

Mapping Success

Evolving to software-defined mobile networks

Software definition as the basis for 5G

It's become clear that a disaggregated and distributed software-defined mobile network is not only possible, but also provides tangible economic and technical benefits. The reduced time to market, reduced cost to deploy and operate, and a web-services like platform to quickly introduce new services are benefits that are all well documented in white papers, blogs, and industry articles. The Rakuten Mobile network designed by Cisco and Rakuten has been the recipient of three industry awards in 2019. However, most networks aren't built from scratch. Instead, they are existing 4G/LTE networks that are implementing or planning to implement a new 5G radio access network and 5G core network to support the new access. It's possible for existing network operators to evolve to a software-defined network and reap the same benefits operators who build an entirely new network.

Contents

- Challenges in evolving to 5G
- A service-oriented 5G network
- The need for automation
- Exposing the 5G network through edge computing
- Private radio use cases
- Mobile broadband
- Decomposition, disaggregation, and standards
- 5G RAN
- Decomposed software-defined mobile core
- Automation
- Redefining the supply chain for mobile networking infrastructure
- Evolution path
- Remain competitive
- Learn more

Challenges in evolving to 5G

Successful operators run profitable businesses that rely on revenue from attractive service products that address both the consumer and enterprise markets. Today, consumer products integrate voice and data plans, whereas enterprise products tend to focus on the needs of businesses. Increasingly, operators are emphasizing capabilities that allow other business to use the network. One example is the Internet of Things (IoT). Directly through the operator or indirectly using a connected device platform, enterprise IT managers access devices that provide benefits to the enterprise. 5G networking is designed to address service creation capabilities with features in both the radio access network known as NR or new radio and the 5G core network known as 5GC or 5G core. The core network aligns with the radio features using standardized interfaces. To unlock the best business value, you need to leverage an end-to-end software-defined approach in architecture, deployment, and procurement practice to build 5G networks.

Mobile operators today face a challenge in evolving to 5G. The promise of 5G is of a system integrating with the NR interface that is defined for a broad variety of use cases. These uses may range from consumer-oriented mobile broadband services to machine-type communications and low-latency use cases.¹ The 5G system architecture opens-up new business-to-business (B2B) and business-to-consumer (B2C) opportunities and potential new revenue. But operators need to understand the business case, including managing the operational lifecycle of the 5G network. The network should offer lower operational expenses (OpEx) and better control over capital equipment (CapEx) costs throughout the mobile supply chain. Enter the software-definition attribute of the emerging 5G solution set.

Cisco experience has shown that it's possible to build an entirely new network. Rakuten Mobile in Japan^{2,3} embraces the principles of disaggregation and decomposition. Both principles are at the heart of what it means for a network to be software-defined. It means that software is a separate purchasing decision from hardware and that commercial off-the-shelf hardware are used wherever possible. Decomposition means that the network architecture is built from modular elements that assemble into a whole through open and interoperable interfaces. By employing the principles of software disaggregation and leveraging decomposition, which in the 5G system architecture is an attribute of both the RAN and the core,⁴ Cisco and partners have shown at Rakuten Mobile that it's possible to build an optimal architecture that addresses the OpEx and CapEx challenges.

The radio signal processing stack is completely decomposed into modular elements. The cell site deploys only the analog radio frequency components along with the lower layers of the LTE stack (Lo-PHY) and is fed by fiber optic cable from a Distributed Unit (DU) deployed in cloud environment implementing the scheduling functions required in multi-access radio (Hi-PHY and MAC/RLC). An active antenna system can be used at the cell-site implementing the virtualized DU functions when necessary or required, such as for massive MIMO⁵ at mid-band frequencies or millimeter-wave (mm-wave). When appropriate, a distributed unit (DU) deployed in cloud environment can be used to implement the critical scheduling functions required in multi-access radio.

Further upstream, a centralized unit (CU) implements the non-real time functions in the radio signal processing stack which can be easily deployed in cloud virtualization environment.^{2,3} A decomposed and completely software-defined set of mobile core functions in the cloud manage enterprise and consumer subscriber services and handle policy, as well as the billing functions. In addition, the core is set up from the start to provide business-to-business (B2B) services to multiple enterprise market segments. The type of network Rakuten Mobile is building is designed to precisely align with 5G objectives. It is flexible using a modular design and fully embraces the opportunity of B2B services.

Two concepts are central to the software-defined 5G network: cloud-RAN that was pioneered by China Mobile⁶ and the 5G core (5GC), which is responsible for subscriber management and services.^{4,7}

Cloud-RAN is the idea that the radio signal processing stack can be decomposed into functions and that these functions (except the analog elements) can be deployed in a virtualized cloud environment. The analog elements are not virtualizable, and they are deployed at the cell site. Interconnecting analog and digital elements is a baseband analog-to-digital and digital-to-analog conversion function. It can vary in terms of sophistication. The analog functions are remote radio heads (RRH) or "remote units" (RU). The 5G system architecture⁴ has embraced this decomposition of the RAN. The decomposition will create new interfaces, known as "RAN splits" which will be used to interconnect cloud functions among themselves and with the cell-site functions.

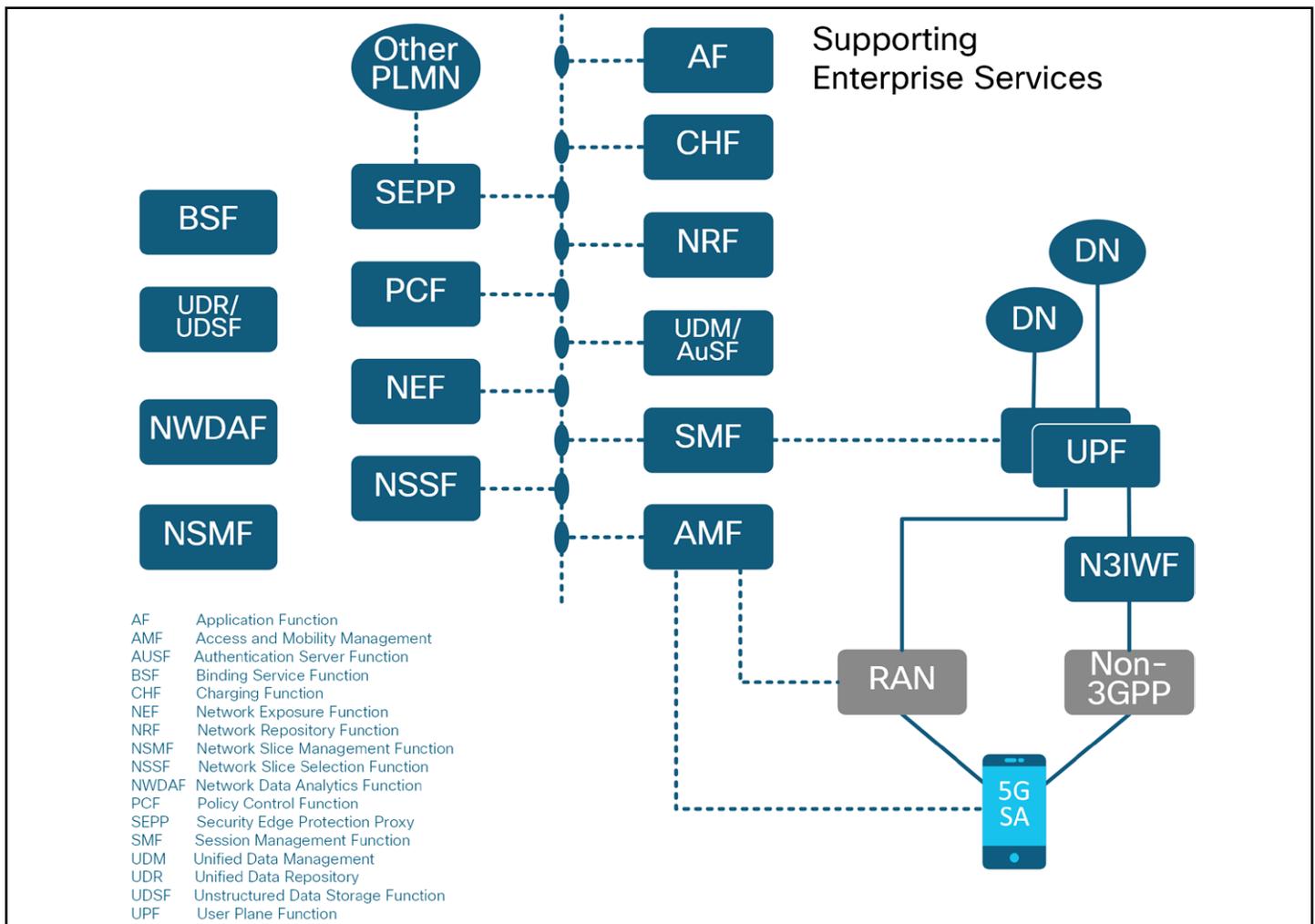
The 5G core (5GC) is designed to be modular. The decomposition, shown in Figure 1, separates the user plane functions (UPFs) from the control plane functions (everything else). The control plane functions are implemented with communications through a software bus using JSON/HTTP2 messaging, which allows mapping into a microservices-based cloud native service environment.^{8,9}

Decomposition and disaggregation are not ends-in-themselves but rather attributes of a competitive network. These two attributes will ultimately ensure the network is adaptable to the broad service needs of 5G. Some of these service needs will materialize over time and will require flexible definition to adapt to market conditions and competitive requirements.

A service-oriented 5G network

Current mobile networks are designed to primarily tackle consumer-oriented mobile broadband services (MBB) but 5G offers more. From the outset, 5G was to address operator needs to provide more value to various vertical segments, including MBB. For example, the 5G NR features flexible numerology,¹¹ which aside from providing better adaptation to spectrum bands, allows the air interface to support lower latency at the expense of spectral efficiency. Flexible numerology allows NR to be a viable solution for latency-sensitive factory automation. It demands low packet loss and low latency for low energy consuming IOT devices, as well as evolved MBB (eMBB) without being suboptimal for other uses.

Figure 1. 5G core (5GC) decomposition



Flexible numerology is matched in the 5GC with a feature known as “network slicing.” With network slicing, the network can support multiple use cases as distinct, separable, and relatively isolated business processes. A set of slices can share a numerology on the radio side and be used to define services for specific vertical segments. The NR, for example, has a mechanism for partitioning a carrier into multiple bandwidth parts (or BWP). Each bandwidth part can be associated with a different numerology. The 5GC correspondingly, through network slicing, also has mechanisms that allow business enterprises to access services defined for each BWP.

The need for automation

Emerging 5G service-oriented models position the network as an innovation platform for developers. Network capabilities are consumed using the Network-as-a-Service (NaaS) framework. With the NaaS model, the operator exposes APIs that allow the consumer to use the network as if it were their own. These APIs, in turn, drive the automatic allocation of network resources. APIs are the engine that drives B2B models and derivatives such as business-to-business-to-consumer (B2B2C) and business-to-business-to-business (B2B2B). Without an underlying end-to-end automation platform, APIs would have limited value. The end-to-end 5G network needs to be fully automated. The automation capabilities are exposed by an API gateway. This gateway performs resource allocation and securely exposes resources to third parties such as other businesses that then consume network services. Edge computing is one example of this dynamic.

Exposing the 5G network through edge computing

In edge computing, the network is used as a vehicle to allow other businesses to take advantage of the low latency services provided by the 5G network. Two factors feature prominently in defining the latency provided by the 5G network. The first factor is the IP

network that is defined between the end-point and the latency-sensitive workload, which is deployed at an appropriate location within the network. An end point may be mobile UE, IoT device, or industrial gear. The IP network contributes a delay that is associated to queuing. This delay is from an accumulation of packets on egress interfaces in routers. The queueing delay can be as high as 1,000 milliseconds (ms) in some oversubscribed networks. It’s a consequence of the oversubscription in IP networks. The bursts in bandwidth are accommodated through queuing. Oversubscription is a normal cost-cutting practice in IP networks.

Latency-sensitive services can be managed with a quality of service (QoS) solution that defines priority queuing disciplines. The packets that require low latency receive preferential treatment or even bypass the queue, such as with strict priority queuing. The second, is the MAC (Media Access Control) layer responsible for scheduling resources in the RAN. A scheduling delay in the RAN MAC layer is a consequence of the multiple access nature of the NR or LTE wireless interface. Scheduling delay can be 10 ms (or higher in LTE) for MBB services.

The origin of scheduling delay is the latency associated to the MAC layer. It assigns digital transport resources to waveforms that address multiple users in the frequency domain grid (in the downlink), or it combines data from multiple users (in the uplink). The MAC layer scheduler operates with a transmit-time interval (TTI). Periodically, multiple waveforms are scheduled, which means elements of the time-frequency grid are populated with modulation symbols (resource elements), assigned to different users in upload or download groupings called physical resource blocks (PRBs).

The flexible numerology feature of NR means the TTI can be reduced to as low as 62.5 microseconds, which also reduces the latency at the expense of a larger subcarrier spacing. In contrast to conventional wisdom, proximity to the endpoint is not the major factor in defining low latency locations. Edge nodes can be comfortably deployed in regional centers. The speed of light in optical fiber is about 5 microseconds per km

(1 ms is 20 km of fiber), and it's not meaningfully large except for larger distances. Beyond flexible numerology, NR also supports flexible scheduling of mini-slots, much faster Hybrid ARQ than LTE, frame pre-emption, and pre-allocation of resources for time-sensitive traffic. Any 5G network needs to anticipate the deployment of this arsenal of features. These features should be software changes to the NR signal processing stack, which reinforces the fact that software-definition is a competitive advantage.

A second important attribute of the edge is bandwidth reduction, which relates to performing computation at the edge to reduce bandwidth. A good example of core network bandwidth reductions is content delivery network (CDN) caching. Content is streamed from the edge without incurring core bandwidth. Access bandwidth also can be reduced by using the edge as an offload platform. For example, for virtual reality (VR) applications, the edge can compute a field of view that would otherwise have to be computed from massive amounts of data associated to all possible horizontal and vertical fields of view.

In edge computing, the operator deploys 5G core (5GC) user plane functions (UPF or UP) in strategically located parts of the network where a hosting environment for cloud computing is also deployed. The same cloud environment may well support 5G infrastructure workloads such as the upper layers of the NR signal processing stack as well as the UPF.

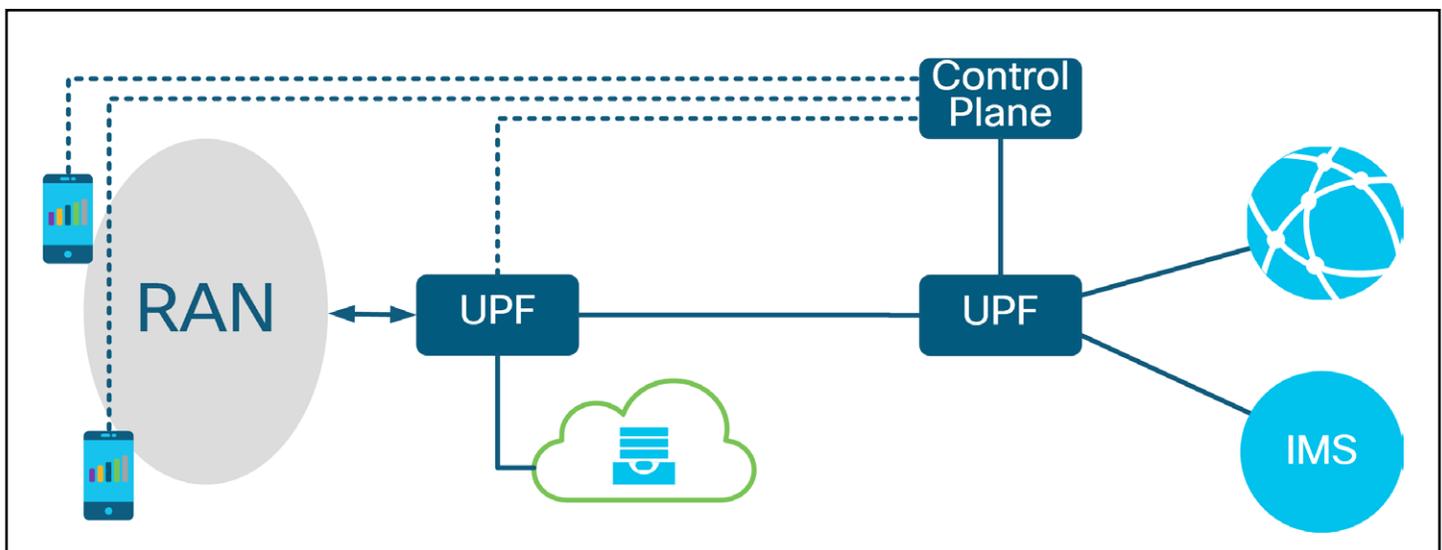
The comprehensive RAN and core feature set is useful in edge computing. Operators can expose an infrastructure that can support a broad variety of latency-sensitive and bandwidth reduction services to their business partners. API exposure will ensure that the business partner has access to the full suite of resources offered by the 5G network, whether those resources are available through signaling via QoS flows and the 5GC¹² or through configuration management (numerology features).

For operators, edge computing opens a new set of revenue-generating services such as content delivery, gaming, and augmented reality.

Private radio use cases

In a special class of use cases, the radio spectrum asset itself is assigned to an enterprise either through localized deployment at customer premises or assigned in the wide area. The enterprise IT manager may choose between 3GPP NR or the 802.11ax¹³ family of wireless LAN specifications to build connectivity. The PHY layer for both radio access methods is very similar. They both use orthogonal frequency division multiple access (OFDMA) in upload and download. They also both support flexible numerology and low latency services and a QoS architecture. However, 802.11ax operates in unlicensed spectrum and is focused on localized environments whereas NR currently deploys in licensed bands and is optimal in the wider area macro-cellular network. The 802.11ax specification will be more friendly to enterprises that choose to deploy by

Figure 2. Edge computing as supported by the 5GC



themselves, within the enterprise, independent of an operator. The 5G solution based on NR will be most relevant to enterprises that are concerned about an interference-prone radio environment. Because the 5G system architecture is defined for massive scale, a private NR solution will likely be delivered as a hosted managed service.

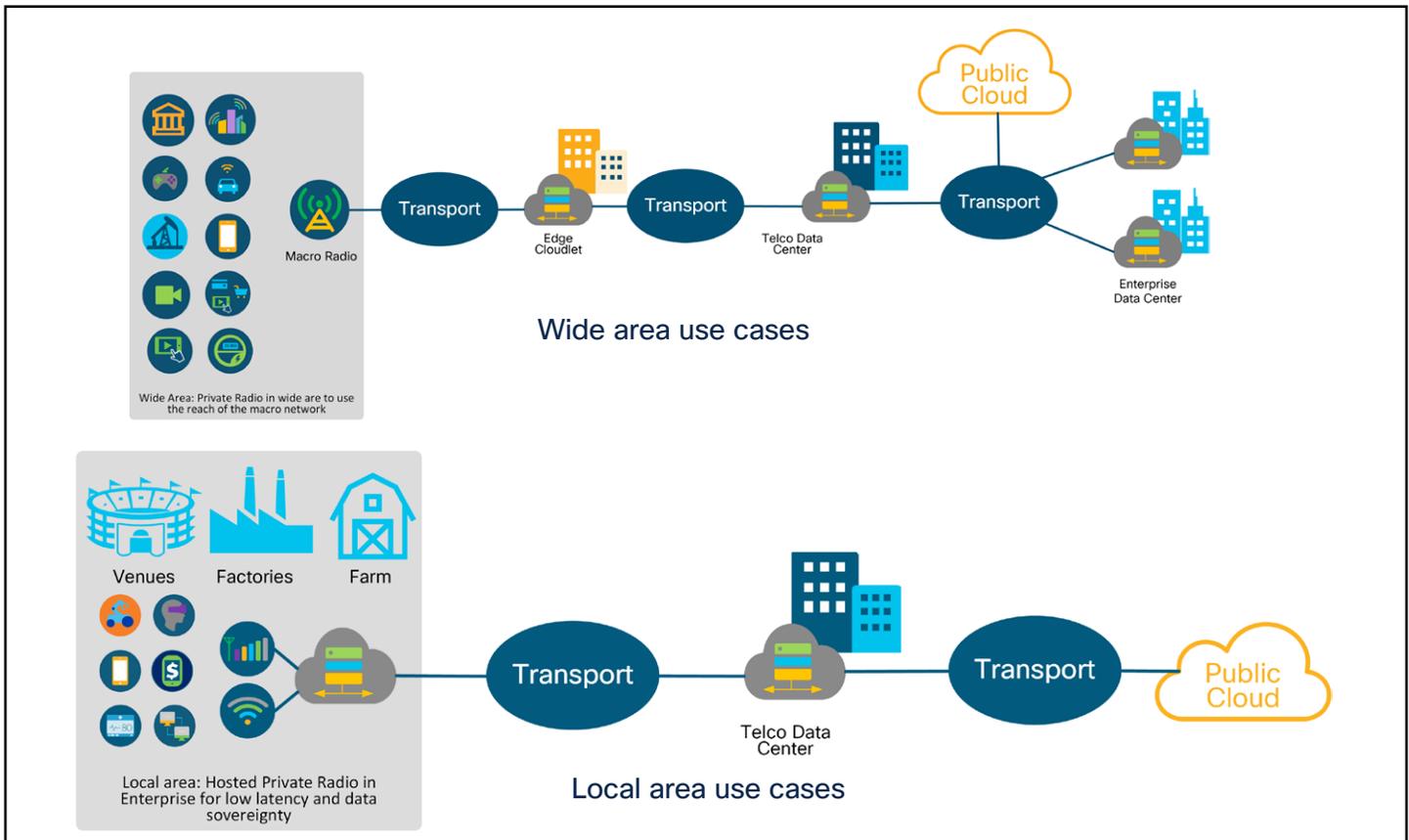
We anticipate Industry 4.0 or the factory of the future.¹⁴ The ability to reconfigure a production line without having to uproot wired connections is valuable to manufacturers. In these use cases, a robotic arm or automatically guided vehicle is equipped with a wireless interface for operation.

Electrical grids or gas and oil fields require a different type of solution because they extend over a wide area and require a high degree of reliability for optimal operation. The 5G network can be configured as a service to electrical utilities to allow them to automatically control electrical grids reliably or as a

service to production operators to monitor and secure oil and gas fields for proper performance.

It's worth emphasizing the difference between the use cases. In the Industry 4.0 use case, the radio is deployed in localized fashion within the factory floor. It isn't desirable for the licensed radio emissions to leak from the inside out. In electrical and oil and gas field use cases, the radio used spans a wide area. The macro network covers an oil and gas field or an electrical grid. In some cases, the radio spectrum needs to be dedicated. But often it can be shared for a lower cost. The question is whether the radio slice can be shared or not. And can a data termination be in the form of a 5GC UPF? Or should it be deployed at the enterprise location alongside a cloud hosting environment? This type of use case is also known as enterprise multi-access edge computing (MEC). Figure 3 illustrates both scenarios. The wide area use cases are instances of the enterprise facing out, whereas the local area use cases are examples of the enterprise facing in.

Figure 3. Two different types of private radio use cases



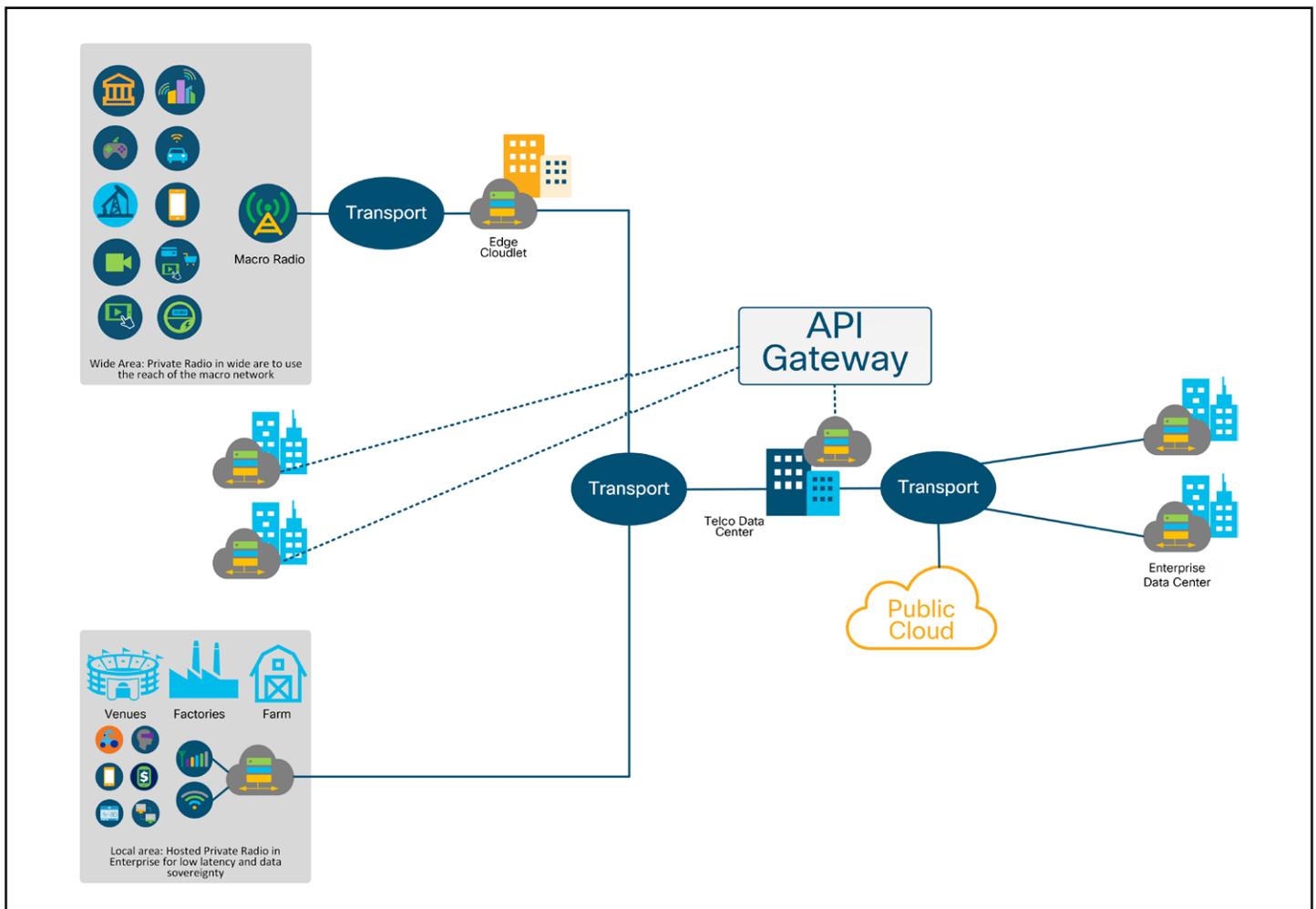
In both cases, the enterprise needs to manage network resources as if they were own, which is a hallmark of a network-as-a-service capability. To support NaaS, an application function is introduced on the 5GC, operating as an API gateway as shown in Figure 4.

To address this situation, Cisco is introducing a solution for API control of private radio solutions called the Cisco Unified Domain Center.¹⁵ With the API gateway, an operator can offer a hosted managed service to enterprises with a service definition that is suitable for that vertical market. The resources are properly isolated, so they don't interfere with services of other customers or interfere with the proper operation of the platform itself.

Mobile broadband

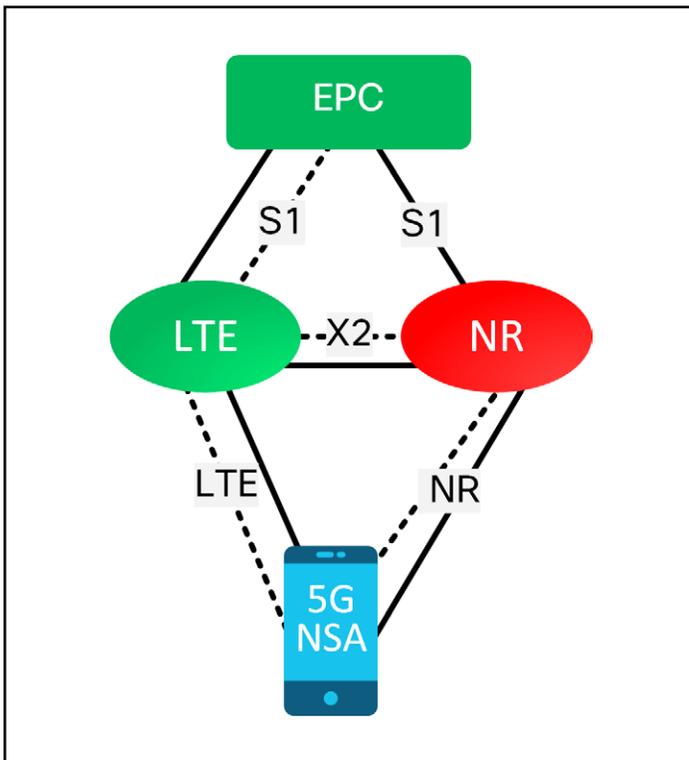
The 5G system is also anticipated to work well for mobile broadband where LTE is already an established system solution. Initial deployments of NR will tightly interwork with LTE through a mechanism called EUTRAN-NR Dual Connectivity (or ENDC) as shown in Figure 5.¹⁶ An eNB supporting LTE radio primary node provides the control plane interfaces into an evolved packet core (EPC). Devices register with a mobility management entity (MME), which supports mobility and session management. The conventional NAS signaling between devices and the core is used.¹⁷ But the devices are multi-radio capable, which means they support NR and LTE radio interfaces concurrently. These devices

Figure 4. API control using APIs of private radio solutions



aren't the conventional LTE devices used today. They are devices equipped with a modem chipset, which can handle both radios concurrently.

Figure 5. ENDC scenario. The NR is deployed as secondary node to an LTE primary node. The devices support both NR and LTE though register into an MME in the EPC.



Although this approach doesn't allow an operator to capture the benefits of the complete 5G system architecture, it's a significant 5G application. The architecture that delivers this use case is called "non-stand alone" or NSA as opposed to "stand alone" or SA. Multi-radio NSA UE devices can provide an immediate benefit to the user with access to more bandwidth that combines LTE and NR for improved service. The benefit to the operator is the simplicity of the upgrade since the element delivering NR, a gNB, is subordinate to the existing eNBs. However, NSA ecosystem devices may not be upgradeable to the SA ecosystem for 5G where the registration is using 5GC.

Introduction of NSA is popular because of the opening of the low band digital dividend spectrum, C-Band outside of the US, and millimeter wave (mm-wave). Existing LTE can be deployed alongside NR, which would then use the newly available spectrum. Although the NSA ecosystem does little for expansive 5G use cases, operators still should implement a 5GC as a software upgrade from an EPC.

Decomposition, disaggregation, and standards

The 5G systems architecture defines decomposition architectures, which can be used to provide flexibility. For example, in the 5GC, the control plane functions can be functionally decomposed from the user plane functions. The user plane functions can be placed at low latency locations with respect to the endpoints, which resulting in an edge computing (or MEC) solution.¹⁸

5G RAN

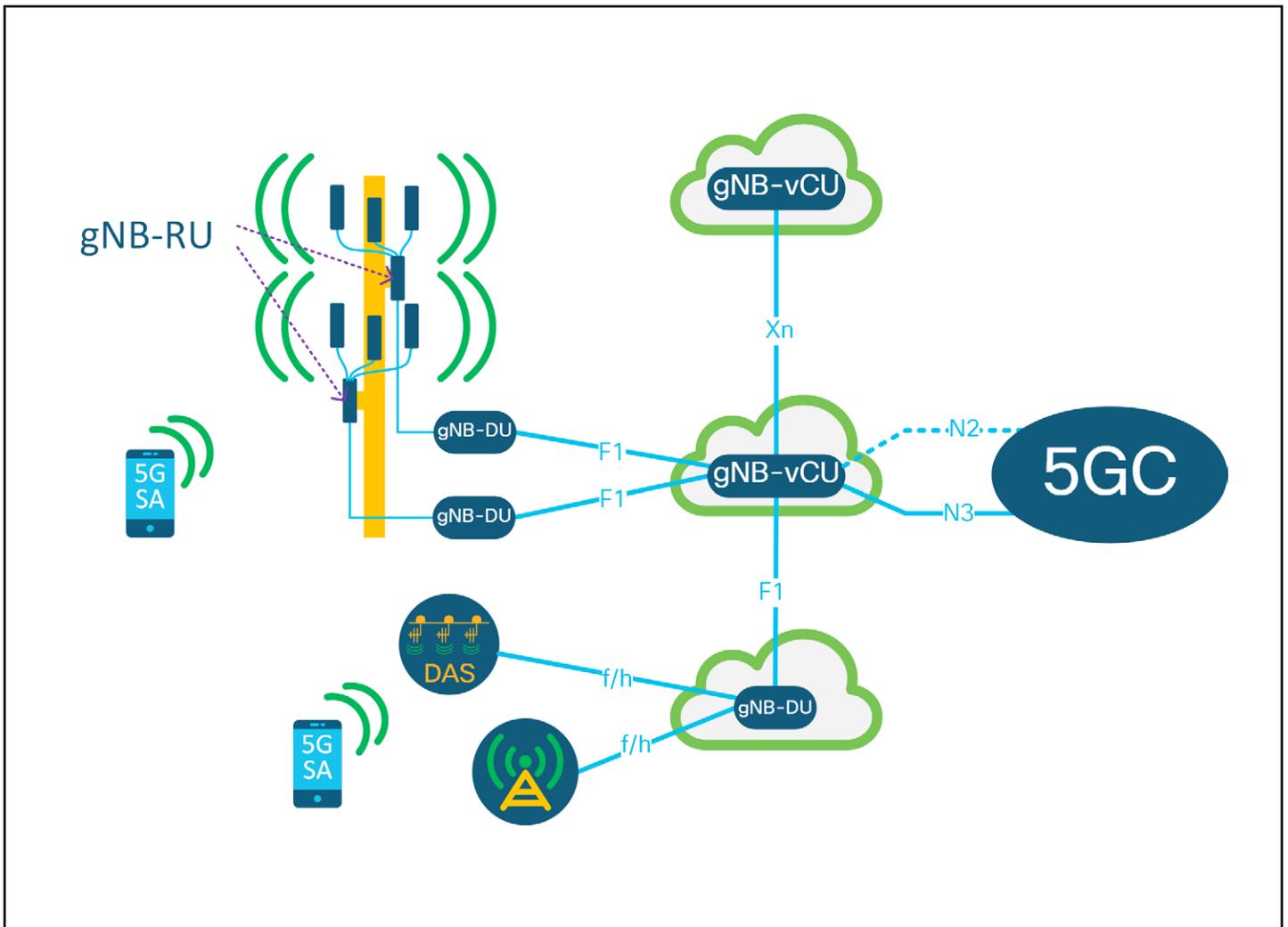
In RAN decomposition, elements of the LTE and NR signal processing stacks can be separated into distinct functions allowing the higher layers to be deployed centrally while the lower layers are pushed out to the edges. When the signal processing is deployed in a virtualized environment, you have what is commonly referred to as cloud-RAN. Because the system is characterized by decomposition and disaggregation attributes the RAN architecture inherently supports multi-vendor RAN, which means a RAN network that is comprised of products from multiple vendors. The software, the hardware, the transport, the management and orchestration systems, and the provisioning systems from various vendors can be modularly replaced as needed. RAN decomposition heralds a transition to the cloud for demanding radio real-time signal processing workloads.^{2,3} In decomposition, operators must demand adherence to interoperability among functional entities (see Figure 6). The 3GPP specification also has more

details.⁴ Separating the DU and the RU is the front haul domain transporting digitized NR or LTE waveforms. An interoperable specification exists for the front-haul transport in O-RAN which covers user, control, management, and synchronization planes. It's based on split 7-2x, which is essentially digitized frequency-domain samples of the NR baseband carried over Ethernet.^{19,20}

Decomposed software-defined mobile core

The 5GC needs to be the enabling platform for an operator to take advantage of the major changes taking place in the data center, networking, and the economics of mobility in a standardized multivendor environment. The 5GC introduces significant changes for

Figure 6. Decomposition in 5G RAN



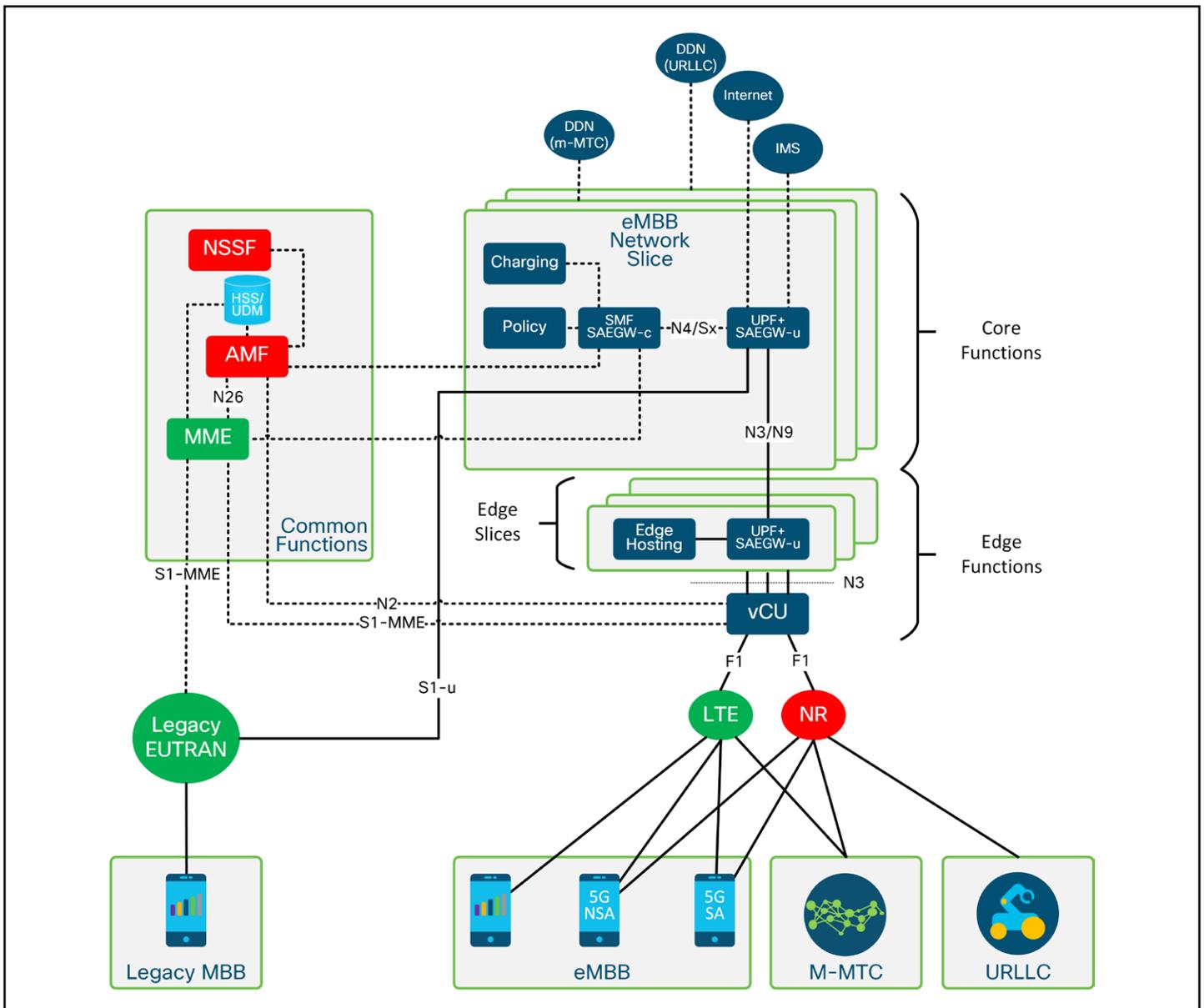
the mobile core that facilitate new opportunities such as personalized networks through slicing and more granular functions are being defined.

The Control, User, and Synchronization Plane Specification (ORAN-WG4.CUS.0-v02.00) from the O-RAN Alliance provides a complete description of the 5GC.^{4,21} The Cisco implementation is the Cisco Ultra 5G Packet Core Solution.¹⁰ Figure 7 shows the basic aspect of implementation and evolution with an almost inevitable scenario when 5G is introduced. The device ecosystem hasn't evolved to the point of fully supporting SA, which

means the state where all devices can register with a 5GC and deriving services defined through it. In that case, an EPC is necessary.

The figure also shows an efficient implementation. The EPC is shown integrated with the 5GC at least, for MB. User plane and control plane functions for mobile broadband each play a dual role: a 5GC UPF is also a System Architecture Evolution (SAE) gateway (SAEGW-u) and a 5GC session management function (SMF) is also an SAEGW-c. The remaining slice functions do not necessarily need to align with the EPC.

Figure 7. Core evolution to support 5GC and EPC.



For MBB, it's useful to allow core network interworking. In core network interworking a common anchor, in this case a combined UPF/SAEGW-u, terminates the RAN connections from both LTE and NR. In that way, IP addresses are preserved as a mobile device moves between an LTE and NR coverage area. In this type of interworking, the device "reselects" one radio technology when the other one isn't available. Reselection can take a few seconds where service is interrupted. An interface, called the N26 is optional – but also highly desirable – between the legacy MME and the 5GC access and mobility management function (AMF) to manage state transfer between the two domains.

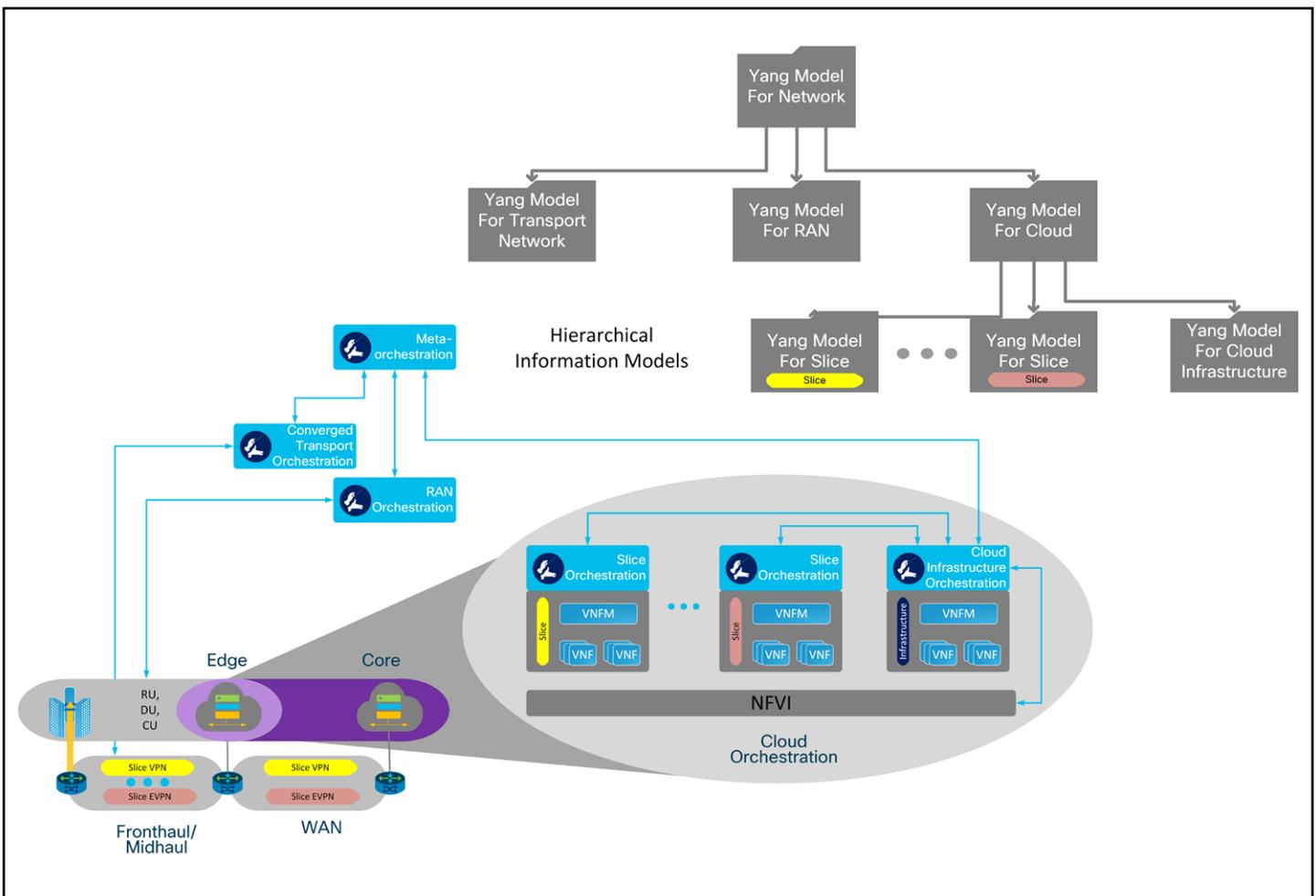
With a software defined EPC, migration to the 5GC and the 5G systems architecture is much simpler. Existing assets can be preserved while new ones can be introduced following the optimal 5G strategy.

Automation

Given the need for automation to build an effective 5G service model based on B2B and NaaS, various industry players (including Cisco and the O-RAN alliance) are advocating a model-driven approach. At the highest level, a model driven approach relies on a network model for a network function or a network domain. It's usually described in human-readable form using a Yang definition.²² Once a model is available, it can be consumed using the network configuration manager system which can then use Netconf²³ to configure either the network function directly or a network domain.

Yang models are stackable. This observation is key to the scalability of Netconf/Yang driven resource configuration management (sometimes called orchestration). A domain such as a RAN cloud, a transport network, or

Figure 8. End-to-end automation in 5G



cell-site equipment can have its own Yang model, which effectively hides information for the lower-level model behind a domain orchestrator.

Yang models are also useful in implementing NFV clouds²⁴ because they can be consumed by Network Function Virtualization Orchestrators (NFVO). In addition, Yang models are extremely useful in managing containerized functions and even Kubernetes clusters in the cloud-native deployments that use Docker⁸ and Kubernetes⁹. A meta-orchestrator can then align resources end-to-end across multiple domains as shown in Figure 8.

Cisco is successfully deploying the Netconf/Yang automation architecture using Cisco Network Services Orchestrator.²⁵

Redefining the supply chain for mobile networking infrastructure

For operators, a 5G system architecture isn't business as usual because 5G is transformational. The need to lower costs means the transformation must start with changes in the global mobile supply chain.

Consider the situation today. A few vendors dominate the RAN market through a "whole ecosystem" approach to infrastructure delivery. In the prevailing approach, the mobile network is delivered through monolithic RAN architectures that often lack standard interfaces that would inherently prevent any kind of multivendor decomposition within the RAN. The vendor supplies a comprehensive set of services that include project and program management and systems integration. Because these vendors often supply a mobile core, they also make the case that their RAN and core work better than a multivendor configuration. They may even offer commercial incentives for aligned procurement of RAN and core. These few vendors dominate the totality of the

RAN procurement and aspire to do the same for core networks.

To maintain a semblance of a multivendor network, operators assign regions (markets) to individual vendors. The net effect is that within a market, only one RAN vendor is deployed, and the assigned vendor controls the deployment. Yet, at the same time, the operator needs to harmonize features and functionality across markets to protect their brand. The effect is that the operator is continually forced to adopt a reduced common set of features across the multiple vendors instead of the features that could add significant value from a competitive point-of-view. New service introduction requires coordination across the multiple vendors. This coordination is made more difficult by the competing interests among them.

The current supply chain practice for mobile network infrastructure won't work in 5G evolution because the success of the systems architecture with B2B and B2C requires the ability to add features more quickly and more adaptability than prior generations. For operators, achieving vendor diversity in their 5G mobile networks is essential. To achieve these goals and embrace the principles of modular decomposition, networks must include software and hardware disaggregation. This practice will enable end-to-end automation, which is required for network-as-a-service. Ensuring vendor competition for each modular component will drive down CapEx and make the network easier to operate.

To build this supply chain, operators must work with standards. Simple adherence to 3GPP isn't enough. Extensions and interoperability profiles as defined in the ORAN Alliance are becoming more relevant. The net outcome of the redefined supply chain is a strategy in which operators focus on the best possible value using best-in-class modular components that together create the best and most adaptable solution.

Evolution path

To deploy 5G, operators need licensed spectrum. To date, growth opportunities in 5G have effectively been overlays to the existing LTE network that use dual connectivity (ENDC) and non-standalone (NSA) access that re-uses the EPC.^{4,16} Given an immature 5G UE ecosystem, the reasonable approach is to deploy 5G NR on low-band digital dividend spectrum such as spectrum made available from the shutdown of broadcast TV. Another option is new mm-wave spectrum above 26 GHz, or mid-band frequencies at C-band typically available outside the US but also being requested by operators in the US.

To achieve the business objectives for 5G, operators must craft an optimal evolutionary path that is customized to spectrum holdings and the markets served. All the spending should be consistent with achieving their objectives. For operators, a broad ecosystem approach to building a harmonized mobile network selecting from best-in-class components is the best approach.

A cutover of an entire network to 5G isn't practical. Investments must be protected during their depreciation cycle, so evolution must be gradual with incremental steps. These evolutionary steps shouldn't be disruptive to the user quality of experience. The increments also should add measurable value in terms of operational and capital expenses and create incremental revenue.

At the highest level, there are basically two approaches that can be pursued. The first is based on gradually evolving each cell-site by migrating carriers or adding carrier to the Cloud-RAN system. The second approach is based on the idea of migrating an entire cell-site one at a time in a complete refurbishment approach.

For the first approach of sector-by-sector evolution, our proposed strategy is based on deploying 5G as an overlay to an existing deployment. The initial overlay can be accomplished in several ways.

- Method one is to overlay cloud-RAN NR with existing LTE and leverage an inter-operable X2 reference point^{26,27} as defined by the O-RAN alliance. LTE and NR can coexist using the Non-Stand Alone (NSA) architecture.
- Method two involves upgrades to the existing LTE to cloud-RAN. No interoperability across X2 is required but it is required to upgrade LTE and then to ensure the upgrade works in conjunction with other LTE and pre-LTE bands that may be deployed.
- In method three, NR is deployed as a pure overlay. This approach is akin to the way LTE was deployed over existing 2G and 3G. It relies on core network interworking, which is a mobile core that can support both 5G NR stand-alone architecture devices using 5G signaling into the core and LTE only devices, which register with an LTE core.

Figure 9 shows the alignment between RAN and core in a software-defined evolutionary strategy. Figure 9 (a) illustrates ENDC to introduce 5G NR but keeps the EPC mobile core. Devices supported are standard LTE UEs as well as NSA UEs with the ENDC capability. To move beyond the initial first step, it will still be necessary to deploy a 5GC and introduce SA ecosystem devices while continuing support of LTE and NSA devices. In fact, many networks will still have 2G and 3G deployed. To achieve the objectives of supporting the plethora of different ecosystem devices, it will be necessary to deploy a 5GC interworked with an EPC as shown in Figure 9(b) and in greater detail in Figure 7. Interworking will require common anchor functions to UEs for service continuity when a device moves from a region where NR is available to one where only LTE is and vice versa.

The second approach, based on migrating a one cell-site at a time, has several important advantages over the

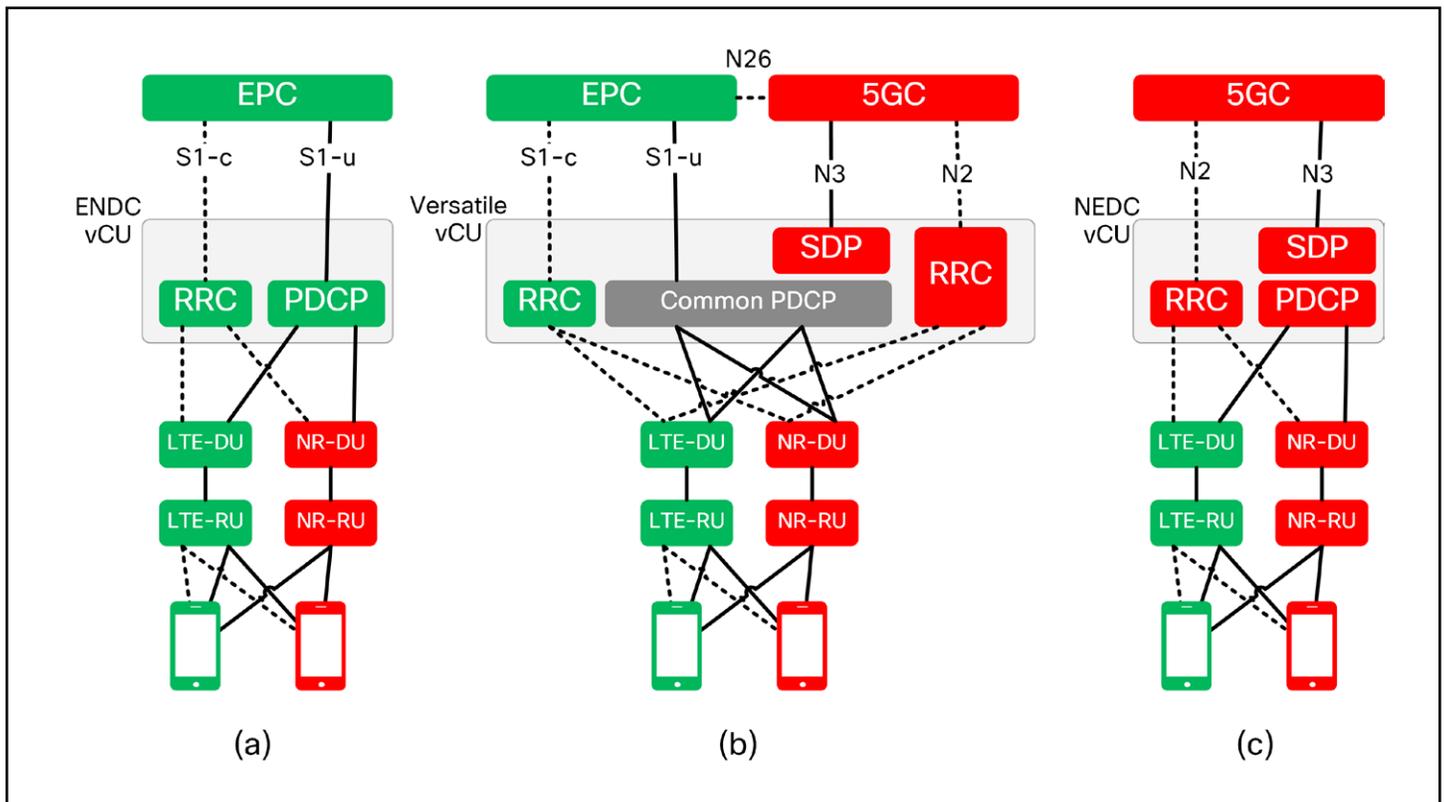
first approach. The most important advantage is that it doesn't rely on interface interworking with the existing network. The second approach also is easiest on OSS migration because operators can cap investment in the legacy OSS while enabling investment in a more suitable environment.

More can be said about network evolution towards 5G, and a comparison of the approaches will be discussed in a future white paper.

Remain competitive

Software definition of the end-to-end 5G network results in significant benefits to operators. It allows them to derive better advantage in service creation. In essence, software definition is the best way to progressively introduce the features that make an operator competitive. Software definition is well aligned

Figure 9. Evolutionary path with cloud-RAN and mobile core alignment



Learn more

Unlock the promise of 5G at
www.cisco.com/go/5g

Find out more about Cisco 5G
Mobility solutions at
www.cisco.com/go/mobile

with procurement practices particularly for the RAN. With software definition, operators can build a best-in-class infrastructure using offerings from multiple vendors. Cisco sees the opportunity for a new mobile infrastructure supply chain to emerge with 5G, which will result in a more cost-efficient, more vendor-diverse, and more attractive way of building networks.

Through it all, the 5G mobile core (5GC) remains the heart of service creation. It must be built modularly and with a strong emphasis on software definition.

1. TS 22.261, Service requirements for the 5G system; Stage 1 (Release 16), 3rd Generation Partnership Project (3GPP), 2019. [Online]. Available: <http://www.3gpp.org/DynaReport/22261.htm>
2. Cisco/AltioStar/Rakuten. "Reimagining the End-to-End Mobile Network in the 5G Era." <https://www.cisco.com/c/en/us/solutions/service-provider/mobile-internet/reimagining-mobile-network.html> (accessed October 2019).
3. Santanu Dasgupta. "Enabling Rakuten Cloud Platform with Cisco NFVI and Orchestration Solutions." <https://blogs.cisco.com/sp/enabling-rakuten-cloud-platform-with-cisco-nfvi-and-orchestration-solutions> (accessed October 2019).
4. TS 23.501, System Architecture for the 5G System (Release 15), 3rd Generation Partnership Project (3GPP). [Online]. Available: <http://www.3gpp.org/DynaReport/23501.htm>
5. E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," IEEE Communications Magazine, vol. 52, no. 2, pp. 186-195, 2014, doi: 10.1109/MCOM.2014.6736761.
6. China Mobile Research Institute. "C-RAN White Paper: The Road Towards Green RAN." <http://labs.chinamobile.com/cran/> (accessed 2019).
7. 3rd Generation Partnership Project (3GPP). "TR 23.799, Study on Architecture for Next Generation System (Release 14)." <http://www.3gpp.org/DynaReport/23799.htm>
8. "What is Docker?" www.docker.com/what-docker (accessed August, 2019).
9. "Kubernetes: Production-Grade Container Orchestration." <http://kubernetes.io> (accessed August, 2019).
10. Cisco Systems Inc. "Cisco Ultra 5G Packet Core Platform." https://www.cisco.com/c/m/en_us/network-intelligence/service-provider/digital-transformation/ultra-5g-packet-core-solution.html (accessed October, 2019).
11. TS 38.300, NR and NG-RAN Overall Description, 3rd Generation Partnership Project (3GPP). [Online]. Available: <http://www.3gpp.org/DynaReport/38300.htm>
12. TS 23.503, Policy and charging control framework for the 5G System (5GS); Stage 2 (Release 15), 3rd Generation Partnership Project (3GPP). [Online]. Available: <http://www.3gpp.org/DynaReport/23503.htm>
13. "IEEE Draft Standard for Information Technology -- Telecommunications and Information Exchange Between Systems Local and Metropolitan Area Networks -- Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment Enhancements for High Efficiency WLAN," IEEE P802.11ax/D4.0, February 2019, pp. 1-746, 2019.
14. Cisco Systems Inc. "Between an idea and production, there's a bridge." https://www.cisco.com/c/m/en_ca/digital-manufacturing/index.html (accessed 2019).
15. Cisco Systems Inc. "Cisco Unified Domain Center." <https://www.cisco.com/c/en/us/solutions/service-provider/mobile-internet/the-cisco-mobile-cloud-control-portal-for-enterprise.html> (accessed October, 2019).
16. 3rd Generation Partnership Project (3GPP). "TS 37.340, Multi-connectivity; Stage 2 (Release 15)." <http://www.3gpp.org/DynaReport/37340.htm> (accessed 2019).
17. TS 24.301 V9.11.0 (2013-03), Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS); Stage 3, 3rd Generation Partnership Project (3GPP). [Online]. Available: <http://www.3gpp.org/dynareport/24301.htm>
18. Cisco Systems Inc. "Multi-Access Edge Computing At-a-Glance." <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/ultra-services-platform/at-a-glance-c45-741450.html> (accessed November, 2019).
19. O-RAN Alliance. "Management Plane Specification (ORAN-WG4.MP.0-v01.00)." <https://www.o-ran.org/> (accessed 2019).
20. Control, User and Synchronization Plane Specification (ORAN-WG4.CUS.0-v02.00), O-RAN Alliance, 2019. [Online]. Available: www.o-ran.org
21. 3rd Generation Partnership Project (3GPP). "TS 23.502, Procedures for the 5G System (Release 15)." <http://www.3gpp.org/DynaReport/23502.htm>
22. RFC 6020, YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF), M. Bjorklund, 2010.
23. RFC6241, Network Configuration Protocol (NETCONF), Internet Engineering Task Force (IETF), 2011. [Online]. Available: <https://tools.ietf.org/html/rfc6241>
24. European Telecommunications Standards Institute, "Network Functions Virtualization White Papers." [Online]. Available: https://portal.etsi.org/nfv/nfv_white_paper.pdf, https://portal.etsi.org/NFV/NFV_White_Paper2.pdf, https://portal.etsi.org/Portals/0/TBpages/NFV/Docs/NFV_White_Paper3.pdf
25. Cisco Systems Inc. "NFV Orchestration with Cisco Network Services Orchestrator." <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/solutions-cloud-providers/white-paper-c11-738702.html>
26. O-RAN NR U-plane profile for EN-DC Version 1.0 - June, 2019 (ORAN-WG5.U.0-v1.00), O-RAN Alliance, 2019. [Online]. Available: www.o-ran.org
27. O-RAN NR C-plane profile for EN-DC Version 1.0 - June, 2019 (ORAN-WG5.C.1-v1.00), O-RAN Alliance, 2019. [Online]. Available: www.o-ran.org