A Brief Introduction to Long Term Evolution (LTE) Architecture

Mohammad Reza Heidarpour and Mohammad Hossein Manshaei

March 18, 2020

1 LTE Entities

As we discussed during the lectures, Figure 1 shows the overall 4G network architecture. LTE network can be at the highest level divided into *radio access network* (RAN), and the *core network* or *evolved packet core* (EPC) network. RAN is the frontier of an LTE network which is responsible for sending/receiving to/from mobile *user equipments* (UEs) through wireless links. On the other hand, EPC bridges RAN to Internet or other IP networks.

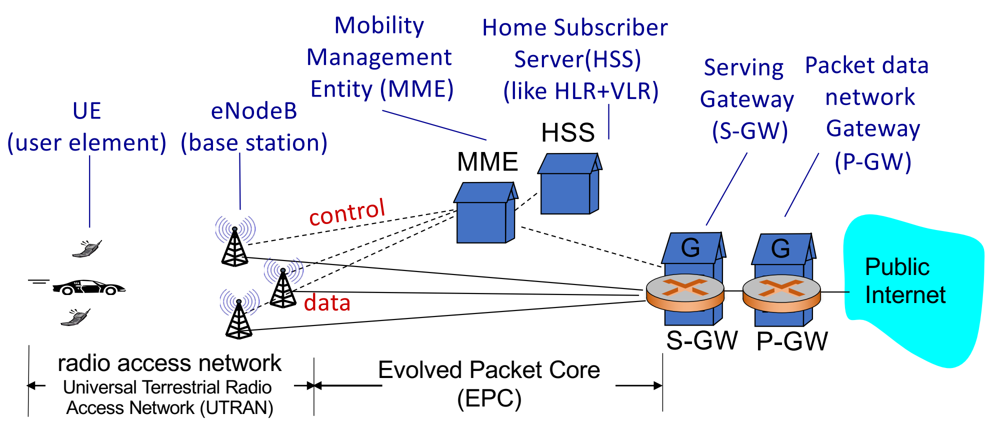

Figure 1: 4G Network Architecture.

1.1 Radio Access Entities

Radio access components of an LTE network are the user equipments (UE), base stations' antennas, and base-stations' cabinets.

UE is composed of the radio interface and the *universal integrated circuit card* (UICC). Radio interface is the L1/L2 LTE wireless bridge and implements the required protocol stacks (*NAS*(*EMM*, *ESM*), *RRC, PDCP, RLC, MAC*, and *PHY*). On the other hand, UICC is a smart card, colloquially known as the SIM card. Astonishingly, a UICC is much more than just a simple memory card as it contains a complete microcontroller system that can be used for a number of additional purposes. The typical properties of a UICC are shown in Fig 2.

EEPROM stores user-specific data such as the user's *International Mobile Subscriber Identity* (IMSI) and the home network identity. The mobile device cannot access the information on the EEPROM directly, but has to request the information from the UICC's CPU. Therefore, direct access to sensitive information is prohibited. The CPU is also used to generate encryption keys from the secret key (Ki) by executing an application known as the *universal subscriber identity module* (USIM). As a result, Ki is always kept inside UICC, and physically, not exposed to disclosure.

CPU	8- or 16-bit CPU
ROM	40–100 kB
RAM	1–3 kB
EEPROM	16–64 kB
Clock rate	10 MHz, generated from clock supplied by mobile device
Operating voltage	3 V or 5 V

Figure 2: UICC's properties

In LTE, UEs are categorized based on their capabilities such as number of antennas and supported modulation schemes. Category information is used to allow the eNB to communicate effectively with all the UEs connected to it (see Fig. 3).

UE category	Release	Maximum # DL bits per ms	Maximum # UL bits per ms	Maximum # DL layers	Maximum # UL layers	Support of UL 64-QAM?
1	R8	10 296	5 160	1	1	No
2	R8	51 024	25 456	2	1	No
3	R8	102 048	51 024	2	1	No
4	R8	150 752	51 024	2	1	No
5	R8	299 552	75 376	4	1	Yes
6	R10	301 504	51 024	4	1	No
7	R10	301 504	102 048	4	2	No
8	R10	2 998 560	1 497 760	8	4	Yes

Source: TS 36.306. Reproduced by permission of ETSI.

Figure 3: UE categories

The base station in the LTE context is referred to as eNB or eNodeB (*e* for "evolved" with respect to 3G, *Node* because the base station is inserted in a network, and *B* for Base station). eNodeB's antennas are tower mounted and route the signal to a cabinet for further processing in the downlink or up-converts/transmits signal in the uplink. The link used for the communication between the antenna and the cabinet is called the *front-haul* link.

✧*Beacon Channel*
Each base station broadcasts a signal (beacon) regularly which indicates its existence, and moreover, gives the characteristics of the network (eg., identity of the operator). By measuring the strength of the signal received, each terminal can indicate if it receives the base station well or not. UEs use this information to know to which eNB send their initial attachment request, and LTE eNBs gather this information from UEs to decide on suitable handover time (UE assisted network triggered handover).

✧*Capacity and Coverage*
The transmission power of a UE is typically 0.2 W (200 mW). The maximum range is typically several kilometers for a signal of this strength. As a result, the operator deploys base stations on the territory to be covered so that a terminal is always less than a few kilometers away from a base station. This is the reason behind the *cellular architecture* of the radio access networks in which the territory is divided up into *cells*, and each cell is served by a base station. In the urban zones (characterized by high user density), the base stations are deployed to provide sufficient *capacity* (i.e. deploying enough base stations so that the capacity in Mbit/s per km2 is superior to the traffic created by the customers). In the rural zones (characterized by low user density), the base stations are deployed to ensure *coverage* (i.e., deploying enough base stations so that, at every point of the territory, a terminal is under the range of a base station and can connect).

1.2 Core Network Entities

We summarize all building blocks of LTE and their relationships in Figure 4. The data packets gathered by eNBs can not directly be piped into Internet. One reason is that the Internet network cannot manage mobility. Therefore, The base stations are interconnected through a private IP network, deployed by the cell phone operator.

This network is connected to the Internet (or to the IP networks of other operators) through a "gateway": PGW, *packet data network gateWay*. The operator's private IP network is called *evolved packet core* (EPC) or the *core network* (CN). EPC is composed of 3GPP specific entities (such as PGW, *SGW*, *HSS*, and *MME*) connected to each other through an IP network of routers and switches. At one end, EPC is connected to the radio access network by eNBs, and at the other end, it is connected to Internet by PGWs.
Equipment of the EPC Participate in the Routing of Data (SGW, PGW)

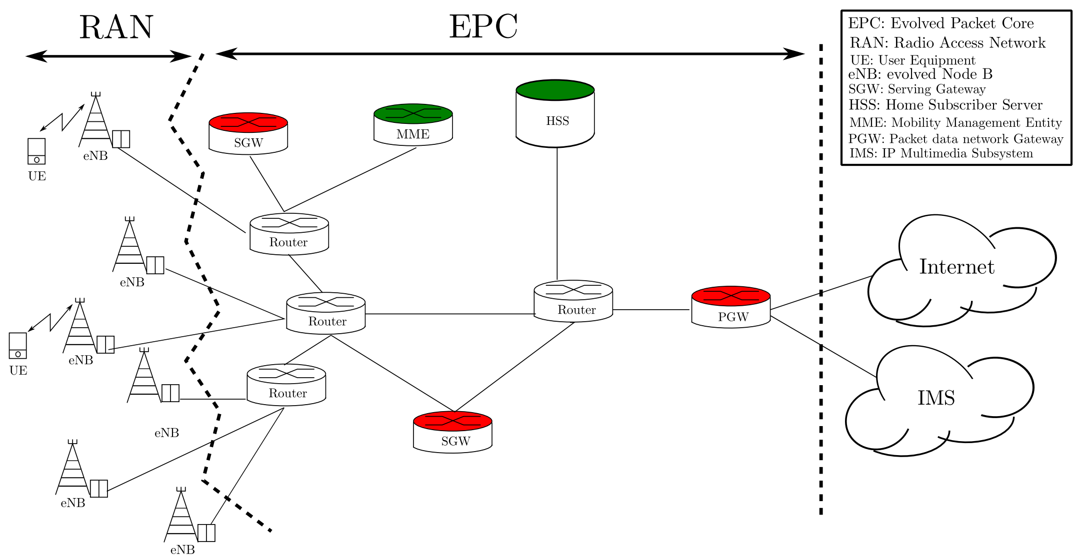

Figure 4: LTE RAN and EPC. The main building blocks and their relationships.

PGWs are the interface points of the EPC to the outside IP networks. The outside IP networks may be Internet or other private IP networks such as *IP multimedia subsystem* (IMS). PGWs establish tunnels to each UE connected to the EPC network. Incoming packets to the EPC, transmitted by a server in the outside world to a UE, are prioritized (tagged) and encapsulated to an IP packet with the destination address which is different from the UE's IP address. Indeed, this intermediate IP address is the IP address of a *serving gateway* (SGW) which UE resides in its geographical coverage. The IP packet is processed by SGW and re-encapsulated towards the eNB which has the radio connection with UE.

An SGW *service area* is an area served by one or more serving gateways, through which the mobile can move without a change of serving gateway. Every base station is connected to all the serving gateways in a service area by means of operator's private IP network. An operator network may have handful PGWs and several SGWs. The existence of SGWs, each for a specific geographical area and connecting to several eNBs, in the EPC architecture is to lower the burden from the PGWs.

As in the downlink, the uplink packets traverse tunnels (through encapsulation) between the eNB and SGWs, and between the SGWs and PGWs to reach to the Internet.

Control Equipment in the EPC (HSS, MME)

Before data can be transmitted by a terminal, there are several access and control procedures.The HSS, *Home Subscriber Server*, contains the users' identifications, their imprecise location information, and the list of accessible outside networks and the level of QoS should be provided for each user. HSS only exchanges control signaling. Mobility and the sporadic

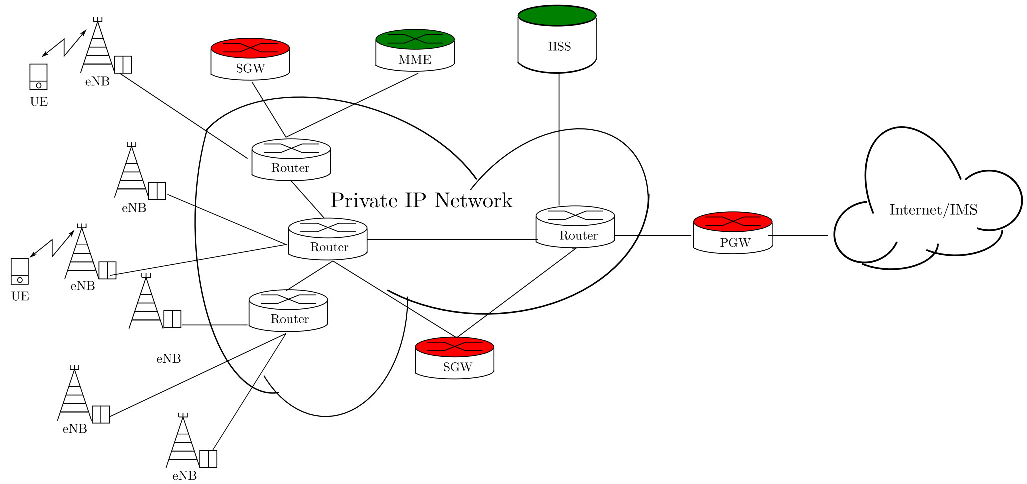

Figure 5: LTE Architecture

nature of terminal activity lead to sending (or receiving) frequent signaling by the terminals. Again to overcome scaling problems, geographical representative called *Mobility Management Entities* (MME) have been inserted into EPC architecture. MME plays a central role in EPC. It (1) communicates with a set of base stations, (2) communicates with the HSS to get the profiles and the security information of the subscribers present in the zone it manages, (3) stores these profiles and security data, manages control mechanisms related to network access, security and mobility for terminals present in its zone, (4) maintains awareness of the location of terminals in its area, (5) selects the PGW and the SGW when the terminal attaches to the network and connects to the Internet, (6) ensures the reachability of the terminal, and (7) is involved in handover (handoff)

An MME *pool area* is an area through which the mobile can move without a change of serving MME. Every pool area is controlled by one or more MMEs, while every base station is connected to all the MMEs in a pool area by operator's IP network. Pool areas can also overlap. Typically, a network operator might configure a pool area to cover a large region of the network such as a major city and might add MMEs to the pool as the signalling load in that city increases.

1.3 Synthesis of the Architecture and Interfaces

eNBs and EPC's entities are all IP connected to each other through the routers and switches of the operator's private IP network. This means that, for an arbitrary pair of nodes, they need to have only their IP addresses in order to communicate with each other, and a direct point-to-point link is not required to be in place (see Figure 5). On top of

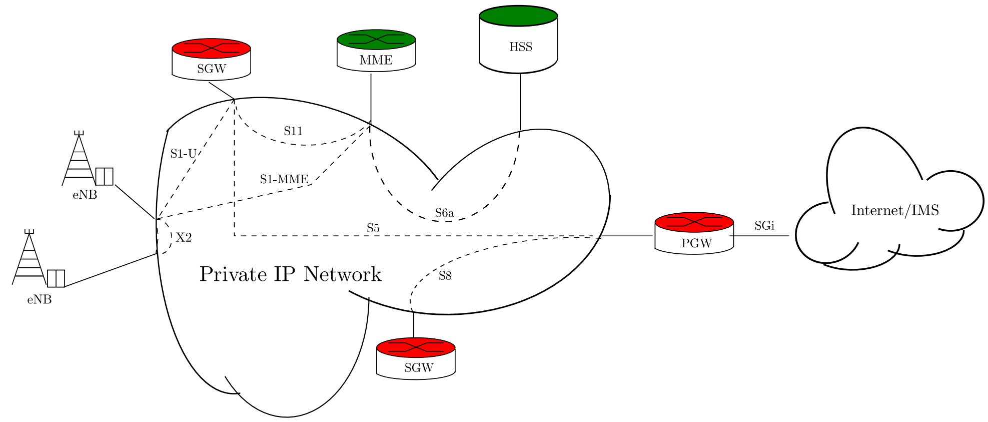

Figure 6: LTE Logical Interfaces

the IP connectivity, an overlay network is formed by defining different logical interfaces among different entities. Each interface is characterized by its end-point entities and the protocol stack used by them to transmit data/control packets to each other. For instance, the *S1-MME* is an interface between an eNB and an MME with a protocol stack of S1-ap (application layer), SCTP (transport layer), IP (network layer), and an arbitrary L2/L1 layer (such as wired Ethernet). All protocol stacks in all interfaces have the IP network layer and SCTP/UDP as their transport protocol of choice. Physically, packets sent and received via an interface may pass through many links and routers. In many cases, the application layers set tags for the packets belonging to different subscribers in order to make (forwarding/QoS provisioning) table inquiries more effective and faster. The list of main interfaces is as follows, and it is illustrated in Figure 6.

- SGi interface: between the PGW and the external IP network (Internet)
- S5 interface: between the SGW and the PGW (same network)
  Transporting user data + a few signaling messages
- S11 interface: between the SGW and the MME
  Transporting signaling messages
- S6a interface: between the MME and the HSS
  Transporting signaling messages
- S1-MME interface: between the eNodeB and the MME
  Transporting signaling messages
- S1-U interface: between the eNodeB and the SGW
  Transporting user data, no signaling exchanges

- X2 interface: between 2 eNodeBs
  Transporting user data and signaling messages
- Uu or radio interface: between the terminal (UE) and the eNodeB
  Transporting user data and signaling messages

2 Generations of Mobile Networks

2.1 1G

Mobile telecommunication systems were first introduced in the early 1980s. The first generation (1G) systems used analogue communication techniques(analogue radios).

Gen.	Principle Services	Name of the technology	Type of radio access	Lifetime
1	Telephony	AMPS, MNT	Analog FDMA	1980 - 1995

2.2 2G

Mobile telecommunications took off as a consumer product with the introduction of second generation (2G) systems in the early 1990s. These systems were the first to use digital technology, which permitted a more efficient use of the radio spectrum and the introduction of smaller, cheaper devices. They were originally designed just for voice, but were later enhanced to sup- port instant messaging through the Short Message Service (SMS). The most popular 2G system was the *Global System for Mobile Communications* (GSM), which was originally designed as a pan-European technology, but which later became popular throughout the world.

2.5G systems was built on the original ideas from 2G, by introducing the core network's packet switched domain and by modifying the air interface so that it could handle data as well as voice. The General Packet Radio Service (GPRS) incorporated these techniques into GSM. At the same time, the data rates available over the Internet were progressively increasing. To mirror this, designers first improved the performance of 2G systems using techniques such as Enhanced Data Rates for GSM Evolution (EDGE) and then introduced more powerful third generation (3G) systems in the years after 2000.

Gen.	Principle Services	Name of the technology	Type of radio access	Lifetime
2	Telephony, SMS	GSM	TDMA, CDMA	1995 -
2.5	Telephony, SMS IP access at 100 kbit/s	GPRS, EDGE	+ packet access new modulation	2000 -

2.3 3G

Since 1990, international mobile union has been publishing specifications for the future mobile technology generations. The first publication was released in 1990 called *Future Public Land Mobile Telecommunication Systems* (FPLMTS) or *International Mobile Telecommunications* (IMT)-2000 as the requirements to be fulfilled by 3G technologies.

The world's dominant 3G system is the *Universal Mobile Telecommunication System* (UMTS). UMTS was developed from GSM by completely changing the technology used on the air interface, while keeping the core network almost unchanged. The system was later enhanced for data applications, by introducing the 3.5G technologies of high-speed downlink packet access (HSDPA) and high-speed uplink packet access (HSUPA), which are collectively known as high-speed packet access (HSPA).

The UMTS air interface has two slightly different implementations. *Wideband code division multiple access* (WCDMA) is the version that was originally specified, and the one that is currently used through most of the world. *Time division synchronous code division multiple access* (TD-SCDMA) is a derivative of WCDMA, which is also known as the low chip rate option of UMTS TDD mode. TD-SCDMA was developed in China, to minimize the country's dependence on Western technology and on royalty payments to Western companies. It is deployed by one of China's three 3G operators, China Mobile.

Long term evolution (LTE) in its first introduction in 3GPP release 8 partially met the requirements set by IMT-advanced (the ITU vision for 4G). For instance, all-IP core network and scalable channel bandwidth 5–20 MHz are provided in LTE while the maximum bit rate is limited to 300 Mbit/s instead of 1 Gbit/s. Therefore, LTE can be categorized as pre-4 G or 3.9 G.

Gen.	Principle Services	Name of the technology	Type of radio access	Lifetime
3	Telephony, SMS, IP access at 1 Mbit/s	UMTS	CDMA	2002 -
3.5	Telephony, SMS, IP access at 10-30 Mbit/s	HSPA (3GPP Rel. 5/6) HSPA+ (3GPP Rel. 7)	CDMA, diversity, MIMO	2005 -
3.9	IP access at 300 Mbit/s with low latency	LTE (3GPP Rel 8/9)	OFDMA, SC-FDMA, MIMO	2008 -

2.4 4G

ITU defined 4G technology in IMT-advanced document in 2008. LTE-Advanced (LTE-A) presented in 3GPP release 10 demonstrate a true 4G system. Enabling technologies used in LTE-A are mainly carrier-aggregation and enhanced MIMO in downlink and uplink. In subsequent releases, further enhancements in techniques, such as interference management, carrier-aggregation, and small cells, resulted in performance improvement and interim 4.5/4.9 generations towards 5G's mobile broadband requirements. Along with mobile broadband, latest releases of 3GPP also address other use-cases of 5G era, such as machine-type communication (MTC), device-to-device, and Internet of things (IoT) among others.

Gen.	Principle Services	Name of the technology	Type of radio access	Lifetime
4	IP access at 1 Gbit/s	LTE-A (3GPP Rel. 10)	Carrier aggregation, enhanced MIMO, ICIC	2011 -
4.5	IP access at 3 Gbit/s, D2D, MTC	LTE-A (3GPP Rel. 11/12)	256QAM, CoMP, 3GPP/WiFi interworking	2013 -
4.9	IP access at 1+ Gbit/s per user (Gigabit LTE), Low latency (2ms), IoT	LTE-A Pro (3GPP Rel. 13/14)	Massive-MIMO, LTE-Unlicensed, NB-IoT	2015 -

2.5 5G

IUT vision on 5G is called IMT-2020 presented in 2015. The main characteristic of IMT-2020 is the definition of different use-cases (or scenarios). Accordingly, requirements are specified separately for different scenarios. For instance, for the case of *machine type communication* (MTC) high power efficiency, large number of users, and low latency are important performance merits while for the *broadband mobile* high data-rate is the crucial parameter. As the amount of resources to meet the requirements becomes prohibitively large for static network planning, 5G candidate solutions are mostly based on *softwarization*, *virtualization*, and *cloudification* techniques such as software-defined network (SDN), *network function virtualization* (NFV), and *software defined radio* (SDR). Considering the broadband mobile use-case, enablers for the extremely high data rates are multi-radio access technology (RAT) connectivity, mmWave and massive MIMO among others.

Broad scope of 5G has brought many organizations to the field each working on specific aspects. Following table names the most important participators.

Working Area	Organization/Technology
Use-cases, requirements, architectures	NGMN, IUT 2020 FG, ITU-R WP5D, 3GPP
Gbps transmission	Wired: ITU-T SG15 Wireless: ETSI mWT, IEEE 802.11.ay
Wireless access protocol functional splits	3GPP, IEEE 802.11, SCF
FH/BH traffic packetization (formatting)	Fronthaul: CPRI/eCPRI, NGFI(IEEE 1914.1) Backhaul VLAN (IEEE 802.1Q), MPLS
FH/BH switching protocol	IEEE 802.1CM TSN, IETF DETNET
SDN control	ONF, OpenDayLight, ONOS, IRTF, SD-NRG, ITU-T SG13, IEEE 802.1CF
NFV	ETSI NFV, IRTF NFVRG, OPNFV, OpenStack
Edge computing	ETSI MEC