The Facts on 5G

How 5G networks are being built in the real world

26 July 2019

David Kennedy

Summary

In brief

This report explains how 5G standards set up a stable, standardized and ongoing separation between the 5G core (5GC) network which controls 5G network operations and the radio-access network (RAN) which manages the air interface. This separation allows operators to build resilient multi-vendor 5G networks that combine a RAN from one or more vendors and a core from another, and there are now numerous examples of this to avoid over-dependence from a single provider and increase competition. Finally, the report addresses the confusion surrounding the concept of “multi-access edge computing” and explains that this has no new implications for core/RAN separation or for network security.

Ovum view

As 5G networks have begun to rollout, operators have been looking for ways to minimize cost, increase network resilience and speed up implementation. In some cases, they have looked to a multi-vendor solution to achieve this.

The 5G standard facilitates this approach. By creating a number of standardized interfaces within the 5G architecture, the designers of 5G have created opportunities for competitive supply of different parts of the network.

5G security is an evolution of the 4G security with enhancements to address known 4G vulnerabilities. It introduces no radical changes to a security architecture based on a clear core/RAN separation. Central to this architecture is the interface between the core network which controls 5G network communications (including access, mobility, session management, authentication and data encryption) and the RAN which conveys signals between the user terminal and the core network. A stable and standardized interface gives vendors confidence that equipment for either RAN or core can interoperate with another vendor’s equipment.

This also has implications for the security of the 5G network. User authentication and data encryption functions are managed and controlled by the core. The RAN, in contrast, functions as a pipe between the core infrastructure and the mobile device. The signals and data conveyed through the RAN remain encapsulated between core and terminals, so that external sources do not have access to unencrypted traffic. The core always retains control of 5G call security. Moreover, the inbound and outbound traffic between RAN and core is always encrypted using a Security Gateway (SGw). The SGw is an essential component – beyond the scope of the 3GPP standards – of the 5G network security architecture to be deployed by mobile operators to protect their security control zone. Choosing a different RAN vendor does not change this.

The upshot is that managing the security of the 5G network is similar to 4G. In particular, any security risk in the RAN can be managed as done for earlier network generations, provided that operators ensure proper configuration of 5G security functions and deploy a 5G network security architecture, end to end. The 3GPP standards set out the security architecture, but it remains imperative that vendors and operators support and implement them consistently.

Key messages

- **5G security is an evolution of 4G security**. 5G introduces new technologies, particularly support for virtualizing core network control plane functions. This requires some new security features. However, the overall 5G security architecture builds on 4G.
- **The Core/RAN distinction is maintained**. The basic security architecture of mobile communications, including RAN/core separation, does not change in 5G.
- **RAN/core separation facilitates multi-vendor operation**. One powerful reason why RAN/core separation has been maintained in 5G standards is to reduce over-dependence on vendors and increase competition. First, the network should not be dependent on just one vendor, as this would render it less resilient. Secondly, competition between vendors will force them to improve their security standards. And it is this raising of the threshold on cyber security standards across the board that is needed, along with compliance to more stringent regulations and enforcement of those standards.
- **The "edge" is not the RAN**. There is some confusion whether shifting core network functions to the "edge" (i.e. closer to the user) implicates the RAN. In fact, it does not. Multi-access Edge Computing (MEC) on the network does not affect the core/RAN separation, as MEC is defined and implemented as an Application Function (AF) of the core network in 5G.

5G architecture and RAN/core separation

5G security architecture

5G technology standards are decided by the 3GPP. The 3GPP is a global standards organization that decides mobile technology standards. It is comprised of many working groups addressing different aspects of the 5G architecture. These standards are defined in a modular way, so that the work of different groups can proceed independently. Thousands of engineers, mainly from equipment vendors and mobile operators, participate in this process.

In particular, the 5G standards architecture relies on a clear modular separation between the 5G core network (5GC) which controls resource allocation and security between the network and the user device, and the radio access network (RAN) which is focused exclusively on managing radio communication between the base station and the user device and encrypt the signaling and user plane traffic, as instructed by the core network. The core encrypts and secures user authentication and signaling traffic across the entire network, and it manages the setting up and tearing down of calls and data connections in response to user requests.

The RAN simply hauls that encrypted signaling between the device and core network, protects traffic across the radio interface using its own algorithms, and protects signaling and data traffic between the core network and device with the aid of a security gateway. To ensure security, the inbound and outbound traffic between RAN and core must encrypted using a Security Gateway (SGw). The SGw is an essential component – beyond the scope of the 3GPP standards – of the 5G network security architecture, and needs to be deployed by mobile operators to protect their security control zone.

The 5G core network (5GC) also has a modular architecture. The main architectural split within the core is between a control plane and user plane. The 5GC is a "service-based architecture" that implements key architectural components in a virtualized and cloud-native manner. From a security perspective, the important thing is that the control plane has exclusive access to the user information.

In contrast, the RAN is confined to the user plane. The service provided by the RAN is data transport between user devices and core network. Unlike the 5G core, it incorporates no control functions associated with user data management, only internal control to manage radio access or direct transfer of encrypted signaling to the 5G core.

The RAN interacts with the core control plane through a standardized interface. It also interacts with the core user plane through another standardized interface. But user devices like smartphones and IoT devices are controlled through their own separate interface with the core, transparent to the RAN. The role of the RAN is simply transfer signaling messages between UE and the core.

The clarity of the separation of core and RAN technology has also been a feature of 4G and earlier generations of mobile technology. It has great advantages:

- It makes standards development much easier. The existence of a stable, well-understood relationship between core and RAN means that standards developers can upgrade core technology without worrying whether it will affect RAN operation, and vice versa. This simplifies the standards development task. It allows the radio technology specialists to focus on what they do best – the RAN - while allowing core technology specialists to maintain their own focus on the core.
- It makes it possible for operators to take multi-vendor approaches to sourcing. The standardized interface between mobile core and RAN means that operators can choose different vendors for their core and RAN technology, in the knowledge that the pieces can be easily integrated. They can also choose multiple operators to supply their RAN equipment as well. This ensures that operators have supply alternatives and increases competitive pressure on vendors. Limiting the field to just one or two, would increase over-dependence and reduce competition, resulting in less resilience and lower security standards. Therefore including a third company will result in higher overall security.

From a security perspective, this separation means that RAN operation cannot affect core security protocols. The RAN is the “idiot savant” of the 5G mobile network. It is brilliant at transferring radio data between user devices and the core, but it does little else. The focus of 5G security concerns is therefore the 5G core.

Operators exploiting RAN/core separation

The benefits of core/RAN separation and the flexibility this gives mobile operators are not theoretical. Many operators have multi-vendor mobile networks, and 5G networks will be no exception. In Australia, Optus has operated a multi-vendor 4G network, using both Nokia and Huawei RAN equipment. Similar modular approaches to 5G implementation have already emerged:

- In the United Kingdom, Vodafone, 3, BT/EE and O2 have all announced they will use Huawei 5G RAN with other vendor’s core networks. 3 is using Nokia core technology. Vodafone will use an Ericsson core network and also uses some Ericsson RAN equipment; Vodafone and O2 have agreed to share 5G infrastructure in the future. EE has not announced a core network supplier yet but has ruled out Huawei as a core network supplier.

- In the ANZ region, Spark in New Zealand has demonstrated multi-vendor operation between Huawei RAN equipment and a Cisco core network in its Auckland laboratory. The network incorporated both mid-band and millimeter wave technology.

Globally, 26 commercial 5G network had been launched as of July 2019. Of those 26, a significant majority (17) were using Huawei RAN equipment, though not necessarily exclusively. However, the number using Huawei's core network was only around half of this (9). The remainder had successfully integrated Huawei RAN with other vendors' core technology.

The competitive benefits of multi-vendor operation are obvious, but they are particularly important for the RAN. RAN equipment supply is less competitive than core equipment supply because RAN supply has become more concentrated in recent years. The UK House of Commons Science and Technology Committee has recently remarked (during its inquiry into network security) that restricting vendors in RAN market will further lessen market competition, and reduce operator leverage over vendors. The Committee noted that this could have the perverse effect of reducing vendor incentives to maximize network security.

Will core and RAN blur in the future?

Concerns have been raised that the architectural separation between core and RAN might be weakened in the future. This would mean that the clarity and stability of the interface between core and RAN would be lost. However, there are strong reasons to believe this will not happen.

The first reason is technical. Core/RAN separation is fundamental to the architecture of 5G standards, as it is for previous generations of mobile technology. Standards development would become much more complex and difficult, slowing down innovation. Weakening this separation would be a major U-turn that has no support in either the vendor or the operator community. And because 5G standards in fact implement core/RAN separation, any mobile technology that broke with this separation would not be 5G and would not be compatible with 5G networks (or any earlier generation of network technology).

The second and more powerful reason is that both operators and vendors would suffer if this separation were to be relaxed.

- Mobile operators that tried to implement networks without a clear core/RAN separation it would difficult to adopt multi-vendor solutions, because it would no longer be possible to guarantee that different core and RAN equipment could be connected in a simple, modular way. Operators would be forced to adopt a single vendor and stick with them indefinitely. This would ease competitive pressure on vendors, and push up operator costs.
- While it might seem that vendors would welcome this, they would still be competing with other vendors offering the standard 5G solution that maintained a clear core/RAN separation. A technology offer that failed to maintain the advantages of core/RAN separation would not be competitive, and not achieve any economy of scale. Even if an operator were to demand such a bespoke and non-standard solution, the additional costs of implementing could not be amortized over the rest of the industry so it would be expensive. No rational vendor would put themselves in this position.

The third reason is because security within a network that has been built to be resilient to attack, such that no single action could disable the system, can be best achieved by diversifying suppliers. The arguments for this are two-fold:

- Reducing over-dependence and increasing competition. First, the network should not be dependent on just one vendor, as this would render it less resilient.
- Secondly, using equipment from more than one vendor increases competition between those vendors. This will force them to improve their security standards. And it is this raising of the bar on cyber security standards across the board that is needed - together with a requirement for more stringent regulation and enforcement of those standards.

The idea of the core/RAN separation being eroded is therefore something of a mirage. These concerns seem to persist largely because of confusion between the RAN (which is a radio access technology) and “multi-access edge computing” on the network (which involves running applications on core network infrastructure). Although the RAN is obviously located at the physical edge of the network, this does not mean that it has anything to do with “MEC”. To see why, it is necessary to understand how multi-access edge computing works and why it is being considered. The idea of the network “edge” needs to be clearly defined to avoid confusion about the difference between edge computing and the RAN. *In fact, 5G multi-access edge computing raises no significant security issues, being defined and implemented as an Application Function (AF) of the 5G core network*.

Edge computing and the RAN

What is edge computing?

Modern telecommunications networks do not simply provide connectivity. They can also host consumer and enterprise applications. A simple example of this is content caching, where content is stored in multiple physical locations around the network. Moving this content closer to users reduces latency and network congestion.

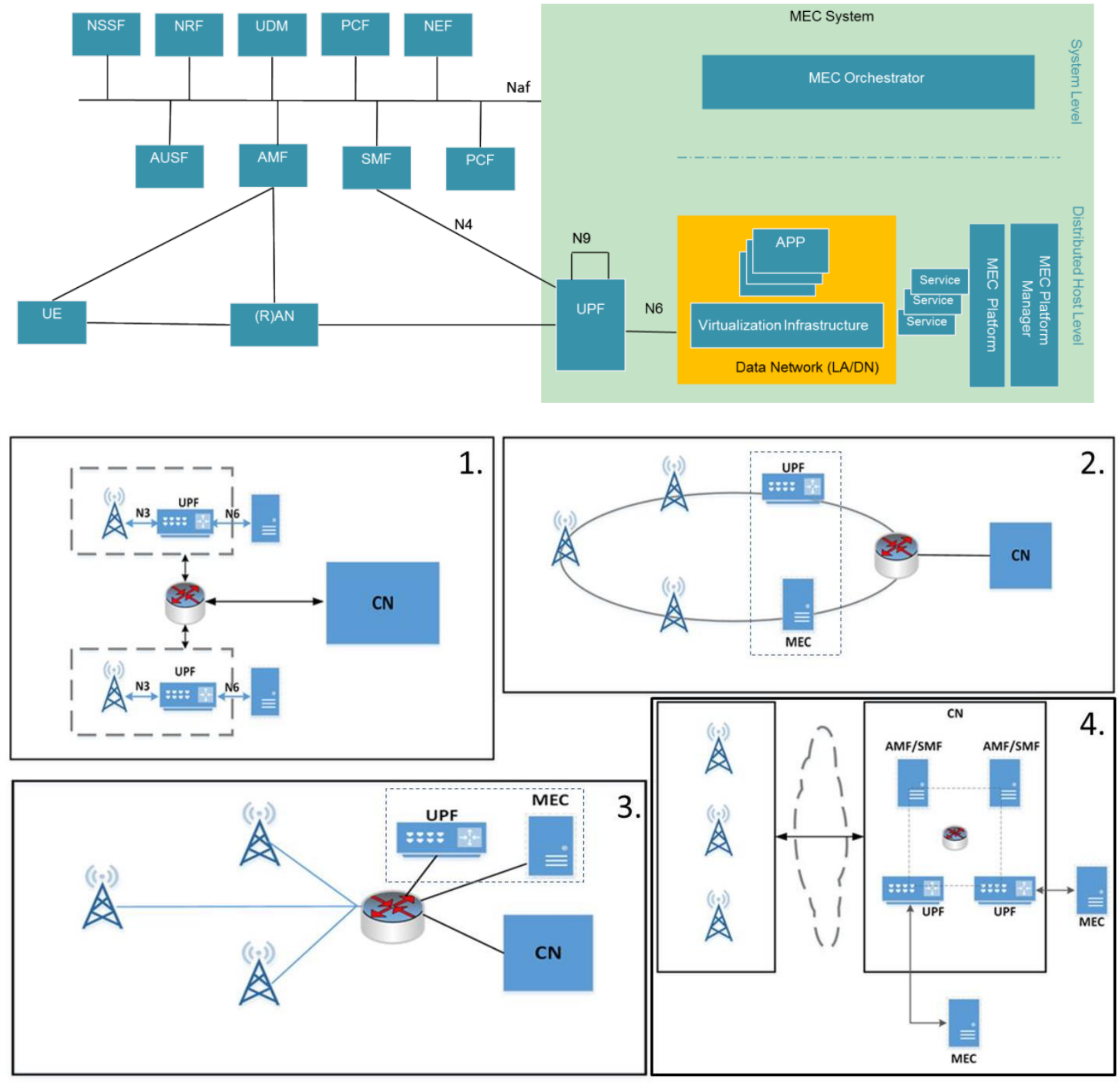
5G networks will be able to host much more sophisticated applications such as autonomous vehicles control, industrial automation, and virtual reality media. Their performance can be enhanced by placing compute, memory and storage resources closer to users. This approach is call “edge computing”.

The available computing, memory and storage capacity on the 5G network will be more constrained the closer to the user that resources are placed, and this means that different applications will require different approaches to distributing resources. Applications that require the caching or processing of large amounts of data at the edge may be better sited at a central office or metro aggregation point. Alternatively, simpler applications targeted to a particular user might be placed closer to that user, at a 5G base station or on customer premises.

The placement of the multi-access edge computing host (MEC applications and platform) is proving to be a key issue, particularly with regard to 5G. In the enterprise context, the "edge" can even refer to the IoT device or a gateway located at the customer premises. In use cases such as virtual reality or autonomous vehicles, low latency can be critical, and the overriding consideration will be one of proximity to the point of service delivery. In such instances, the base station might provide the optimum location.

Use cases where a higher level of latency can be tolerated, but where the demands of local content caching or data processing require more capacity, can potentially be supported further back in the network, such as at a central office or aggregation point (see Figure 1). Common to all deployments is that they are controlled by the core network.

Figure 1: Integrated MEC deployment in 5G network



Source: ETSI

- MEC and the local User Plane Function collocated with the Base Station.
- MEC co-located with a transmission node, possibly with a local UPF.

- MEC and the local UPF collocated with a network aggregation point.
- MEC co-located with the Core Network functions (i.e. in the same data center).

The options presented above show that MEC can be flexibly deployed in different locations from near the base station to the central data network. *Common to all deployments is the user plane function (UPF) of the 5G core that is deployed and used to steer the traffic towards the targeted MEC applications and towards the network*.

Edge computing security and the RAN

First and foremost, it is important to understand that edge computing functions use services and information offered by the 5G core functions. The term "edge" simply means that core resources (virtual network functions) are distributed physically across the core network, closer to users, rather than being centralized nationally. In theory, these resources could be co-located with RAN equipment in a single cabinet. But as part of the core network, they remain subject to core security protocols. Moving compute, memory and storage resources physically closer to the user or the RAN does not change this. All MEC applications remain enveloped by the core security protocols set up in the 5G standard.

In particular, they remain subject to the core/RAN separation discussed above. The security of user authentication and data communications between the core network and devices like autonomous cars, IoT sensors or smartphones is still managed by the core network's security protocols, and the RAN 's role remains exactly the same as before: to convey encrypted traffic as efficiently as possible. Multi-access edge computing gives the RAN no new access to authentication information or data traffic, and creates no new security issues apart from the physical one of securing exchange buildings and/or cabinets where edge resources are located. But again, this has nothing to do with the RAN.

This means that the security of multi-access edge computing depends on primarily on the selection of the core network vendor, and how well the operator manages the security options built into the core network technology and deploy the network security architecture to protect its security control zone. The security significance of the RAN is still limited; its role remains to convey traffic between base station and the mobile device.

Appendix

Author

David Kennedy, Practice Leader, Asia-Pacific

david.kennedy@ovum.com

Ovum Consulting

Ovum is a market-leading data, research, and consulting business focused on helping digital service providers, technology companies, and enterprise decision-makers thrive in the connected digital economy.

Through our 150 analysts worldwide, we offer expert analysis and strategic insight across the IT, telecoms, and media industries.

We create business advantage for our customers by providing actionable insight to support business planning, product development, and go-to-market initiatives.

Our unique combination of authoritative data, market analysis, and vertical industry expertise is designed to empower decision-making, helping our clients to profit from new technologies and capitalize on evolving business models.

Ovum is part of Informa Tech, a B2B information services business serving the technology, media, and telecoms sector. The Informa group is listed on the London Stock Exchange.

We hope that this analysis will help you make informed and imaginative business decisions. If you have further requirements, Ovum's consulting team may be able to help your company identify future trends and opportunities, please contact us on

https://ovum.informa.com/contact/contact-us

consulting@ovum.com

Copyright notice and disclaimer

Whilst reasonable efforts have been made to ensure that the information and content of this product was correct as at the date of first publication, neither Informa Telecoms and Media Limited nor any person engaged or employed by Informa Telecoms and Media Limited accepts any liability for any errors, omissions or other inaccuracies. Readers should independently verify any facts and figures as no liability can be accepted in this regard – readers assume full responsibility and risk accordingly for their use of such information and content.

Any views and/or opinions expressed in this product by individual authors or contributors are their personal views and/or opinions and do not necessarily reflect the views and/or opinions of Informa Telecoms and Media Limited.

CONTACT US

ovum.informa.com

askananalyst@ovum.com

INTERNATIONAL OFFICES

Beijing
Boston
Chicago
Dubai
Hong Kong
Hyderabad
Johannesburg
London
Melbourne
New York
Paris
San Francisco
Sao Paulo
Shanghai
Singapore
Sydney
Tokyo