use cases for its own subscribers on 4G, including consumers, enterprises, automotive and fixed wireless access (FWA), for example. A CSP may decide to migrate all APNs applicable for a use case to the hybrid core, or it may decide to migrate only one APN – enterprise, for example – to hybrid core, leaving other APNs on legacy PGWs.

If a CSP decides to support legacy 4G devices on the hybrid core, for example, it may choose to migrate one use case after the other. Prior to each migration, it needs to determine the prerequisites for that particular use case and then perform the actual migration. For example, a CSP may choose to migrate FWA devices first, then enterprise devices, followed by consumer smartphone devices. Once the CSP has migrated the first use cases, it may not be necessary to perform every step again.

A key benefit of the proposed hybrid core architecture is that it makes it possible to modernize user profiles, policy and charging.

Our checklist of prerequisites for the migration is based on the architecture depicted in Figure 2. One of the first steps is to provision the subscription data for the legacy 4G and 5G NSA devices in the 5GC UDM. For these devices there is subscription data accessed through the HSS and the UDM. The legacy HSS is typically evolved to support the UDICOM (NU1) interface to the UDM. As a result, the legacy HSS deployment realizes the HSS part in the HSS+UDM function shown in Figure 1 and is no longer needed as a separate network function.

The required capabilities to handle the particular use case must be enabled in the SMF+PGW-C, the MME must be enhanced and the DNS configured to select an SMF+PGW-C for legacy 4G and 5G NSA devices.

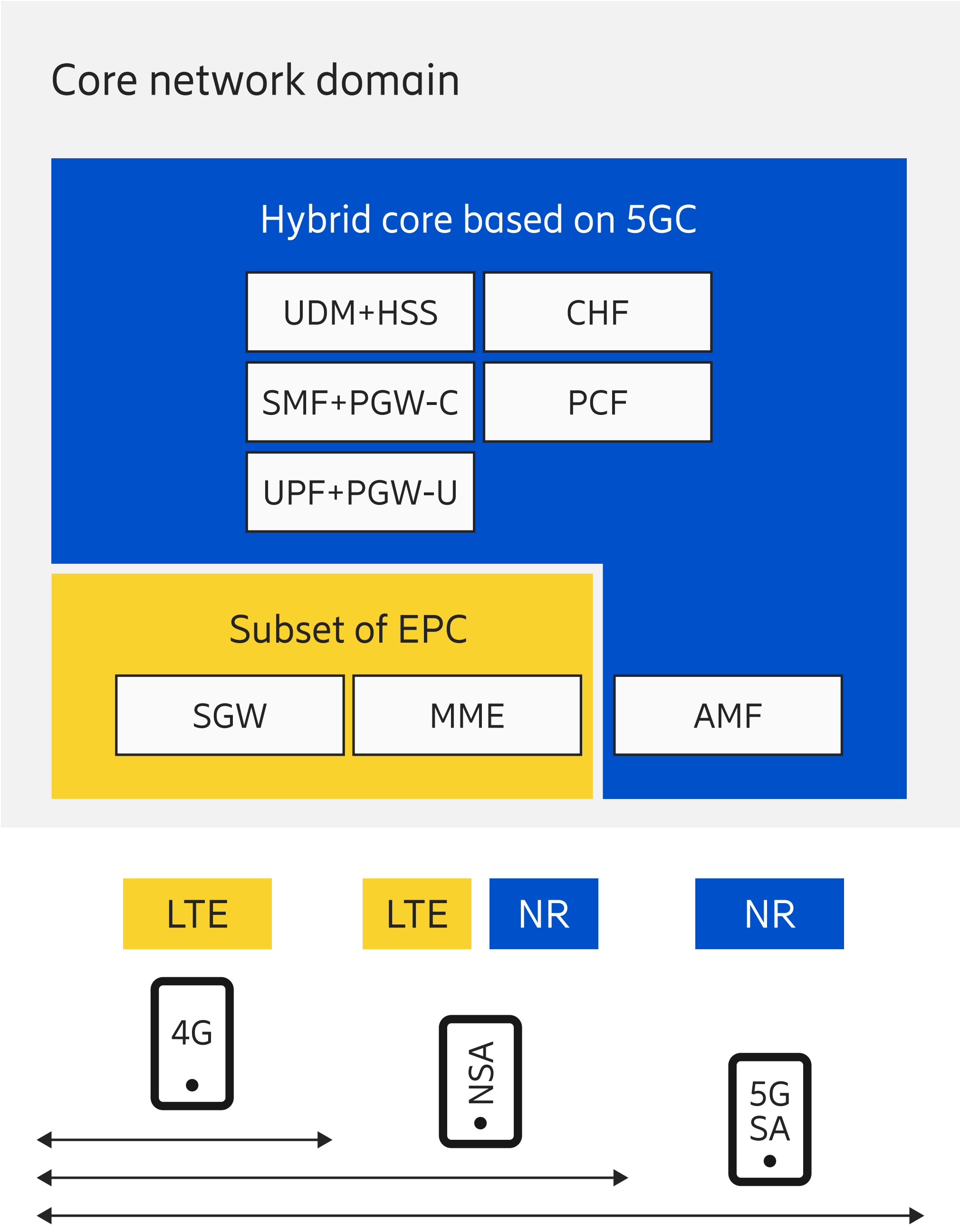


**Figure 3**: Network simplification through migration to a hybrid core architecture

The capacity needs on legacy nodes may reduce before the migration is finished. However, the CSP may only start to decommission legacy nodes that are no longer needed (such as PGW-C, PGW-U, OCS and PCRF) once all use cases (and therefore all devices) have been migrated to the hybrid core. When the decommissioning is complete, the network simplification shown in **Figure 3** will be the result. When this simplification has been achieved, the core network domain includes only the minimal set of network functions needed to serve all the devices shown in the figure. In particular, only the SGW and MME are mandatory as a subset of EPC to maintain the interface toward E-UTRAN and 4G devices.

## Conclusion

Migration toward our proposed hybrid core architecture offers many benefits to communication service providers (CSPs) that want to modernize their network architecture without introducing disruptive ecosystem changes late in the 4G life cycle. We believe that architecture evolution should not require any changes to the legacy 4G and 5G non-standalone (NSA) devices or to the already widely deployed LTE (Long Term Evolution) radio access network.

Based on features specified in 3GPP (3rd Generation Partnership Project) Release 16, our hybrid core architecture proposal makes it possible to anchor legacy 4G and 5G NSA devices on the 5G Core (5GC), creating the opportunity to apply 5GC operational methods for handling subscription, policy and charging to both 5GC-capable and legacy 4G and 5G NSA devices. It further provides possibilities for savings in system integration and operational costs when aligning with the 5GC ways of working, such as using Service-Based Interfaces instead of Diameter interfaces.

The migration of 5G NSA devices to the hybrid core architecture is the most obvious early use case, as the adoption of 5GC capabilities delivers a variety of benefits that enable continued network and service evolution.
In the longer term, CSPs will want to use the hybrid core architecture to secure cost-efficient support for the long tail of legacy 4G devices in the network, including Massive Internet of Things devices. Hence, the hybrid core architecture provides a single solution for all 4G and 5G devices.

# The authors

**Torbjörn Cagenius** is a senior expert in network architecture at Business Area Cloud Software and Services. He joined Ericsson in 1990 and has worked in a variety of technology areas such as fiber-to-the-home, main-remote radio base stations, fixed-mobile convergence, IPTV and network architecture evolution. In his current role, he focuses on 5G and network architecture evolution toward 6G. Cagenius holds an M.Sc. in electrical engineering from KTH Royal Institute of Technology in Stockholm, Sweden.

**Ralf Keller** joined Ericsson in 1996. He is an expert in core network migration and his current focus is on packet core architecture and technology. His work includes both technology studies and contributions to product strategies for mobile communication, including communication services in 5G, migration to the 5GS, migration to 6G, and interworking and coexistence with legacy networks. He is also active in the GSMA, where he works on the profiling of 5G and 5G roaming and network slicing. Keller holds a Ph.D. in computer science from the University of Mannheim in Germany.

**Lasse Olsson** is an expert in core network evolution realization at Business Area Cloud Software and Services. He has over 30 years' experience from telecommunication, joined Ericsson for the first time in 1991 and has worked in a variety of technology areas such as fixed telephony, mobile telephony (circuit switched and packet switched) and network architecture evolution. Olsson has had many different roles ranging from tester to system owner. In his current role, he focuses on how to make 5G simple, programmable and the network architecture evolution toward 6G.

## References

1. Ericsson Mobility Report, June 2023 ↗
2. Ericsson Technology Review, 5G migration strategy from EPS to 5G System, February 24, 2020, Keller, R; Cagenius, T; Ryde, A; Castellanos, D ↗

## Further reading

- Ericsson, Future network architecture ↗
- Ericsson blog, How to unlock the full potential of mission-critical 4G and 5G ↗
- Ericsson, Discover the power of 5G ↗
- Ericsson blog, Non-standalone and Standalone: two standards-based paths to 5G ↗
- Ericsson, Evolve your core network for 5G ↗
- Ericsson blog, Enabling 5G revenue opportunities through charging ↗
- Ericsson Technology Review, 5G BSS: Evolving BSS to fit the 5G economy ↗
- 3GPP TS 23.501; System architecture for the 5G System (5GS) ↗

## Acknowledgements

The authors would like to thank David Castellanos, Lars Frid and Annika Åhlberg for their contributions to this article.