
2.2 LTE ARCHITECTURAL OVERVIEW

EPS provides the user with IP connectivity to a PDN for accessing the Internet, as well as for running services such as VoIP. An EPS bearer is typically associated with a QoS. Multiple bearers can be established for a user in order to provide different QoS streams or connectivity to different PDNs. For example, a user might be engaged in a voice (VoIP) call while at the same time performing web browsing or File Transfer Protocol (FTP) download. A VoIP bearer would provide the necessary QoS for the voice call, while a best-effort bearer would be suitable for the web browsing or FTP session. The network must also provide sufficient security and privacy for the user and protection for the network against fraudulent use.

Release 9 of LTE introduced several additional features. To meet regulatory requirements for commercial voice, services such as support of IMS, emergency calls and UE positioning (see Chapter 19) were introduced. Enhancements to Home cells (HeNBs) were also introduced in Release 9 (see Chapter 24).

All these features are supported by means of several EPS network elements with different roles. Figure 2.1 shows the overall network architecture including the network elements and the standardized interfaces. At a high level, the network is comprised of the CN (i.e. EPC) and the access network (i.e. E-UTRAN). While the CN consists of many logical nodes, the access network is made up of essentially just one node, the evolved NodeB (eNodeB), which connects to the UEs. Each of these network elements is inter-connected by means of interfaces which are standardized in order to allow multivendor interoperability.

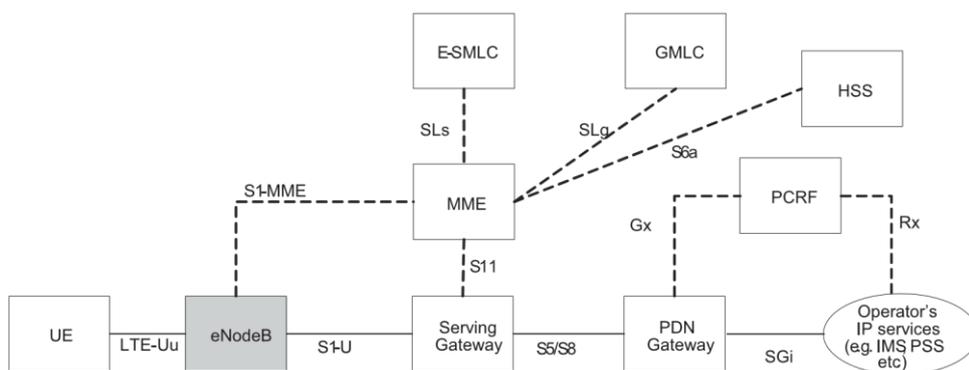


Figure 2.1: The EPS network elements.

The functional split between the EPC and E-UTRAN is shown in Figure 2.2. The EPC and E-UTRAN network elements are described in more detail below.

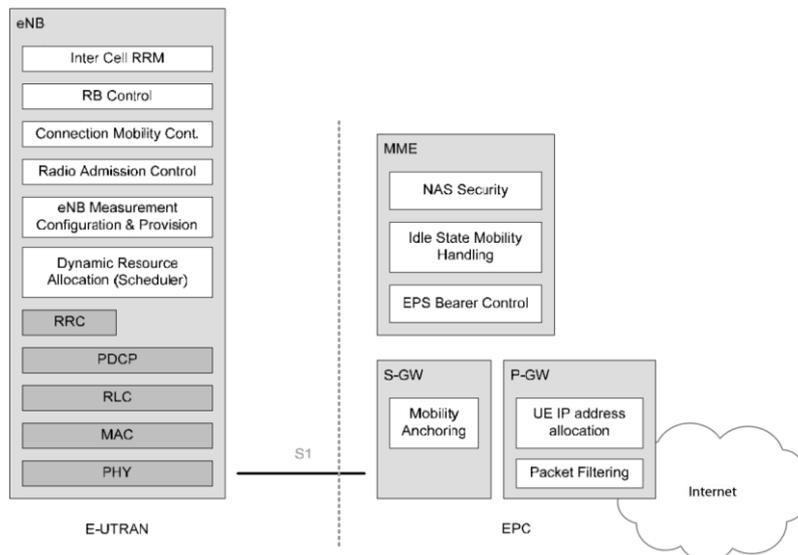


Figure 2.2: Functional split between E-UTRAN and EPC.

2.2.1 The Core Network

The CN (called the EPC in SAE) is responsible for the overall control of the UE and the establishment of the bearers. The main logical nodes of the EPC are:

- PDN Gateway (P-GW);
- Serving GateWay (S-GW);
- Mobility Management Entity (MME) ;
- Evolved Serving Mobile Location Centre (E-SMLC).

In addition to these nodes, the EPC also includes other logical nodes and functions such as the Gateway Mobile Location Centre (GMLC), the Home Subscriber Server (HSS) and the Policy Control and Charging Rules Function (PCRF). Since the EPS only provides a bearer path of a certain QoS, control of multimedia applications such as VoIP is provided by the IMS which is considered to be outside the EPS itself. When a user is roaming outside his home country network, the user's P-GW, GMLC and IMS domain may be located in either the home network or the visited network. The logical CN nodes (specified in [1]) are shown in Figure 2.1 and discussed in more detail below.

- PCRF. The PCRF is responsible for policy control decision-making, as well as for controlling the flow-based charging functionalities in the Policy Control Enforcement Function (PCEF) which resides in the P-GW. The PCRF provides the QoS authorization (QoS class identifier and bit rates) that decides how a certain data flow will be treated in the PCEF and ensures that this is in accordance with the user's subscription profile.
- GMLC. The GMLC contains functionalities required to support LoCation Services (LCS). After performing authorization, it sends positioning requests to the MME and receives the final location estimates.
- Home Subscriber Server (HSS). The HSS contains users' SAE subscription data such as the EPS-subscribed QoS profile and any access restrictions

for roaming (see Section 2.2.3). It also holds information about the PDNs to which the user can connect. This could be in the form of an Access Point Name (APN) (which is a label according to DNS¹ naming conventions describing the access point to the PDN), or a PDN Address (indicating subscribed IP address(es)). In addition, the HSS holds dynamic information such as the identity of the MME to which the user is currently attached or registered. The HSS may also integrate the Authentication Centre (AuC) which generates the vectors for authentication and security keys (see Section 3.2.3.1).

- P-GW. The P-GW is responsible for IP address allocation for the UE, as well as QoS enforcement and flow-based charging according to rules from the PCRF. The P-GW is responsible for the filtering of downlink user IP packets into the different QoS-based bearers. This is performed based on Traffic Flow Templates (TFTs) (see Section 2.4). The P-GW performs QoS enforcement for Guaranteed Bit Rate (GBR) bearers. It also serves as the mobility anchor for inter-working with non-3GPP technologies such as CDMA2000 and WiMAX networks (see Section 2.4.2 and Chapter 22 for more information about mobility).
- S-GW. All user IP packets are transferred through the S-GW, which serves as the local mobility anchor for the data bearers when the UE moves between eNodeBs. It also retains the information about the bearers when the UE is in idle state (known as EPS Connection Management IDLE (ECM-IDLE), see Section 2.2.1.1) and temporarily buffers downlink data while the MME initiates paging of the UE to re-establish the bearers. In addition, the S-GW performs some administrative functions in the visited network, such as collecting information for charging (e.g. the volume of data sent to or received from the user) and legal interception. It also serves as the mobility anchor for inter-working with other 3GPP technologies such as GPRS² and UMTS³(see Section 2.4.2 and Chapter 22 for more information about mobility).
- MME. The MME is the control node which processes the signalling between the UE and the CN. The protocols running between the UE and the CN are known as the *Non-Access Stratum* (NAS) protocols.

The main functions supported by the MME are classified as:

Functions related to bearer management. This includes the establishment, maintenance and release of the bearers, and is handled by the session management layer in the NAS protocol.

Functions related to connection management. This includes the establishment of the connection and security between the network and UE, and is handled by the connection or mobility management layer in the NAS protocol layer.

¹ Domain Name System.

² General Packet Radio Service.

³ Universal Mobile Telecommunications System.

NAS control procedures are specified in [1] and are discussed in more detail in the following section.

Functions related to inter-working with other networks. This includes handing over of voice calls to legacy networks and is explained in more detail in Section 2.4.2.

- E-SMLC. The E-SMLC manages the overall coordination and scheduling of resources required to find the location of a UE that is attached to E-UTRAN. It also calculates the final location based on the estimates it receives, and it estimates the UE speed and the achieved accuracy. The positioning functions and protocols are explained in detail in Chapter 19.

2.2.1.1 Non-Access Stratum (NAS) Procedures

The NAS procedures, especially the connection management procedures, are fundamentally similar to UMTS. The main change from UMTS is that EPS allows concatenation of some procedures so as to enable faster establishment of the connection and the bearers.

The MME creates a *UE context* when a UE is turned on and attaches to the network.

It assigns to the UE a unique short temporary identity termed the SAE-Temporary Mobile Subscriber Identity (S-TMSI) which identifies the UE context in the MME. This UE context holds user subscription information downloaded from the HSS. The local storage of subscription data in the MME allows faster execution of procedures such as bearer establishment since it removes the need to consult the HSS every time. In addition, the UE context also holds dynamic information such as the list of bearers that are established and the terminal capabilities.

To reduce the overhead in the E-UTRAN and the processing in the UE, all UE-related information in the access network can be released during long periods of data inactivity. The UE is then in the ECM-IDLE state. The MME retains the UE context and the information about the established bearers during these idle periods.

To allow the network to contact an ECM-IDLE UE, the UE updates the network as to its new location whenever it moves out of its current Tracking Area (TA); this procedure is called a 'Tracking Area Update'. The MME is responsible for keeping track of the user location while the UE is in ECM-IDLE.

When there is a need to deliver downlink data to an ECM-IDLE UE, the MME sends a paging message to all the eNodeBs in its current TA, and the eNodeBs page the UE over the radio interface. On receipt of a paging message, the UE performs a service request procedure which results in moving the UE to the ECM-CONNECTED state. UE-related information is thereby created in the E-UTRAN, and the bearers are re-established. The MME is responsible for the re-establishment of the radio bearers and updating the UE context in the eNodeB. This transition between the UE states is called an 'idle-to-active transition'. To speed up the idle-to-active transition and bearer establishment,

EPS supports concatenation of the NAS and AS⁴ procedures for bearer activation (see also Section 2.4.1). Some inter-relationship between the NAS and AS protocols is intentionally used to allow procedures to run simultaneously, rather than sequentially as in UMTS. For example, the bearer establishment procedure can be executed by the network without waiting for the completion of the security procedure.

Security functions are the responsibility of the MME for both signalling and user data. When a UE attaches with the network, a mutual authentication of the UE and the network is performed between the UE and the MME/HSS. This authentication function also establishes the security keys which are used for encryption of the bearers, as explained in Section 3.2.3.1. The security architecture for SAE is specified in [2].

The NAS also handles IMS Emergency calls, whereby UEs without regular access to the network (i.e. terminals without a Universal Subscriber Identity Module (USIM) or UEs in limited service mode) are allowed access to the network using an ‘Emergency Attach’ procedure; this bypasses the security requirements but only allows access to an emergency P-GW.

2.2.2 The Access Network

The access network of LTE, E-UTRAN, simply consists of a network of eNodeBs, as illustrated in Figure 2.3. For normal user traffic (as opposed to broadcast), there is no centralized controller in E-UTRAN; hence the E-UTRAN architecture is said to be flat.

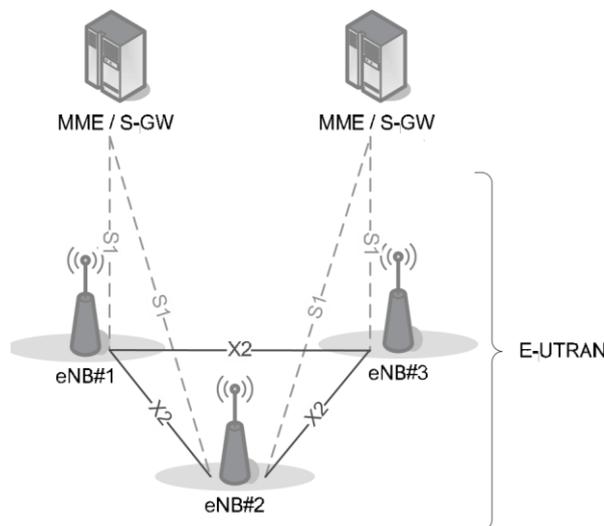


Figure 2.3: Overall E-UTRAN architecture. Reproduced by permission of © 3GPP.

The eNodeBs are normally inter-connected with each other by means of an interface known as X2, and to the EPC by means of the S1 interface – more

⁴ Access Stratum – the protocols which run between the eNodeBs and the UE.

specifically, to the MME by means of the S1-MME interface and to the S-GW by means of the S1-U interface.

The protocols which run between the eNodeBs and the UE are known as the *Access Stratum* (AS) protocols.

The E-UTRAN is responsible for all radio-related functions, which can be summarized briefly as:

- Radio Resource Management. This covers all functions related to the radio bearers, such as radio bearer control, radio admission control, radio mobility control, scheduling and dynamic allocation of resources to UEs in both uplink and downlink.
- Header Compression. This helps to ensure efficient use of the radio interface by compressing the IP packet headers which could otherwise represent a significant overhead, especially for small packets such as VoIP (see Section 4.2.2).
- Security. All data sent over the radio interface is encrypted (see Sections 3.2.3.1 and 4.2.3).
- Positioning. The E-UTRAN provides the necessary measurements and other data to the E-SMLC and assists the E-SMLC in finding the UE position (see Chapter 19).
- Connectivity to the EPC. This consists of the signalling towards the MME and the bearer path towards the S-GW.

On the network side, all of these functions reside in the eNodeBs, each of which can be responsible for managing multiple cells. Unlike some of the previous second- and third-generation technologies, LTE integrates the radio controller function into the eNodeB. This allows tight interaction between the different protocol layers of the radio access network, thus reducing latency and improving efficiency. Such distributed control eliminates the need for a high-availability, processing-intensive controller, which in turn has the potential to reduce costs and avoid ‘single points of failure’. Furthermore, as LTE does not support soft handover there is no need for a centralized data-combining function in the network.

One consequence of the lack of a centralized controller node is that, as the UE moves, the network must transfer all information related to a UE, i.e. the UE context, together with any buffered data, from one eNodeB to another. As discussed in Section 2.3.1.1, mechanisms are therefore needed to avoid data loss during handover. The operation of the X2 interface for this purpose is explained in more detail in Section 2.6.

An important feature of the S1 interface linking the access network to the CN is known as *S1-flex*. This is a concept whereby multiple CN nodes (MME/S-GWs) can serve a common geographical area, being connected by a mesh network to the set of eNodeBs in that area (see Section 2.5). An eNodeB may thus be served by multiple MME/S-GWs, as is the case for eNodeB#2 in Figure 2.3. The set of MME/S-GW nodes serving a common area is called an *MME/S-*

GW pool, and the area covered by such a pool of MME/S-GWs is called a *pool area*. This concept allows UEs in the cell(s) controlled by one eNodeB to be shared between multiple CN nodes, thereby providing a possibility for load sharing and also eliminating single points of failure for the CN nodes. The UE context normally remains with the same MME as long as the UE is located within the pool area.

2.2.3 Roaming Architecture

A network run by one operator in one country is known as a Public Land Mobile Network (PLMN). Roaming, where users are allowed to connect to PLMNs other than those to which they are directly subscribed, is a powerful feature for mobile networks, and LTE/SAE is no exception. A roaming user is connected to the E-UTRAN, MME and S-GW of the visited LTE network. However, LTE/SAE allows the P-GW of either the visited or the home network to be used, as shown in Figure 2.4. Using the home network's P-GW allows the user to access the home operator's services even while in a visited network. A P-GW in the visited network allows a 'local breakout' to the Internet in the visited network.

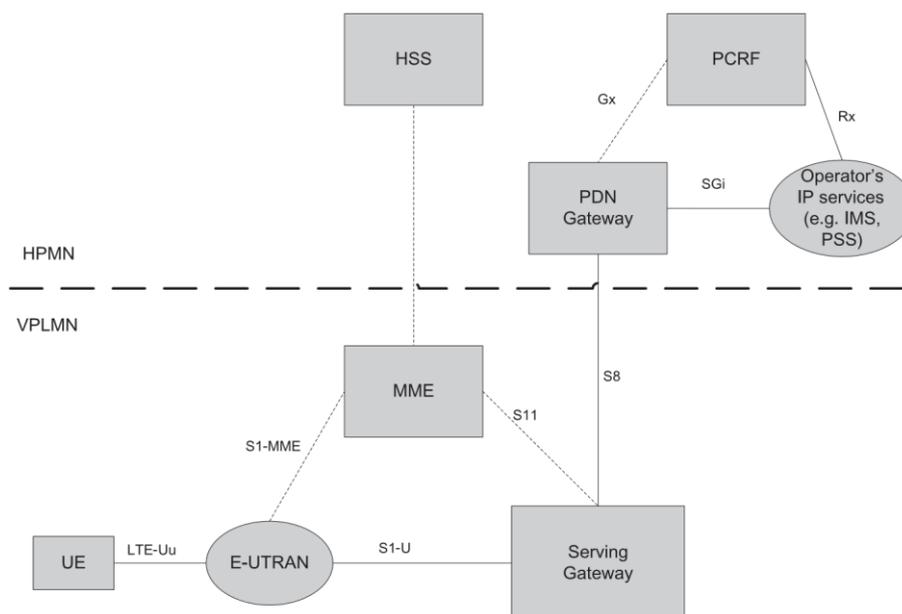


Figure 2.4: Roaming architecture for 3GPP accesses with P-GW in home network.

2.3 Protocol Architecture

We outline here the radio protocol architecture of E-UTRAN.

2.3.1 User Plane

An IP packet for a UE is encapsulated in an EPC-specific protocol and tunnelled between the P-GW and the eNodeB for transmission to the UE. Different tunnelling protocols are used across different interfaces. A 3GPP-specific tunnelling protocol called the GPRS Tunnelling Protocol (GTP) [4] is used over the core network interfaces, S1 and S5/S8.⁵

The E-UTRAN user plane protocol stack, shown greyed in Figure 2.5, consists of the Packet Data Convergence Protocol (PDCP), Radio Link Control (RLC) and Medium Access Control (MAC) sublayers which are terminated in the eNodeB on the network side. The respective roles of each of these layers are explained in detail in Chapter 4.

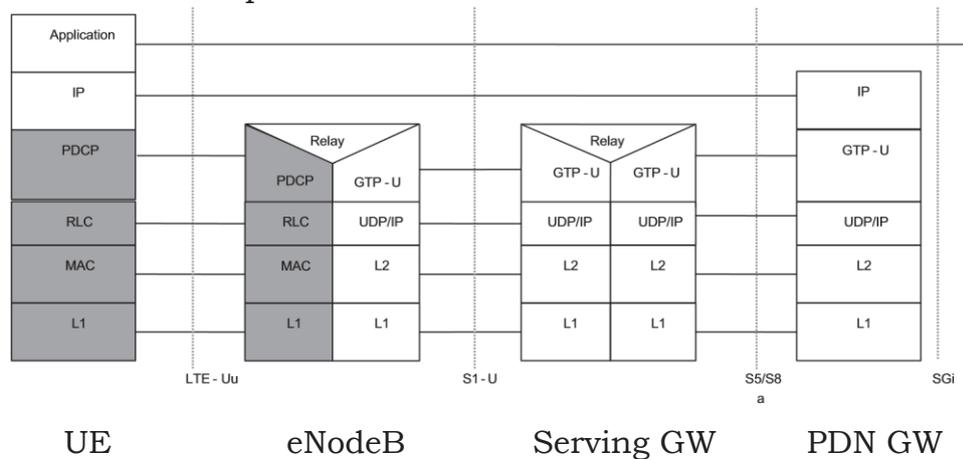


Figure 2.5: The E-UTRAN user plane protocol stack. Reproduced by permission of © 3GPP.

2.3.1.1 Data Handling During Handover

In the absence of any centralized controller node, data buffering during handover due to user mobility in the E-UTRAN must be performed in the eNodeB itself. Data protection during handover is a responsibility of the PDCP layer and is explained in detail in Section 4.2.4.

The RLC and MAC layers both start afresh in a new cell after handover is completed.

2.3.2 Control Plane

The protocol stack for the control plane between the UE and MME is shown in Figure 2.6.

⁵ SAE also provides an option to use Proxy Mobile IP (PMIP) on S5/S8. More details on the MIP-based S5/S8 interface can be found in [3].

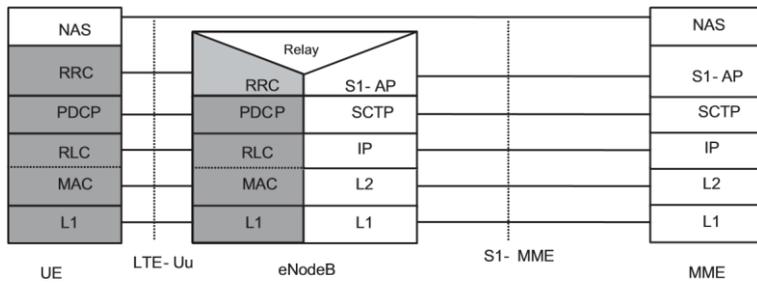


Figure 2.6: Control plane protocol stack. Reproduced by permission of © 3GPP.

The greyed region of the stack indicates the AS protocols. The lower layers perform the same functions as for the user plane with the exception that there is no header compression function for control plane.

The Radio Resource Control (RRC) protocol is known as ‘Layer 3’ in the AS protocol stack. It is the main controlling function in the AS, being responsible for establishing the radio bearers and configuring all the lower layers using RRC signalling between the eNodeB and the UE. These functions are detailed in Section 3.2.