The evolution of 5G mobile networks:

the what, why and how

white paper

The evolution of 5G mobile networks: the what, why and how

white paper

What is 5G

5G is the 5th generation of mobile networks specified by the 3GPP organization. It promises capabilities within a mobile wireless network that were only thought possible through local area networks (LANs) that use fixed infrastructure. From a feature viewpoint, 5G offers higher throughput (25x), lower latency (1/25th) and higher capacity (500x) than 4G networks. This allows new applications to be developed such as high-definition 3D video, real-time virtual reality environments, autonomous vehicles, remote surgery and connectivity to billions of devices to support the Internet of things (IoT).

5G also promises higher reliability, improved energy efficiency and easier reconfiguration of the network to deploy additional services.

5G also promises higher reliability, improved energy efficiency and easier reconfiguration of the network to deploy additional services.

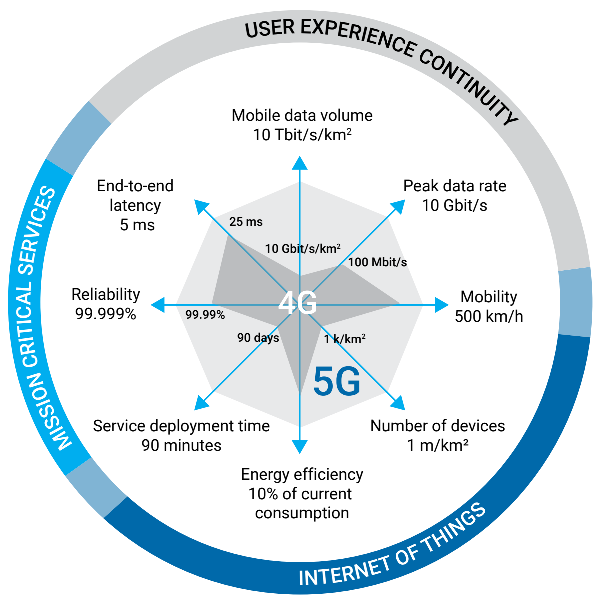


Figure 1. 4G vs 5G network features

In order to deliver these capabilities, 5G introduces new technologies including:

- The use of higher radio frequency (RF) bands to support additional bandwidth (mmWave)
- Faster and more efficient fronthaul connections (eCPRI)
- More reliable and cost effective timing provided over the network (IEEE1588/PTP)
- More granularity in the distribution of network functions (CU, DU, RU)
- The ability to dynamically configure the network infrastructure through network function virtualization (NFV)
- Cost savings by running virtual network functions (VNFs) over “white box” hardware
- The ability to implement separate quality of service levels for specific applications over a single physical network (network slicing)

These new technologies require testing to ensure the network can meet end user demands—and all of this must be achieved while supporting existing 4G and legacy infrastructure.

In this white paper we will provide an overview of these technologies and their role in implementing a 5G network.

5G non-standalone (NSA) architecture allows a 5G NR to use 5G for radio to handset communication (downlink or DL) while relying on existing 4G communication for the handset to radio head uplink (or UL) communication.

The evolution from 4G to 5G

At the end of December 2017, the 3GPP standards body approved an interim set of specifications for 5G networks focused on implementing enhanced mobile broadband (eMBB) features. This interim specification defined 5G new radio (5G NR) and a way of leveraging the existing 4G network to provide improved bandwidth and slightly improved latency. This 5G non-standalone (NSA) architecture allows a 5G NR to use 5G for radio to handset communication (downlink or DL) while relying on existing 4G communication for the handset to radio-head uplink (or UL) communication. eMBB features are the first that will be offered through 5G networks. The 5G NSA architecture will be followed by 5G standalone (SA), the architecture that will enable ultra low latency (uRLLC) and massive machine to machine communication (mMTC) applications. As shown in figure 2, this release will support the full core—supporting 5G on the UL and DL as well as providing even further improvements in latency and device connectivity.

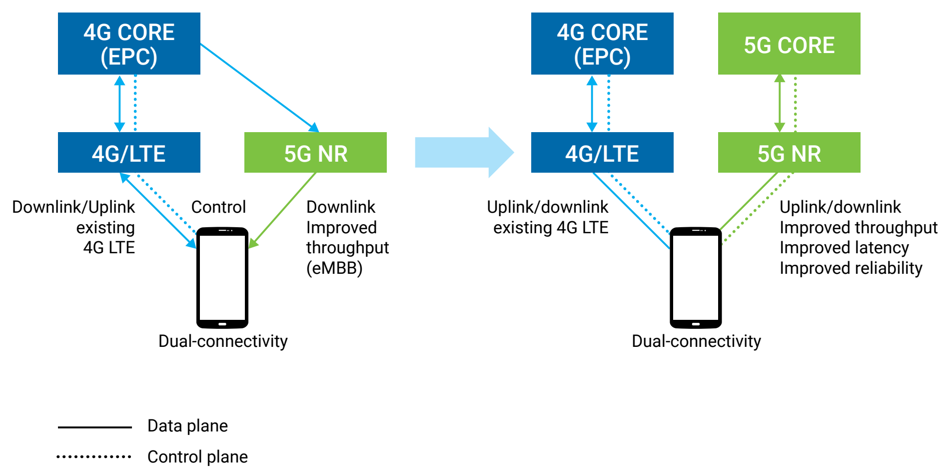


Figure 2. 5G NSA to 5G SA migration

During the roll-out of 4G networks we witnessed the migration of the radio access network (RAN) architecture into a distributed radio access network (D-RAN) where copper co-axial cables were replaced by optical fibers—referred to as fiber-to-the-antenna (FTTA). This transformation implemented a split of the radio elements,—generally referred to as the remote radio head (RRH),—and the baseband functions, generally referred to as the baseband unit (BBU). These were connected together with one of two competing digital RF communication protocols: common public radio interface (CPRI) and open base atation architecture initiative (OBSAI). These elements are typically referred to as fronthaul technology (see figure 3). This split of RRH and BBU functionality also allowed the grouping of several BBUs, from multiple cell sites, into a single location to save on infrastructure costs, creating a centralized RAN (C-RAN) architecture.

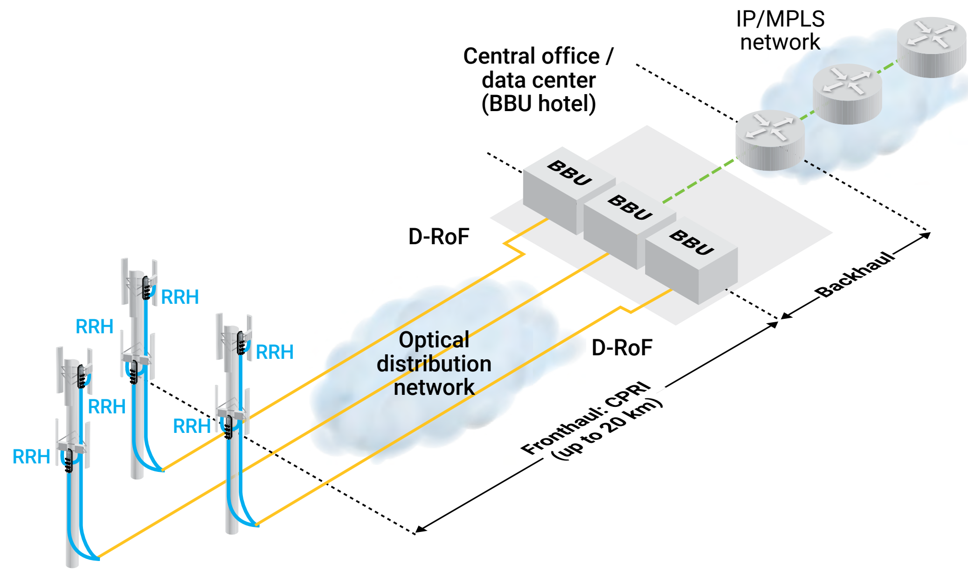


Figure 3. An example of a 4G fronthaul architecture

5G introduces a further split of the radio access architecture into three, rather than two elements: centralized unit, distributed unit and the radio unit.

5G introduces a further split of the radio access architecture into three, rather than two elements. These are called the centralized unit (CU), the distributed unit (DU) and the radio unit (RU). The network designer can decide where to place network functionality within these building blocks to meet specific network requirements, such as latency and throughput (as seen in figure 4).

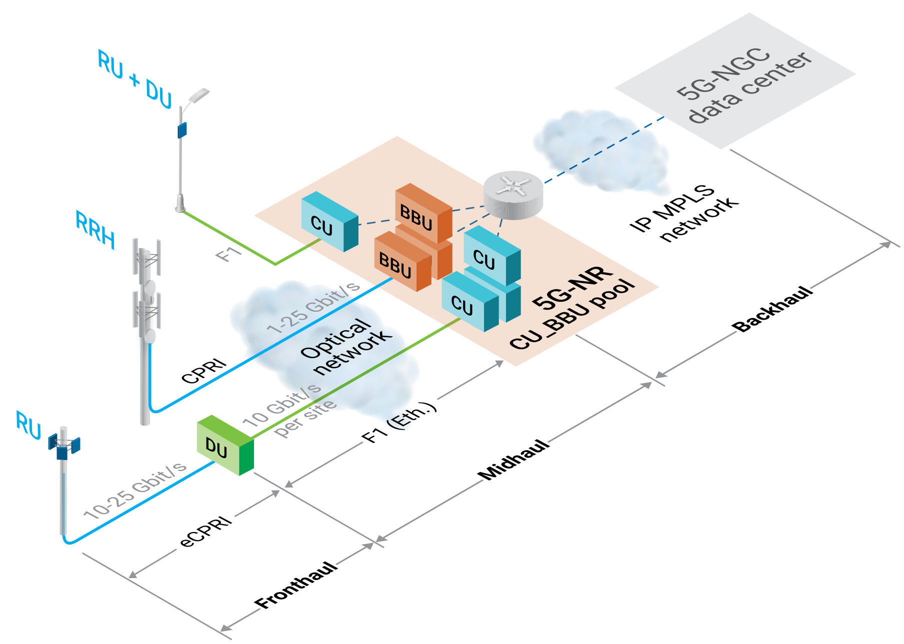


Figure 4. An example of a 5G fronthaul architecture

The presence of the DU introduces a new network element between the DU and the CU, referred to as the midhaul. Together the backhaul, midhaul and fronthaul networks are referred to as the xhaul.

5G interfaces

To improve the throughput of the fronthaul connection, 5G introduces eCPRI—a new protocol to connect the DU to the RU. eCPRI uses Ethernet as its physical layer and runs over an FTTA connection. This provides faster link rates of 10 Gbit/s and 25 Gbit/s, and more efficient use of fronthaul bandwidth.

Communication between the CU and the DU is covered by another Ethernet-based protocol stack called the F1 interface. Since RU, DU and even CU functions can be grouped together in a single device, cell sites may be connected to the rest of the infrastructure through either of these interfaces, depending on the operators chosen network architecture.

Mobile edge computing (MEC) allows the operator to move computation resources closer to the edge of the network.

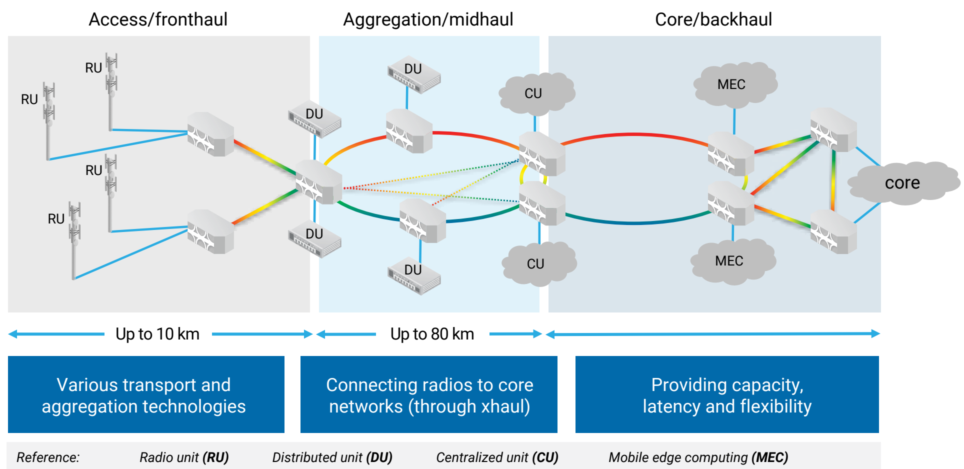


Figure 5. Xhaul network components

Another architectural improvement is called mobile edge computing (MEC). MEC allows the operator to move computation resources closer to the edge of the network. Doing so provides two benefits. First, latency is reduced providing faster response times for application. Second, as this reduces packets traveling through the core, network congestion in the core is reduced.

Comparison of 4G and 5G cell sites

4G cell sites

4G remote radio heads (RRHs) are typically mounted at the top of cell towers or on building rooftops. The ability to cover a large geographic area using radio frequencies in the sub-3 GHz spectrum was ideal for these installations. A single sector would consist of an RRH connected through a small form-factor pluggable (SFP) device to a single CPRI link using an optical cable between the RRH and another SFP in the BBU. A short, co-axial jumper cable connects the RF output ports on the RRH to the antenna. A single cell site might cover a range of about 10 km.

5G cells using mmWave will see a densification in the number of cell sites required to cover the same geographic area.

The split of the RRH and BBU functions allowed high loss coaxial cable running up the towers to be eliminated, reducing amplifier power requirements, which in turn led to smaller HVAC and back-up power requirements.

However, the lack of contiguous spectrum around 4G frequency bands limits the amount of bandwidth that can be offered to the end user.

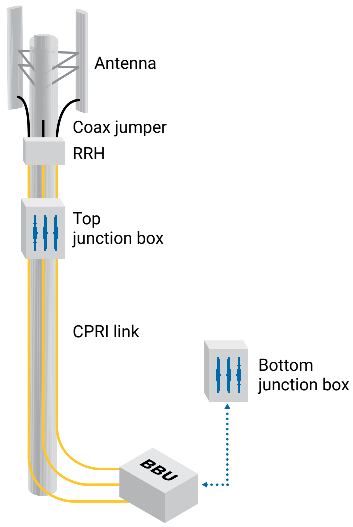


Figure 6. Typical 4G RRH

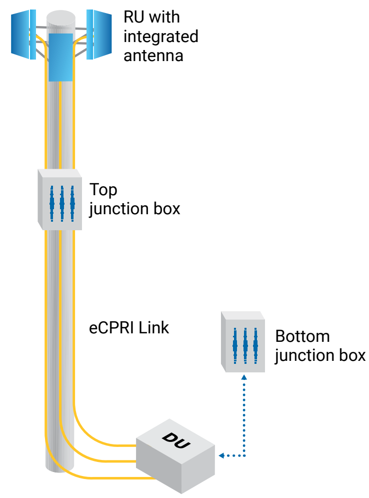


Figure 7. Typical 5G RU

5G cell sites

In 5G two frequency ranges are offered, FR1 (450-7125 MHz) and FR2 (24250-52600 MHz). The FR1 frequencies offer channel bandwidths of up to 100 MHz depending on the operator's spectrum holding while FR2 frequencies—referred to as millimeter wave (mmWave)—are in a less crowded area of the frequency spectrum and allow channel bandwidths of up to 400 MHz. The higher the channel bandwidth, the greater the potential throughput of the cell.

However, the higher the frequency the shorter the distance the RF signal will travel. So 5G cells using mmWave will see a densification in the number of cell sites required to cover the same geographic area. This densification is also required to support the higher capacity for the number of connected devices offered by 5G.

RUs using mmWave are typically mounted on street lamps, utility poles or the sides of buildings. This is due to the relatively short distances that mmWave can support. A single site might cover a range of only a few hundred feet.

In many 5G RU's the antenna will be integrated into the RU (see figure 7). This is due to massive MIMO and the mmWave frequencies that form part of 5G.

Timing and synchronization

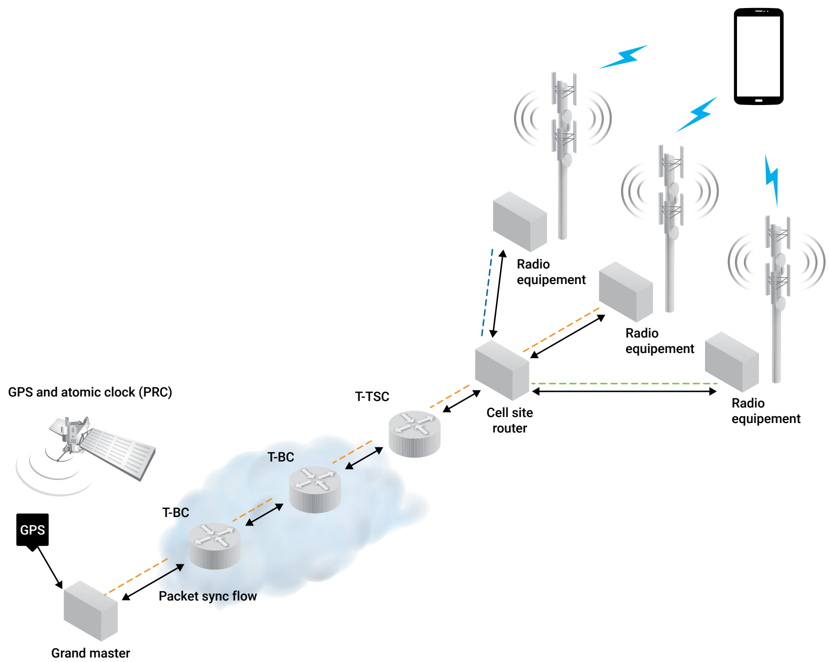


Figure 8. 5G timing and synchronization

The specific timing limits for each network will depend on the network architecture and the type of applications being supported.

The network densification in 5G and the positioning of 5G small cells in urban areas also creates new challenges in network timing and synchronization. LTE Advanced and 5G already place more stringent timing demands on the fronthaul network for technologies such as co-ordinated multipoint (CoMP), emergency location services (e.g., E911) and the use of time division duplexing (TDD) in RF transmission and reception.

In TDD the same frequency is used for the transmission and reception of data. Adjacent cell sites must transmit or receive in the agreed-upon time windows. If a cell site transmits during a receive window, then interference is experienced. To avoid this, an absolute timing accuracy of about 1.5 us is required from the core to the RU.

The principle of CoMP is for a UE to communicate with multiple, geographically separated cell sites, at the same time, with the goal of enhancing system performance and quality of service. Timing is critical in being able to reconcile the signals received at the coordinating cell sites. CoMP requires a relative timing accuracy between these sites of around 260 ns to 350 ns.

Emergency location services rely on observed time difference of arrival (OTDOA) technology where the UE measures the arrival time difference of the positioning reference signal (PRS) between a reference cell and several neighboring cells. The more accurate the time measurement, the more accurate the location information that is derived. The timing accuracy required to meet the requirements of OTDOA based applications can be as low as 100 ns.

In all cases the need for stable, accurate and reliable timing is an essential requirement for the network to operate correctly.

Multiple input multiple output technology can be used to increase the capacity of a radio head and the reliability of the wireless transmission.

The use of a global network satellite system (GNSS) receiver to provide timing at each cell site would be expensive and not always practical when small cell locations do not have line of site to a sufficient number of satellites, such as may be the case in a city center (e.g., San Francisco or New York City).

5G networks will use IEEE1588/Precision Time Protocol (PTP) and SyncE to allow cell sites to be synchronized in both frequency and phase using the Ethernet network to distribute this information to each cell site and ensuring that all devices are in-sync with each other (see figure 8).

The specific timing limits for each network will depend on the network architecture and the type of applications being supported.

Massive MIMO

Multiple input multiple output (MIMO) technology can be used to increase the capacity of a radio head and the reliability of the wireless transmission. The most common technique used is spacial diversity which allows radio waves to be transmitted along multiple paths to multiple antennas, increasing the probability that if one path fails the other paths can provide sufficient signal to maintain the connection, hence increasing the reliability.

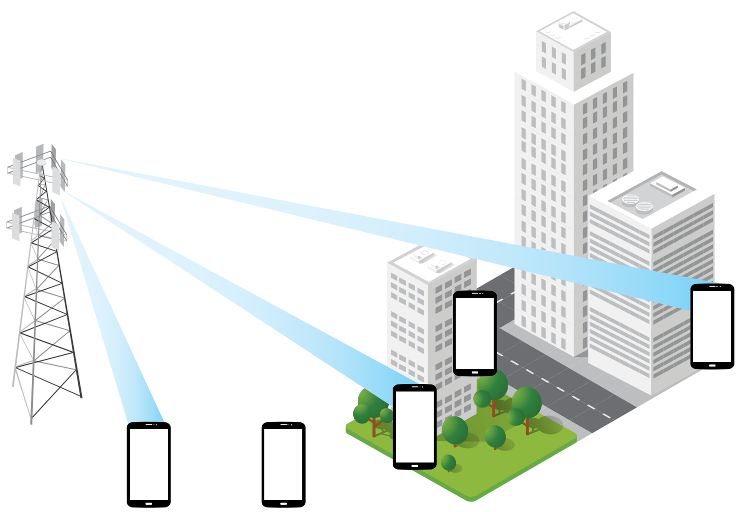


Figure 9. Beamforming

In 4G, two transmit and two receive (2x2) and four transmit and four receive (4x4) configurations are the most prevalent. To increase capacity MIMO's ability to use spatial multiplexing to transmit different information on these multiple paths, on the same frequency, without interference can be used. In 5G massive MIMO (mMIMO), antenna array sizes in excess of 64x64 are being implemented. 5G RU's supporting mMIMO have integrated antenna arrays due to the cabling and connector issues that would otherwise arise connecting a separate antenna, and due to the signal losses that would be encountered at FR2 frequencies. This eliminates the co-ax cable used in smaller MIMO configurations.

The use of mMIMO also facilities the use of beamforming technology which allows the RU to focus the RF energy of a subset of the MIMO array towards a specific UE.

5G network architecture

The architecture of a 5G cell site depends on the architectural split chosen by the mobile network operator (MNO). The 3GPP and eCPRI standards provide multiple options, or splits, of functionality based on the building blocks shown in figure 10. The various functions can then be performed by the CU, DU or RU, depending on the option chosen. Some examples of the most common splits are shown below.

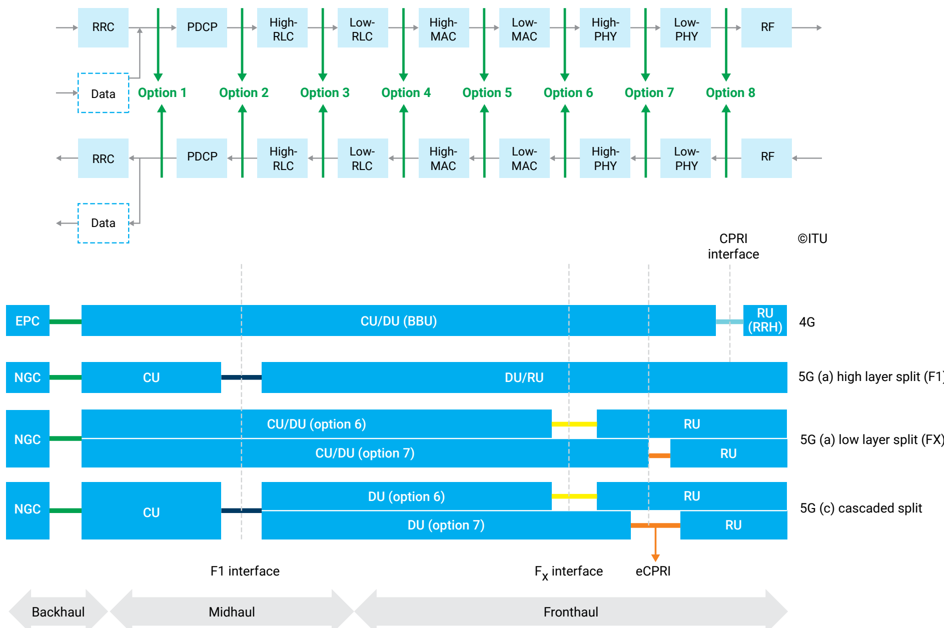
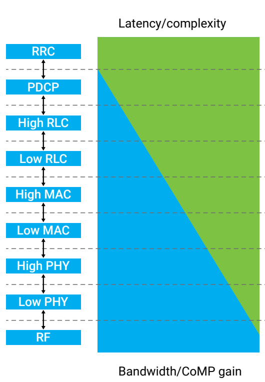


Figure 10. 3GPP and eCPRI functional splits mapped onto CU/DU/RU

As we journey through 5G we will see additional standards that will allow interoperability between multi-vendor equipment.

The choice of which option to use depends on the desired throughput and latency requirements of the overall network. Figure 11 gives some examples of the performance expected.



	Connection	Test results*		3GPP requirements		Impact on CoMP
		Latency	Bandwidth	Latency	Bandwidth	
Option 1 (split A)	RRC-PDCP	50 ms	110 Mbit/s	10 ms	DL: 4 Gbit/s UL: 3 Gbit/s	−75%
Option 2 (split B)	PDCP-RLC	30 ms	151 Mbit/s	1.5~10 ms	DL: 4016 Mbit/s UL: 3024 Mbit/s	−55%
Option 3	RLC-RLC	24 ms	151 Mbit/s	1.5~10 ms	Lower than option 2 for UL/DL	−40%
Option 4 (split C)	RLC-MAC	15 ms	151 Mbit/s	approximate 100 us	DL: 4000 Mbit/s UL: 3000 Mbit/s	−35%
Option 5	Split MAC	6 ms	151 Mbit/s	hundreds of microseconds	DL: 4000 Mbit/s UL: 3000 Mbit/s	−25%
Option 6 (split D)	MAC-PHY	2 ms	152 Mbit/s	250 us	DL: 4133 Mbit/s UL: 5640 Mbit/s	−15%
Option 7	Split PHY	0,25-2 ms	1075 Mbit/s	250 us	DL: 10.1~22.2 Gbit/s UL: 16.6~21.6 Gbit/s	−5%
Option 8 (split E)	CPRI	<250 us	2500 Mbit/s	250 us	DL: 37.8~86.1 Gbit/s UL: 53.8~86.1 Gbit/s	0%

*Test results are based on current work and might be updated in the future.

Figure 11. Throughout and latency characteristics of network split options

Splitting network functions between CU, DU and RU requires the definition of an additional standard for interconnecting these devices. Just as the splitting of RRH and BBU functions introduced the use of the CPRI protocol, the new functional splits introduce eCPRI and F1, both based on Ethernet technology. Traditionally, the devices used to implement the network functions have been supplied by a single manufacturer. As we journey through 5G we will see additional standards that will allow interoperability between multi-vendor equipment. Initiatives such as the open radio access network (O-RAN) and the Telecom Infra Project (TIP), among others, are actively working towards this goal, with initial deployments targeted in late 2020. EXFO is an active participant in these organizations and will support test equipment for compliant devices as the market evolves.

5G introduces the concept of network slicing which avoids the need to construct separate networks with different architectures to support the latency and throughout requirements of the various applications.

The future of 5G

Not only will future 5G implementations support different network splits but they will also allow MNOs to select the best devices to implement that split from a selection of multiple vendors.

5G introduces the concept of network slicing which avoids the need to construct separate networks with different architectures to support the latency and throughout requirements of the various applications. This allows the bandwidth on the physical network to be split into numerous slices, each slice dedicated to providing a specific capability.

5G network slicing

5G network slicing enables service providers to build virtual end-to-end networks tailored to application requirements.

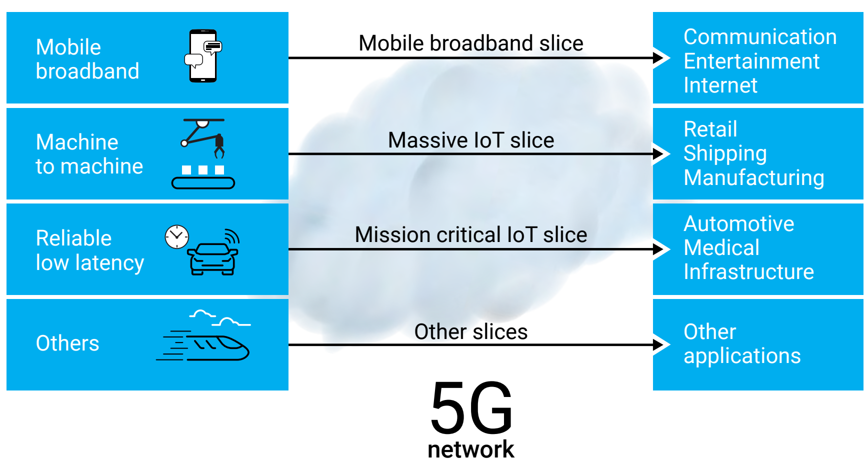


Figure 12. Network slicing

Network virtualization

Another concept that 5G embraces is network function virtualization (NFV). Imagine that each of the functions in a 5G network could be modeled in software and become a virtualized network function (VNF), and that these models could be coupled together, using the standardized communications protocols described above, to create a complete model of all or part of the network. If we also virtualize the computer hardware (i.e., provide a common interface, regardless of the underlying hardware) we can then run the network on off-the-shelf computer hardware, so called "white boxes", reducing overall network costs.

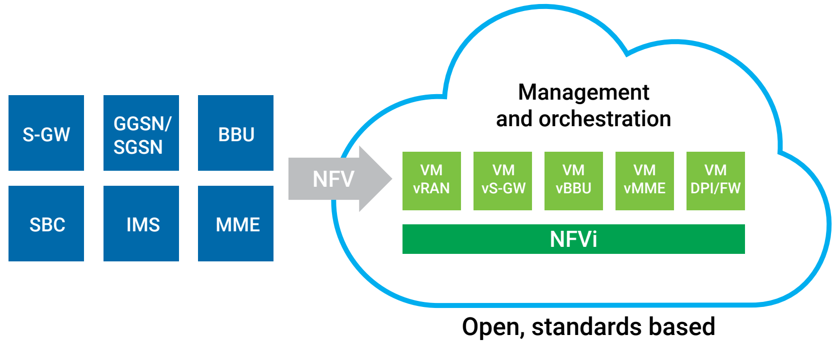


Figure 13. Network function virtualization and orchestration

A further concept in NFV is the orchestration process. To build highly reliable and scalable services, NFV requires that the network be able to create new instances of VNF's and monitor, repair, and bill for the services rendered. These functions are allocated to an orchestration layer to ensure high availability and security, with low operation and maintenance costs. Importantly, the orchestration layer must be able to manage VNFs from different vendors, irrespective of the underlying technology used, allowing the end user to pick and choose the VNFs needed to build their network.

Another complementary technology is software defined networking (SDN). In SDN the control and data planes of the network equipment are separated allowing the network equipment to be configured through an SDN controller which would form part of the orchestration function. This would allow the MNO to be able to change the network configuration through orchestration when defining new virtually defined services.

Eventually, when the mobile network is virtualized, the MNO will be able to dynamically create new network slices and services through software tools saving time and resources associated with the deployment of proprietary hardware.

Conclusion

As you can see, there is much more to 5G than an upgrade in the speed of the network, with many changes to the way mobile networks are implemented and used. This brave new world also brings with it new challenges in constructing, operating and maintaining these networks.

Rather than 5G being a single solution implemented by all operators, it offers a selection of features from which operators can choose. Operators will choose different solutions to meet cost, coverage, latency, throughput and other parameters required to meet their customer requirements and their operating objectives. For example, in North America we see 5G networks implemented to provide the most coverage, but with lower throughput as well as networks that provide limited coverage but with very high throughput. This will evolve as mobile operators expand their 5G coverage.

As 5G capabilities are expanded beyond higher throughput, to include low latency and increased capacity, there will be new user applications being developed—particularly in the areas of intelligent vehicles and smart devices connected through the mobile network.

EXFO continues to be an active participant in the 5G standardization process and is providing tools and innovations to help you meet the challenges ahead. Please contact us if you have questions or would like further information.

Additional resources

Discover our products

[FTB 5GPro—5G test solution for mobile networks]

[Nova SkyRAN—fronthaul remote access and monitoring]

Read our blog post

[On the road to 5G: upgrading the transport network]

[5G technology challenges and other (5G survey) highlights]

[5G testing challenges: EXFO solutions for top 5 in 5G networks]

Glossary of terms

3D	three dimensional	**mMIMO**	massive MIMO
3GPP	third generation partnership project	**mMTC**	massive machine to machine communication
5G NR	5G new radio	**mmWave**	millimeter wave
BBU	baseband unit	**MNO**	mobile network operator
CoMP	coordinated multi point	**NFV**	network functions virtualization
CPRI	common public radio interface	**NSA**	non-standalone
C-RAN	centralized RAN	**OBSAI**	open base station architecture initiative
CU	centralized unit	**O-RAN**	open RAN
DL	downlink	**OTDOA**	observed time difference of arrival
D-RAN	distributed RAN	**PRS**	positioning reference signal
DU	distributed unit	**PTP**	precision time protocol
eCPRI	enhanced CPRI	**RAN**	radio access network
eMBB	enhanced mobile broadband	**RF**	radio frequency
EPC	evolved packet core	**RRH**	remote radio head
FTTA	fiber to the antenna	**RU**	remote unit
GNSS	global navigation satellite system	**SA**	stand alone
HD	high definition	**SDN**	software defined network
HVAC	heating, ventilation and air conditioning	**SFP**	small form-factor pluggable
IEEE	Institute of Electrical and Electronic Engineers	**SyncE**	synchronous ethernet
IoT	Internet of things	**TDD**	time division duplex
LAN	local area network	**TIP**	telecom infra project
LTE	long-term evolution	**UL**	uplink
MEC	mobile edge computing	**uRLLC**	ultra-reliable low latency communication
MIMO	multiple input multiple output	**VNF**	virtualized network function

EXFO serves over 2000 customers in more than 100 countries.
To find your local office contact details, please go to **EXFO.com/contact.**