5G Deployment: Standalone vs. Non-Standalone from the Operator Perspective

Guangyi Liu, Yuhong Huang, Zhuo Chen, Liang Liu, Qixing Wang, and Na Li

Abstract

The fifth generation (5G) mobile network is standardized and developed to explore the mobile market beyond 2020. In response to the diverse strategies of 5G deployment, five alternative network architectures have been proposed to 3GPP by different mobile operators. To fulfill the urgent deployment requirement from some operators, an early drop of 5G, termed as non-standalone (NSA) new radio (NR), was completed at the end of 2017. After that, the standardization of a new 5G system, including th standalone (SA) new radio access network, was finished in June 2018. This article analyzes and compares the SA NR and NSA NR deployment modes in terms of coverage, network capability, interworking between 4G and 5G, complexity and cost of network deployment, and the latest industry progress. NSA NR performs better in interworking performance in the initial phase, while SA NR performs better in network capabilities, device performance, simple network deployment, and cost efficiency. 5G SA NR is recommended for operators who have the ambition to explore new opportunities in the vertical and enterprise markets.

Introduction

The fourth generation mobile communication (4G, also referred to as Long Term Evolution/Evolved Packet Core, LTE/EPC) has opened opportunities for mobile broadband communication and played an important role in promoting new mobile services and applications. Some of the communication requirements of the new services and applications, including enhanced mobile broadband (eMBB), massive machine type communication (mMTC), and ultra-reliable low-latency communication (URLLC), have exceeded the capabilities of the 4G network [1]. To address these new demands, 5G is being developed.

To deliver the highly reliable, ultra-low-latency, multi-gigabit connectivity that 5G portends, sufficient radio spectrum is a prerequisite. To offer premium 5G experience, a significant amount of new harmonized spectrum is demanded for the 5G network deployment, ideally 80–100 MHz of contiguous spectrum per operator in below 6 GHz bands and around 1 GHz per operator in millimeter-wave (mmWave) bands. For 5G's nationwide coverage, spectrum below 2 GHz (typically frequency-division duplex, FDD, bands) is also needed [2], especially for extremely-low-latency use cases.

At the International Telecommunication Union (ITU) World Radiocommunication Conference 2015 (WRC-15), some new bands below 6 GHz were identified for mobile service in different regions and countries: 3.3~3.4 GHz, 4.4~4.5 GHz, and 4.5~5 GHz. At WRC-19, two new bands for mobile service were identified: 26 GHz (24.25–27.5 GHz) and 40 GHz (37–43.5 GHz). Two other frequency bands received mobile identification: 66 GHz (66–71 GHz, for unlicensed use) and 50 GHz (45.5–47 GHz and 47.2–48.2 GHz, in designated countries) [2]. Early rollouts of 5G in Europe and Asia will likely use spectrum below 6 GHz (e.g., 2.6 GHz/3.5 GHz), while operators in the United States will mainly start with the bands above 6 GHz for 5G, and refarm some 4G spectrum [3].

With the guidance of ITU Radiocommunication Standardization Sector (ITU-R) on the technical requirements and evaluation methodology, the Third Generation Partnership Project (3GPP) began 5G standardization in 2015 and scheduled its first release of specifications on a 5G system in June 2018 [4], including both the new air interface (New Radio, NR) and 5G Core Network (5GC). The new 5G system is referred to as standalone (SA), which could bring the full 5G capabilities and is regarded as the target 5G architecture.

However, as a huge amount of money has been invested to build a 4G network in the last decade, the mobile operators face a challenge to have sufficient financial support for large-scale 5G deployment. Because of the different forecasts on service opportunities brought by 5G and the concerns on the return of investment on 5G networks, many operators are pessimistic about launching a large-scale 5G network in 2020. Due to the different considerations on 5G deployment, five diverse network architectures [5] for 5G NR deployment have been proposed to 3GPP by different operators, which may lead to a fragmented 5G industry and market.

On the other hand, the intense competition drives the operators to compete for regional leadership on 5G deployment. It seems a good compromise between capacity expanding and investment scale to add 5G NR air interface to the legacy 4G (i.e., LTE/EPC) network as an extra data pipe. Thus, a non-standalone (NSA) 5G deployment based on dual connectivity between 4G and 5G was standardized in 3GPP by the end of 2017, in which the device anchors at the 4G

The authors analyze and compare the SA NR and NSA NR deployment modes in terms of coverage, network capability, interworking between 4G and 5G, complexity and cost of network deployment, and the latest industry progress.

The authors are with the China Mobile Research Institute.

Digital Object Identifier: 10.1109/MCOM.001.2000230

The SA NR deployment not only brings a new E2E network architecture, but also enables E2E network capabilities, for example, fast initial access and data transmission, network slicing and mobile edge computing (MEC), which is essential to explore the new service opportunities from enterprise and vertical markets.

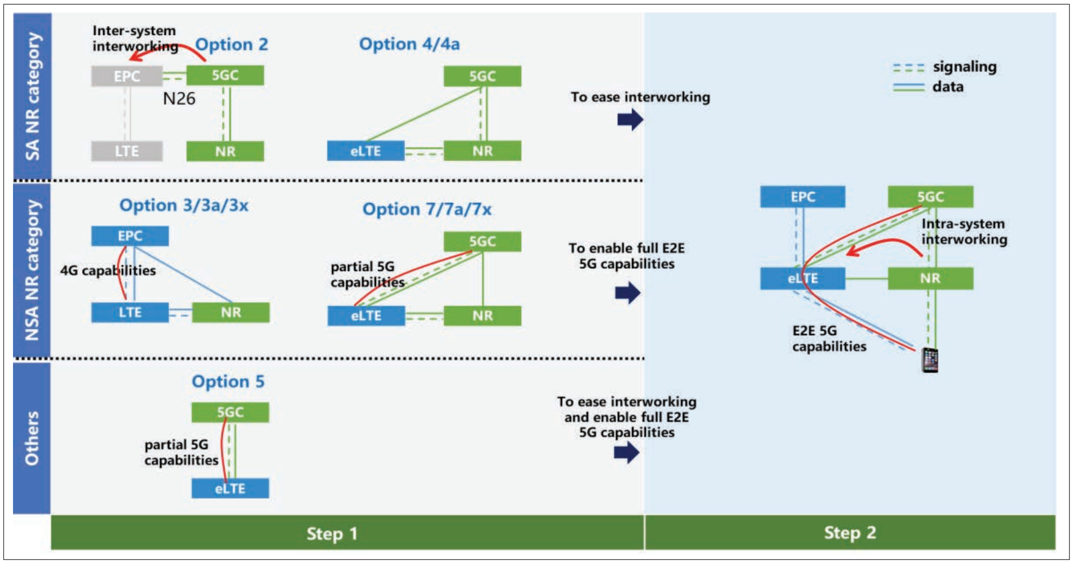


Figure 1. Five network architecture options proposed to 3GPP and the potential migration paths.

network, and 5G NR could work as an extra data pipe when NR capability is required and coverage is available.

In general, the five network architecture options identified by 3GPP fall into two major categories: SA NR and NSA NR. With so many options and two different categories, most of the operators are puzzled by selecting from them. The 5G chipset and device will face a big challenge to support all the diverse architectures and the corresponding global roaming. To facilitate the understanding and determination of SA and NSA, it is necessary to clarify the pros and cons of SA NR and NSA NR carefully.

For the operators that have deployed LTE at 900 MHz or 1800 MHz, LTE's coverage will be much better than that of 3.5 GHz NR. If the operators do not aim for full NR coverage from the beginning, NSA option 3 is the natural way to deploy 5G since it can avoid the frequent interworking between LTE and NR, and avoid the deployment of the 5G core network. For other operators who are ambitious to provide full NR coverage and ultimate service experience, SA option 2 may be a better choice. From the network deployment aspect, SA option 2 and NSA option 3 are two typical architectures supported by mobile network vendors and operators [6]. Hence, these two alternatives are the focus of this article. The detailed concepts of SA and NSA are introduced in the following section, and SA option 2 and NSA option 3 are discussed carefully in terms of coverage, network capabilities, terminal performance, 4G/5G interoperation, and deployment cost following that, and the recommendation of the network architecture is suggested.

Concepts of 5G SA NR and NSA NR

In general, SA NR architecture refers to a 5G system consisting of NR and 5GC. NR is the control plane anchor, while NSA NR architecture refers to a system that uses LTE/evolved LTE (eLTE) as the control plane anchor for NR. Actually, both SA NR and NSA NR architectures consist of some variants [5]. The five sets of architecture alternatives proposed to 3GPP are given in Fig.1, where the option 2 and option 4 sets fall into the SA NR category, while the option 3 and option 7 sets belong to the NSA NR category. The difference between the options in one set is that option *N* supports a split bearer, option *Na* supports a secondary cell group (SCG) bearer, and option *Nx* supports an SCG split bearer, where *N* could be 3, 4, or 7. Option 5 considers the case that the eLTE base station (ng-evolved Node B, ng-eNB) connects to 5GC, and this deployment mode is not related to NR.

The network migration steps of NSA NR and SA NR are also illustrated in Fig. 1. In our understanding, no matter which one is deployed in the initial stage, the ultimate deployment mode is the same.

Considering that SA option 2 and NSA option 3 are the typical architectures supported by mobile network vendors and operators and thus the focus of this article, we use NSA to denote NSA option 3 and SA to denote SA option 2 in the remainder of the article for brevity.

In the first step of SA NR, to enable interworking between the EPC and 5GC to guarantee service continuity, an N26 interface is a must. When a 5G device is within the coverage of NR, it anchors at the 5GC and its mobility is managed by the 5GC. When it moves out of NR coverage, the user equipment (UE) is handed over to an LTE/EPC network, where UE connects to LTE/EPC like a general LTE device. This means that a 5G SA UE only works in either 5G mode or 4G mode. As a further evolution of LTE, the LTE eNB could be further updated by certain new features as an ng-eNB and connected to the 5GC (i.e., step 2). The legacy LTE UE accesses and anchors at the legacy EPC, while the eLTE UE can be anchored at the 5GC and enjoy the new services provided by the 5GC, for example, end-to-end (E2E) network slicing.

The SA NR deployment not only brings a new E2E network architecture, but also enables E2E network capabilities, for example, fast initial access and data transmission, network slicing, and mobile edge computing (MEC), which is essential to explore the new service opportunities from enterprise and vertical markets.

As illustrated in step 1 of NSA in Fig. 1, the device is anchored to the LTE/EPC system, and NR is utilized as an extra data pipe when NR coverage is available. When moving outside of NR coverage, the NSA UE only works as an LTE UE. The mobility management of NSA NR is completely controlled by the LTE system, which provides a better user experience since the interworking between LTE and NR is intra-system handover. When service data transmission is accomplished by LTE and NR simultaneously, the data can be directly offloaded to the NR gNB from the LTE eNB or from the EPC to the NR gNB. For eNB offloading, the hardware of the legacy LTE eNB must be upgraded to support higher throughput. EPC offloading does not require hardware upgrades of the legacy LTE eNB.

NSA NR requires the device to support dual connectivity, which means the device should maintain LTE and NR radio transmission links simultaneously. Sharing of the total transmission power of the device between the two transmission links leads to more limited UL coverage than the LTE system and shorter battery life. In some paired bands for NSA NR, interference [7] between the two communication links may happen and lead to worse terminal performance.

Analysis of SA vs. NSA

Typically, to select one network architecture from five alternatives, many issues need to be considered, for example, radio coverage, network capabilities, terminal performance, 4G/5G interworking, complexity of network deployment, and cost for further evolution. Details of these aspects are given in this section.

Coverage

Coverage is one of the key factors that an operator considers when commercializing cellular communication networks due to its direct impact on service quality as well as capital expenditure (CAPEX). If 5G NR cannot provide continuous coverage like that of the legacy LTE network, it will cause frequent inter-system inter-radio access technology (RAT) mobility between 4G and 5G in SA when UE moves into/out of NR coverage, leading to degradation of the user experience. Comparably, only intra-system intra-RAT mobility happens in NSA as an NSA UE anchors in LTE/EPC, and the mobility management is done by LTE/EPC. In this case, NSA may outperform SA. Otherwise, SA will be better if NR's coverage is better than that of LTE. On the other hand, to simplify the network deployment and control the CAPEX of a 5G network, mobile operators desire to deploy the 5G NR network by reusing existing LTE sites. Considering NR is designed to operate at much higher frequencies compared to LTE, and it is inevitable that the wireless channel will be subject to higher path loss, it will be more challenging to maintain an adequate quality of service (QoS) that is at least equal to that of LTE.

As described above, the coverage performance of NR does impact the selection of NSA and SA. Therefore, it is highly important to evaluate 5G NR's coverage comprehensively by comparing it with that of LTE.

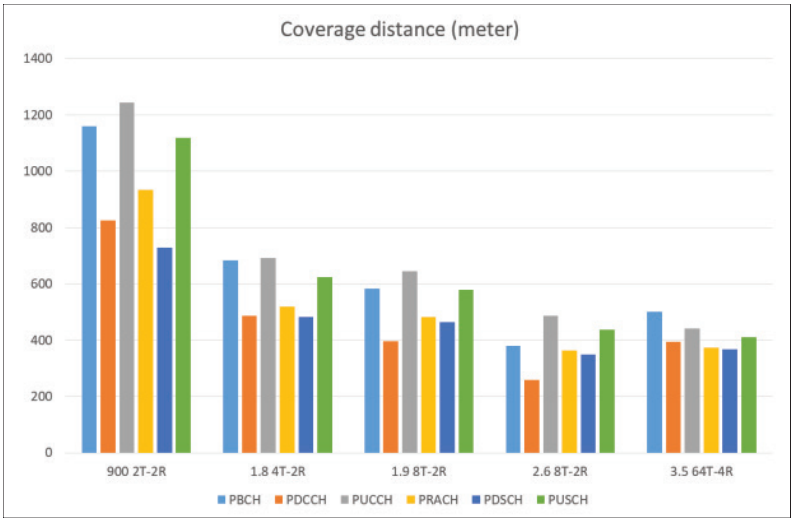


Figure 2. Coverage distance of different channels.

Considering that the NR leg of NSA and SA share almost the same layer 1 and layer 2 protocol functions and configurations, the performance of their data channels or UE-specific channels are almost the same. Their difference is determined by the common channel (e.g., broadcast channel, cell-specific control channel). As the common channel of NSA and SA is transmitted by LTE and 5G NR, respectively, the coverage comparison of NSA and SA will be done between LTE and 5G NR.

Since the candidate bands for the initial deployment of 5G is the high part of sub-6 GHz (e.g., 3.5 GHz), it usually means poor coverage due to much larger propagation loss. Many schemes to enhance the coverage of NR are standardized in 3GPP specifications [8], for example, massive multiple-input multiple-output (MIMO), beam sweeping of MIMO, larger control channel elements (CCE) of 16, 4Rx/2Tx of device, 26 dBm transmission power of UE, and extended format for the physical random access channel (PRACH). Comprehensively, the schemes above offer the 5G NR system a coverage gain more than 10 dB/5 dB in downlink (DL) and UL, respectively. Considering this, the link budgets of TD-LTE and 5G NR are calculated and compared based on the assumptions given by [9] (Fig. 2). To facilitate the comparison, the different cases of LTE and 5G NR are defined as follows:

- 3.5 GHz NR (64T-4R): 64-antenna port at the base station, 4Rx/2Tx at the terminal with 26 dBm
- 900 MHz FDD LTE(2T-2R): 2Tx/2Rx at the base station and 2Rx/1Tx at the terminal
- 1.8 GHz FDD LTE (4T-2R): 4Rx/2Tx antennas at the base station and 2Rx/1Tx at the terminal
- 1.9 GHz TD-LTE(8T-2R): 8Tx/Rx at the base station and 2Rx/1Tx at the terminal
- 2.6 GHz TD-LTE (8T-2R): 8Tx/Rx at the base station and 2Rx/1Tx at the terminal

To get the comprehensive cell radius, the coverage radius of the physical broadcast channel (PBCH), physical dedicated control channel (PDCCH), physical random access channel (PRACH), physical downlink shared channel (PDSCH), and physical uplink shared channel

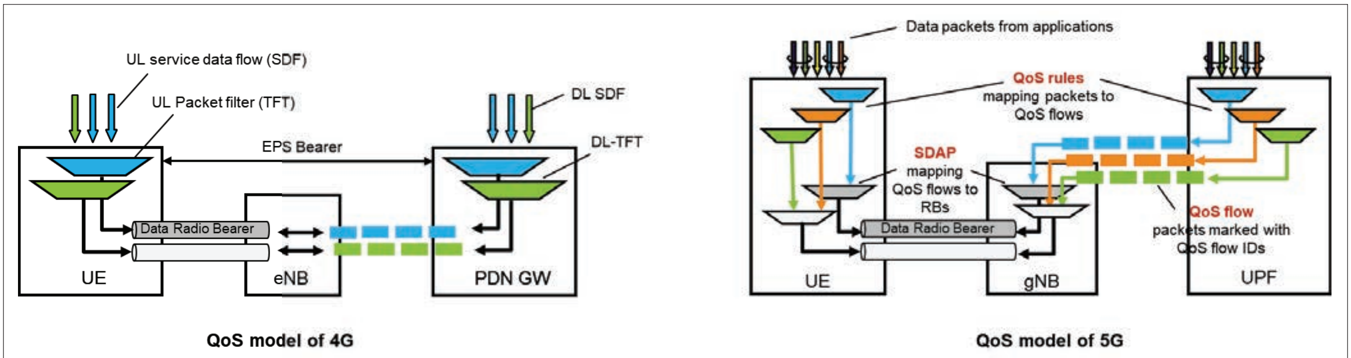


Figure 3. QoS model comparison between 4G and 5G.

(PUSCH) are calculated. For PUSCH coverage calculation, the typical case of 128 kb/s data rate is assumed [9].

Because higher bands lead to larger propagation loss and penetration loss, the coverage of 3.5 GHz NR seems to be much worse than that of 900 MHz, 1.8 GHz FDD LTE. However, due to the coverage enhancement features introduced, the coverage performance of 3.5 GHz NR looks better than that of 2.6 GHz TD-LTE and very close to that of 1.9 GHz TD-LTE.

Furthermore, the case of China Mobile is analyzed as an example. Considering that China Mobile has adopted 2.6 GHz for TD-LTE deployment in its urban area, it seems that 3.5 GHz 5G NR achieves better coverage than the legacy 4G network in an urban area by reusing the legacy 4G sites. Thus SA will not lead to frequent 4G/5G interworking in China Mobile's early 5G deployment in an urban area.

Network Capabilities

Since NSA mainly targets meeting the capacity demand from eMBB, it will not have the same network capabilities as SA (e.g., the network slicing and finer QoS treatment) [10].

Network Slicing: 5GC introduces a service-based architecture and targets for building a Software Defined Network (SDN) [11] and Network Function Virtualization (NFV) [12] platform. SDN/NFV based 5GC facilitates the elastic capacity expansion, fast service launch and software update. The most important capability of 5GC is to provide the E2E network slicing, which is the enabler to timely offer the service and deployment requirements from the much diverse vertical industry and enterprise. The different slices can also be separated each other logically or physically to protect the privacy and security as requested.

In some case, the EPC could also be virtualized and enjoy the corresponding benefits, like the fast software update and service launch in some degree. Nevertheless, due to constraints from the EPC architecture, it can't support the thorough split of control plane and data plane as the service-based architecture proposed in 5GC, and thus the transmission efficiency of the user data will be lower and the latency is higher than that of 5GC.

5GC can be deployed based on the general CPU platform, the entire network can be sliced as three types of virtual network, for example, eMBB, mMTC and URLLC, to meet the requirements from the diverse vertical industries. Further, the 5GC can also be deployed in a distributed or centralized way on demand, and the flexible MEC can be supported as requested to meet the stringent performance requirements of devise services, especially the low latency requirement of URLLC services. Such features are the key for the operators to exploit the potential vertical and enterprise markets besides the traditional personal communication market and monetize their 5G network and grow their revenue.

Finer QoS Treatment: For NSA, the QoS management is aligned with that in 4G, since QoS is controlled by the core network, and the core network in NSA is EPC. Similarly, QoS in SA is aligned with that in 5G.

In 4G, the EPC takes full responsibility for QoS functions with one-level mapping, that is, service data flow (SDF) to Evolved Packet System (EPS) bearer at the gateway (GW) (Fig. 3). In this framework, the eNB does not participate in QoS management but provides radio bearers (RBs) as part of the EPS bearers and enforces the QoS by radio resource management (RRM) tools such as scheduling. To better cope with the heterogeneous service requirements, instead of 4G SDF-based QoS, 5G employs a QoS-flow-based framework with two-level mapping (Fig. 3): non-access-stratum (NAS)-level packet filters in the UE and 5GC associate UL and DL packets with QoS flows according to QoS rules, and the newly introduced SDAP layer in the UE and gNB associate UL and DL QoS flows with data RBs.

Note that in this two-level framework, the 5GC only defines the treatment requirements of service packets by filtering them into QoS flows, and how to bear them at the radio interface is up to the gNB. It allows the gNB to decide/modify the L2 configurations (i.e., mapping/remapping QoS flows to RBs) for data transmission according to the actual channel status at the radio interface, making QoS management in 5G more accurate and flexible. By independent and transparent separation of QoS management at the NAS and AS levels, this framework provides more options and agility for E2E slice implementation as well.

Device Performance

From the device perspective, the NSA and SA behave in different ways and lead to different user experience. NSA keeps two radio links to LTE and NR simultaneously, so it seems that it could enjoy higher peak data rate . However, the practical performance of NSA is much influenced by the paired bands for LTE and NR, in which some of them may introduce mutual interference and degrade the device performance due to the constraint on the implementation of the NSA device.

Interference Analysis on NSA's Dual Connectivity: Typically, NSA deployment reuses the legacy LTE network in lower frequency, for example, 700 MHz to 2.6 GHz, and NR is deployed in new frequency, for example, 3 GHz~5 GHz. The intermodulation and harmonic may happen between LTE and NR of NSA when UE works in a dual connectivity mode. In this article, 900/1800 MHz for FDD, 1.9/2.6 GHz for TD-LTE, and 3.4~3.6 GHz for 5G NR are taken as references for the interference analysis since these bands are the typical frequency for the mobile industry.

From the analysis [7], we can see there is no interference between 1.9/2.6 GHz and 3.4~3.6 GHz. However, the interference happens between 900 MHz/1.8 GHz and 3.5 GHz:

- The 2nd order harmonic of 1800 MHz interferes with 3.5 GHz (3.42–3.47 GHz), and the 2nd order intermodulation interferes with the downlink of 1.8 GHz.
- The 4th order harmonic of 900 MHz interferes with 3.5 GHz (3.556–3.616 GHz)

When implemented, to avoid the interference above, simultaneous transmission of LTE and 5G NR links should not be scheduled in NSA, which leads to lower peak data rate. To improve the peak data rate, some complicated or advanced scheduling could avoid the possible interference by muting the resource blocks, which could introduce interference with each other.

Although there is no interference between 1.9 GHz/2.6 GHz TD-LTE and 3.5 GHz NR, the transmission alignment between two time-division duplex (TDD) radio links assumed in 3GPP standardization work leads to longer air interface latency. As TD-LTE is a legacy network, it is very difficult to change its frame structure. The only way to achieve the transmission alignment is to match the frame structure of the NR system with that of TD-LTE. The widely used frame duration of TD-LTE is 5 ms, which makes the NSA suffer from double latency of the air interface when compared to the 2 ms or 2.5 ms frame duration of 5G SA. The longer latency will put a constraint on the NR leg in NSA to address some extremely low-latency use case.

For SA, NR and LTE will not work simultaneously, and thus there is no intermodulation or harmonic interference happening to the 5G device.

Power Consumption

For the initial commercial launch of 5G, the battery life of the terminal will impact the user experience much. So the power consumption of the 5G device is also a very important aspect to be considered.

Due to much larger bandwidth and much higher transmit data rate, 5G NR leads to high power consumption. Due to dual connectivity of the device, NSA may have higher power consumption.

For the power consumption analysis, three typical modes of a device should be considered: idle mode, connected mode, and data transmission mode. Since there is no RRC-inactive state

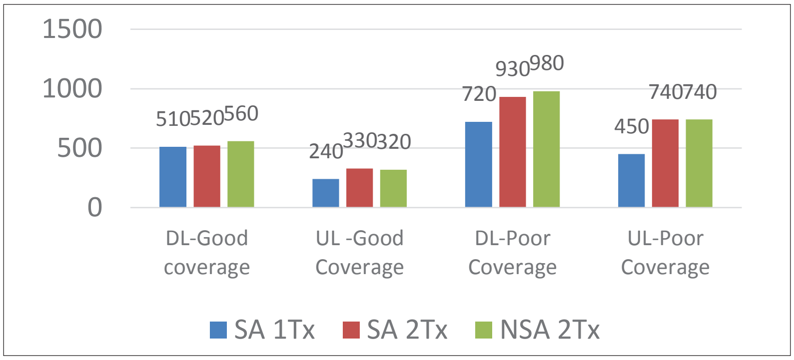


Figure 4. Power consumption of SA vs. NSA.

in NSA, the performance comparison of this state is not considered. To be fair, only the power consumption of the communication module of a terminal is considered and compared.

For the idle mode, only the necessary broadcasting or paging information is received and the power consumption have nothing to do with the UL. Because a similar amount of data is received from a base station, the power consumption of the NSA device and SA device is almost the same. A similar conclusion is drawn for the connected mode, where only necessary and limited data is received and transmitted from the device.

For the data transmission mode, a simulation is conducted, and the bandwidth of 100 MHz is considered for both NSA and SA. UL data transmission with necessary DL feedback and vice versa are considered. For the poor coverage scenario, the device is assumed to transmit with the maximum power (23 dBm for LTE and 26 dBm for 5G NR). The same amount of data is transmitted, and the power consumption of different device is tested in milliampere and compared in Fig. 4.

It is concluded that the SA device consumes less power than the NSA terminal in DL service transmission mode. Although 2Tx of SA device leads to higher data rate in a good coverage scenario and shorter transmission duration, two independent power amplifiers (PAs) of 2Tx consume more working power (10 mA) than that of 1Tx. For the poor coverage case, the PA works with maximum power; 2Tx means double power consumption. So 1Tx outperforms 2Tx of SA on power consumption. For NSA, the power consumption of LTE baseband processing and RF power consumption have to be considered together with NR's power consumption. Hence, it is reasonable that NSA has higher power consumption.

For the UL data transmission, NSA has similar power consumption as 2Tx of SA, but much higher power consumption than 1Tx of SA. This can be explained similarly by the two PAs of two radio links and the power consumption of LTE baseband processing.

From the consumer behavior perspective, a heavy DL is typical. Hence, an SA device outperforms an NSA device on power consumption

4G/5G Interworking

In general, the initial deployment of 5G NR cannot provide seamless coverage compared to the legacy 4G network, which has been optimized for many years. To guarantee the user experience

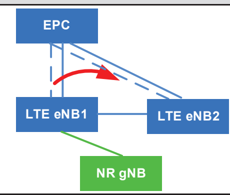
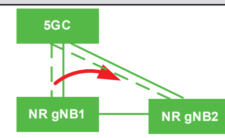
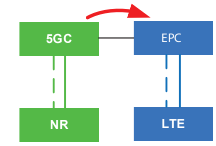
Scenario categories	1) NSA	2) SA	
Scenarios			
Handover latency	<100ms	<100ms	400~500ms
Service interruption	20ms	0~20ms	30~50ms
Duration of paging unreachable	0	0	350~450ms

Table 1. Handover performance analysis in different scenarios.

and service continuity, the interworking between 5G and 4G networks is necessary to provide service continuity. The interworking performance is also a very important aspect for the selection of SA and NSA.

For the NSA deployment, the device always anchors in the LTE network, and there is no 4G/5G interworking. Naturally, voice service can be supported by existing voice over LTE (VoLTE).

For the SA deployment, the device anchors in the 5G NR network when its coverage is available, while it hands over to the LTE network when it moves out of 5G's coverage. The interworking between 4G and 5G is performed through interface N26.

The voice service of SA could be provided in two ways. One is voice over NR (VoNR), and the second is EPS fallback. For VoNR, the service is served by NR directly, and when the device moves out of the coverage of the 5G NR network, the voice is handed over to the LTE network. For VoNR, the voice service and the other 5G service can be supported simultaneously. For EPS fallback, the device falls back to the LTE network when the voice is initiated, and accordingly the other 5G service will be interrupted. Hence, EPS fallback cannot support voice and 5G data services simultaneously.

To do a valuable comparison, the latency and interruption time caused by UE movement in different NSA and SA scenarios are compared in Table 1:

- 1. UE moves from one LTE eNB to another LTE eNB in the NSA scenario where intra-EPC system handover happens.
- 2A. UE moves from one SA NR coverage to another SA NR coverage in the SA scenario where intra-5GC system handover happens.
- 2B. UE moves from SA NR coverage to LTE coverage where inter-system handover happens.

It is concluded that NSA performs better because the interworking of the NSA system is done as intra-system handover of LTE. The performance of SA is similar as that of LTE's CS fallback, which is acceptable for a human customer.

Complexity of Network Deployment

In the operator community, there is a consensus that SA would be the target network evolution direction since only SA NR can offer E2E 5G capabilities.

For SA NR deployment, a new mobile network is built independently, and only an N26 interface is introduced for interworking between 4G and 5G. Hence, SA NR has very little impact on the legacy network. The supplier of the 5G network could be decoupled from that of the legacy network, which facilitates the operator's price negotiation when selecting the supplier.

For NSA deployment, no new core network is introduced, but the legacy EPC of 4G will have to be updated for 5G NR's access, and the capacity should be expanded. In order to introduce 5G NR, a hardware upgrade to LTE eNB is necessary. To guarantee the good performance of dual connectivity, the same vendor for 5G and 4G networks is preferred, which means it will be very difficult for the operator to bargain with vendors. The investment in EPC modification and capacity expansion cannot be reused for further evolution from NSA to SA NR, where the new 5GC based on the software defined networking/network function virtualization (SDN/NFV) platform will be introduced. Besides, the gNBs of NSA have to be further upgraded when NSA is transited to SA.

In summary, SA enables the deployment of one step for all, while NSA deploys the 5G network in two steps. Since NSA deployment is only an interim stage of 5G's development, its entire deployment will be much more complicated than that of SA, which also implies a higher total cost from the CAPEX point of view.

Progress on NSA NR and SA NR

Recently, sub-6 GHz has been the focus of 5G initial deployment due to its good coverage and maturity of key components and chipsets.

To promote the maturity of the E2E 5G system, lab trials and field trials have been conducted by global operators [13-14]. The pre-commercial products of NSA NR (e.g., base station, terminal chipset, and device) matured by 2019, and the operators of China and South Korea launched their 5G commercial network in 2019. The commercial infrastructure of SA NR (e.g., the 5GC) is making fast progress and was ready for commercial launch by the third quarter of 2020.

Previous research by Gartner reveals that 66 percent of commercial organizations planned to deploy 5G by 2020 [15]; now 46 operators worldwide have made 5G commercially available and another 79 operators have announced plans to launch mobile services [2]. It is expected that initial 5G deployments will mostly be NSA NR as this approach enables operators to make use of their existing 4G LTE/EPC network, with a software upgrade [6]. Widespread commercial 5G services are expected in the post-2020 period, which will mark the start of the 5G era [6].

Conclusion

5G is designed to explore the new business opportunities of eMBB, mMTC, and URLLC scenarios. Five alternatives for the SA NR and NSA NR network architecture have been proposed to meet the demand of diverse deployment strategies. In this article, option 2 and option 3 are compared from the operator's perspective. To facilitate the understanding of which one should be selected as the first step, the coverage, network capabilities, device performance, 4G/5G interworking, and complexity of network deployment are analyzed in detail. From the analysis, it is proved that SA outperforms NSA in terms of device power consumption, network deployment complexity, and cost. NSA option 3 has some advantages in 4G/5G interworking and CAPEX for the initial deployment. For SA, the interworking performance of 4G/5G is similar to that of LTE's CS fallback, which is acceptable for human customers. For operators who have the ambition to explore the vertical and enterprise markets as soon as possible, SA is recommended for large-scale deployment. For operators who are not willing to introduce the 5GC for initial NR deployment, NSA is a compromise between adding new capacity for eMBB and saving some cost of initial deployment. When it is necessary, the NSA NR can be evolved to SA NR.

References

[1] IMT-2020 Vision and Requirement white paper; http://www.imt-2020.org.cn/.

[2] GSMA, "The Mobile Economy 2020."

[3] Ericsson, "5G Deployment Considerations"; https://www.ericsson.com/en, accessed July 5, 2020.

[4] 3GPP RP-180554, "Plan for Finalizing All NR Architecture Options," RAN Chairman, RAN1/2/3/4 Chairmen.

[5] 3GPP TR 38.801, "Radio Access Architecture and Interfaces," V2.0.0 Release 14.

[6] GSMA, "5G Implementation Guidelines: NSA Option 3," Feb. 2020.

[7] 3GPP TR 38.817-01, "General Aspects for UE RF for NR."

[8] 3GPP TR 38.802, "Physical Layer Aspects," V 2.0.0 Release 14.

[9] 3GPP RP-191527, "Link Budget Template SRIT."

[10] 3GPP TR 23.501, "System Architecture for the 5G System," Release 15.

[11] F. Hu, Q. Hao, and K. Bao, "A Survey on Software-Defined Network and OpenFlow: From Concept to Implementation," *IEEE Commun. Surveys & Tutorials*, vol. 16, no. 4, 2014.

[12] Y. Li and M. Chen, "Software-Defined Network Function Virtualization: A Survey," *IEEE Access*, vol. 3, 2015.

[13] J. Gozalvez, "5G Worldwide Developments [Mobile Radio]," *IEEE Vehic. Tech. Mag.*, vol. 12, no. 1, 2017.

[14] M. Shafi *et al.*, "5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice," *IEEE JSAC*, vol. 35, no. 6, 2017.

[15] GSMA, "GSMA Future Networks 5G Global Launch Update: Revenue generated to hit $4.2 Billion by Next Year (Gartner)," Oct. 2019, accessed July 5, 2020.

Biographies

Guangyi Liu [M] (liuguangyi@chinamobile.com) received his Ph.D. degree from Beijing University of Posts and Telecommunications. He is currently a Fellow and CTO of the Wireless Department, China Mobile Research Institute (CMRI), where he is in charge of the wireless technology research and development, including 5G and 6G.

Yuhong Huang (huangyuhong@chinamobile.com) acts as the Deputy General Manager of CMRI, where she is responsible for the work on wireless technology and terminal technology. She has participated in technical and strategic research, specification, test, and trial of 3G/4G/5G.

Zhuo Chen (chenzhuo@chinamobile.com) received a Ph.D. from Beijing University of Posts and Telecommunications in 2006. He is currently with CMRI. His research interests include 4G/5G systems, RAN network architecture, radio interface, and protocol stack as well as signaling procedures in wireless and mobile systems.

Liang Liu (liuliang@chinamobile.com) received his Ph.D. degree in signal and information processing from Beihang University, China, in 2011. He is currently with CMRI as a delegate of 3GPP. His research interests include RAN architecture and radio protocols, network automation, and signal processing for LTE-A and 5G/5G+.

Qixing Wang (wangqixing@chinamobile.com) received his B.S and Ph.D. degrees from Beijing University of Posts and Telecommunications in 2004 and 2008, respectively. He is currently with CMRI. His research interests include massive MIMO, multiuser beamforming, and interference cancellation.

Na Li (linawx@chinamobile.com) received her M.S. degree from Beijing Institute of Technology in 2014. She is currently with CMRI. Currently, she is working on the architecture and protocol functions of the next-generation RAN.

For operators who have the ambition to explore the vertical and enterprise markets as soon as possible, SA is recommended for large-scale deployment. For operators who are not willing to introduce the 5GC for initial NR deployment, NSA is a compromise between adding new capacity for eMBB and saving some cost of initial deployment.