Long-term evolution network architecture

Conference Paper · December 2009

DOI: 10.1109/COMCAS.2009.5385947 · Source: IEEE Xplore

1 author:

Ronit Nossenson

Long-Term Evolution Network Architecture

Ronit Nossenson

RNWC, Kfar-Sava, 44405, Israel, ronit.nossenson@gmail.com

***Abstract* — This paper provides a technological overview of the System Architecture Evolution (SAE) of LTE networks. The target of this overview is to provide a basic knowledge on 4G cellular network structure, entities responsibilities and protocol-stack. The overview is mainly based on the relevant 3GPP standards [2] and on Chapter 3 of the book "LTE for UMTS: OFDMA and SC-FDMA Based Radio Access" edited by Harri Holma and Antti Toskala [1].**

***Index Terms* — LTE, SAE.**

I. Introduction

Cellular operators are competing traditional broadband operators by offering mobile broadband access and IP services such as rich multimedia (e.g., video-on-demand, music download, video sharing) to laptops, PDAs, smart-phones and other advanced handsets. They offer these services through access networks such as High-Speed Packet Access (HSPA), Evolution-Data Optimized (EV-DO) and soon, Long-Term Evolution (LTE). These access networks promise to deliver performance comparable to today's ADSL services, but with the added benefits of mobility and ubiquitous coverage. The new technologies offer mobile operators significantly improved data speeds, short latency and increased capacity.

LTE is the next major step in mobile radio communications, and is introduced in 3rd Generation Partnership Project (3GPP) Release 8. LTE uses Orthogonal Frequency Division Multiplexing (OFDM) as its radio access technology, together with advanced antenna technologies.

When the evolution of the radio interface started, it soon became clear that the system architecture would also need to be evolved. Therefore, in addition to LTE, 3GPP is also defining IP-based, flat network architecture as presented in Figure 1. In the User Plane (UP) of the Evolved Packet System (EPS), for instance, there are only two types of nodes (Base Stations and Gateways); while in current hierarchical networks there are four types (Node B, RNC, SGSN, GGSN). The gateway consists of two logical UP entities, Serving Gateway and Packet Data Network Gateway (PDN-GW), collectively called the SAE-GW. Flat architecture with less involved nodes reduces latencies and improves performance.

Another simplification is the separation of the Control Plane (CP), with a separate mobility-management network element. A key difference from current networks is that the EPS is defined to support packet-switched traffic only.

This architecture is defined as part of the System Architecture Evolution (SAE) effort. The LTE–SAE

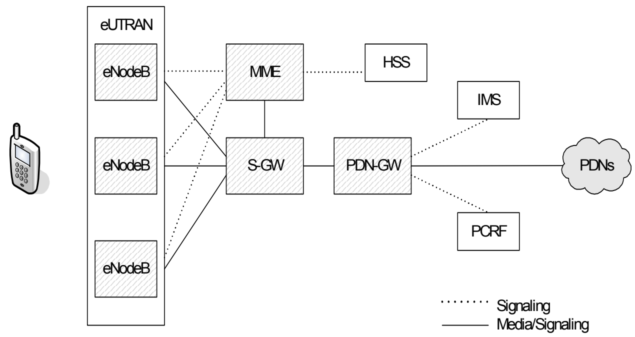

Fig. 1. LTE High-Level Network Architecture.

architecture and concepts have been designed for efficient support of mass-market usage of any IP-based service. The architecture is based on an evolution of the existing GSM/WCDMA core network, with simplified operations and smooth, cost-efficient deployment.

Moreover, work was initiated between 3GPP and 3GPP2 (the CDMA standardization body) to optimize inter-working between CDMA and LTE–SAE. This means that CDMA operators will be able to evolve their networks to LTE–SAE and enjoy the economies of scale and global chipset volumes that have been such strong benefits for GSM and WCDMA.

II. Basic System Architecture Configuration

This section introduces the logical network elements for the Basic System Architecture configuration.

A. User Equipment (UE)

UE is the device that the end user applies for communication. Typically it is a hand held device such as a smart phone or a data card such as those used currently in 2G and 3G, or it could be embedded, e.g. to a laptop. UE also contains the Universal Subscriber Identity Module (USIM) that is a separate module from the rest of the UE, which is often called the Terminal Equipment (TE). USIM is an application placed into a removable smart card called the Universal Integrated Circuit Card (UICC). USIM is used to identify and authenticate the user and to derive security keys for protecting the radio interface transmission.

Functionally the UE is a platform for communication applications, which signal the network to set up, maintain and remove the communication links the end user needs. This includes mobility management functions such as handovers and reporting the terminals location, and in these the UE performs as instructed by the network. Maybe most importantly, the UE provides the user interface to the end user.

B. E-UTRAN Node B (eNodeB)

The only node in the Evolved Universal Terrestrial Radio Access (eUTRAN) is the eUTRAN Node-B (eNodeB). It is a radio base station that is in control of all radio related functions in the fixed part of the system. Typically, the eNodeBs are distributed throughout the networks coverage area, each residing near the actual radio antennas.

A noteworthy fact is that most of the typical protocols implemented in today's Radio Network Controller (RNC) are moved to the eNodeB. The eNodeB is also responsible for header compression, ciphering and reliable delivery of packets. On the control plane, functions such as admission control and radio resource management are also incorporated into the eNodeB. Benefits of the RNC and Node-B merger include reduced latency with fewer hops in the media path, and distribution of the RNC processing load.

C. Mobility Management Entity (MME)

The Mobility Management Entity (MME) is a signalling-only entity, thus user's IP packets do not go through the MME. Its main function is to manage the users mobility. In addition, the MME also performs authentication and authorization; idle-mode user tracking and reaching abilities; security negotiations; and Network-Architecture Specific (NAS) signalling. An advantage of a separate network element for signalling is that operators can grow signalling and traffic capacity independently.

D. Serving Gateway (S-GW)

In the basic system architecture configuration, the high level function of S-GW is tunnel management and switching of the UP. The S-GW is part of the network infrastructure maintained centrally in operation premises.

The S-GW has a very minor role in control functions. It is only responsible for its own resources, and it allocates them based on requests from other network entities, such as MME, PDN-GW, or PCRF which in turn are acting on the need to set up, modify or clear bearers for the UE. If the request was received from the PDN-GW or PCRF, the S-GW will also relay the command on to the MME so that it can control the tunnel to eNodeB. Similarly, when the MME initiated the request, the S-GW will signal on to either the PDN-GW or the PCRF.

During mobility between eNodeBs, the S-GW acts as the local mobility anchor. The MME commands the S-GW to switch the tunnel from one eNodeB to another. The MME may also request the S-GW to provide tunneling resources for data forwarding, when there is a need to forward data from source eNodeB to target eNodeB during the time UE makes the radio handover. The mobility scenarios also include changing from one S-GW to another, and the MME controls this change accordingly, by removing tunnels in the old S-GW and setting them up in a new S-GW.

For all data flows belonging to a UE in connected mode, the S-GW relays the data between eNodeB and PDN-GW. However, when a UE is in idle mode, the resources in eNodeB are released, and the data path terminates in the S-GW. If S-GW receives data packets from PDN-GW on any such tunnel, it will buffer the packets, and request the MME to initiate paging of the UE. Paging will cause the UE to re-connect, and when the tunnels are re-connected, the buffered packets will be sent on. The S-GW will monitor data in the tunnels, and may also collect data needed for accounting and user charging.

E. PDN Gateway (PDN-GW)

Packet Data Network Gateway (PDN-GW) is the edge router between the EPS and external packet data networks. It is the highest level mobility anchor in the system, and usually it acts as the IP point of attachment for the UE. It performs traffic gating and filtering functions as required by the service in question. Similarly to the S-GW, the PDN-GWs are maintained in operator premises in a centralized location.

Typically the PDN-GW allocates the IP address to the UE, and the UE uses that to communicate with other IP hosts in external networks, e.g. the internet. It is also possible that the external PDN to which the UE is connected allocates the address that is to be used by the UE, and the PDN-GW tunnels all traffic to that network. The IP address is always allocated when the UE requests a PDN connection, which happens at least when the UE attaches to the network, and it may happen subsequently when a new PDN connectivity is needed. The PDN-GW performs the required Dynamic Host Configuration Protocol (DHCP) functionality, or queries an external DHCP server, and delivers the address to the UE. Also dynamic auto-configuration is supported by the standards. Only IPv4, only IPv6 or both addresses may be allocated depending on the need, and the UE may signal whether it wants to receive the address(es) in the Attach signalling, or if it wishes to perform address configuration after the link layer is connected.

The PDN-GW performs gating and filtering functions as required by the policies set for the UE and the service in question, and it collects and reports the related charging information.

The UP traffic between PDN-GW and external networks is in the form of IP packets that belong to various IP service flows. If the interface towards S-GW is based on tunneling, the PDN-GW performs the mapping between the IP data flows to tunnels, which represent the bearers. The PDN-GW sets up bearers based on request, either through the PCRF or from the

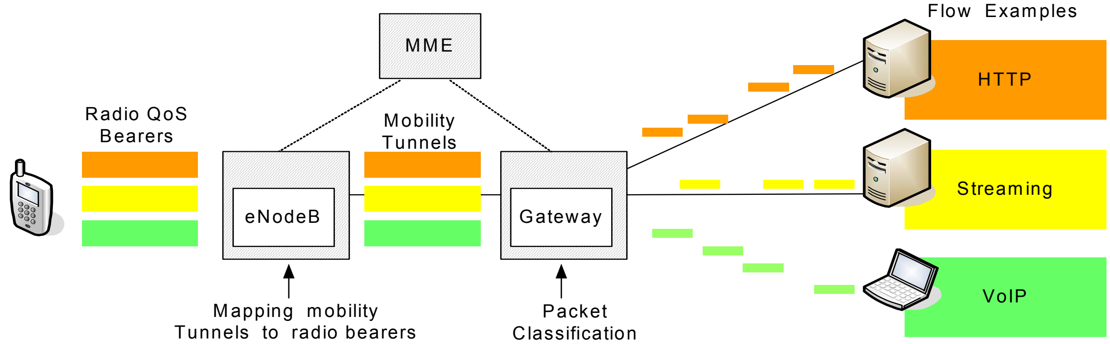

Fig. 2. QoS Procedure [4]

S-GW, which relays information from the MME. In the latter case, the PDN-GW may also need to interact with the PCRF to receive the appropriate policy control information, if that is not configured in the PDN-GW locally. The PDN-GW also has functionality for monitoring the data flow for accounting purposes.

PDN-GW is the highest level mobility anchor in the system. When a UE moves from one S-GW to another, the bearers have to be switched in the PDN-GW. The PDN-GW will receive an indication to switch the flows from the new S-GW.

Each PDN-GW may be connected to one or more PCRF, S-GW and external network. For a given UE that is associated with the PDN-GW, there is only one S-GW, but connections to many external networks and respectively to many PCRFs may need to be supported, if connectivity to multiple PDNs is supported through one PDN-GW.

F. Policy and Charging Resource Function (PCRF)

The Policy and Charging Resource Function (PCRF) is the network element that is responsible for Policy and Charging Control (PCC). It makes decisions on how to handle the services in terms of QoS, and provides information to the PDN-GW, and if applicable also to the S-GW, so that appropriate bearers and policing can be set up. PCRF is a server usually located with other core network elements in the operator switching centers.

G. Home Subscription Server (HSS)

Home Subscription Server (HSS) is the subscription data repository for all permanent user data. It also records the location of the user in the level of visited network control node, such as MME. It is a database server maintained centrally in the home operator's premises.

The HSS stores the master copy of the subscriber profile, which contains information about the services that are applicable to the user, including information about the allowed packet data connections, and whether roaming to a particular visited network is allowed or not.

H. Services Domain

The Services domain may include various sub-systems, which in turn may contain several logical nodes. The following is a categorization of the types of services that will be made available, and a short description of what kind of infrastructure would be needed to provide them:

- IMS based operator services: The IP Multimedia Sub-system (IMS) is a service machinery that the operator may use to provide services using the Session Initiation Protocol (SIP) see [3].
- Non-IMS based operator services: The architecture for non-IMS based operator services is not defined in the standards. The operator may simply place a server into their network, and the UEs connect to that via some agreed protocol that is supported by an application in the UE. For example, a video streaming service.
- Other services not provided by the mobile network operator, e.g. services provided through the internet.

III. IP QoS Support

An important aspect for any packet network is a mechanism to guarantee differentiation of packet flows based on its QoS requirements. Applications such as video streaming, HTTP, or video telephony have special QoS needs, and should receive differentiated service over the network. With EPS, QoS flows (so-called EPS bearers) are established between the user and the PDN-GW. Each EPS bearer is associated with a QoS profile, composed of a radio bearer and a mobility tunnel, and the network can prioritize packets accordingly.

The QoS procedure for packets arriving from the Internet is as follows (see Figure 2). When receiving an IP packet, the PDN-GW performs packet classification based on parameters such as rules received, and sends it through the proper mobility tunnel. Based on the mobility tunnel, the eNodeB can map packets to the appropriate radio QoS bearer.

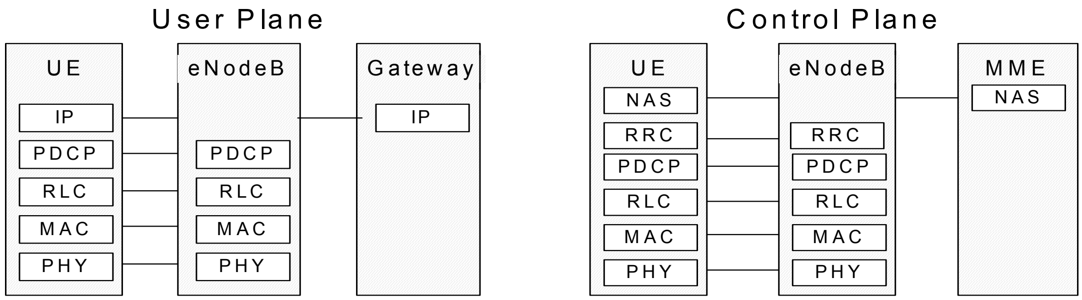

Fig. 3. LTE Simplified Protocol Stack [4]

IV. Interfaces and Protocols

Figure 3 shows the CP and UP protocols related to a UE's connection to a PDN.

The topmost layer in the CP is the Non-Access Stratum (NAS), which consists of two separate protocols that are carried on direct signalling transport between the UE and the MME. The content of the NAS layer protocols is not visible to the eNodeB, and the eNodeB is not involved in these transactions by any other means, besides transporting the messages, and providing some additional transport layer indications along with the messages in some cases. The NAS layer protocols are:

1) EPS Mobility Management (EMM): This protocol is responsible for handling the UE mobility. It includes functions for attaching to and detaching from the network, and performing location updating in between. Note that the handovers in connected mode are handled by the lower layer protocols, but the EMM layer does include functions for re-activating the UE from idle mode. The UE initiated case is called Service Request, while Paging represents the network initiated case. Authentication and protecting the UE identity, are also part of the EMM layer, as well as the control of NAS layer security functions, encryption and integrity protection.

2) EPS Session Management (ESM): This protocol may be used to handle the bearer management between the UE and MME, and it is used in addition for E-UTRAN bearer management procedures. Note that the intention is not to use the ESM procedures if the bearer contexts are already available in the network and E-UTRAN procedures can be run Application Function in the network, and the relevant information has been made available through the PCRF.

The radio interface protocols are:

- Radio Resource Control (RRC): This protocol is in control of the radio resource usage. It manages UE's signalling and data connections, and includes functions for handover.
- Packet Data Convergence Protocol (PDCP): The main functions of PDCP are IP header compression (UP), encryption and integrity protection (CP only).
- Radio Link Control (RLC): The RLC protocol is responsible for segmenting and concatenation of the PDCP-PDUs for radio interface transmission. It also performs error correction with the Automatic Repeat Request (ARQ) method.
- Medium Access Control (MAC): The MAC layer is responsible for scheduling the data according to priorities, and multiplexing data to Layer 1 transport blocks. The MAC layer also provides error correction with Hybrid ARQ.
- Physical Layer (PHY): This is the Layer 1 of LTE-Uu radio interface.

Regarding the UP, below the end user IP, the protocol structure is very similar to the CP. This highlights the fact that the whole system is designed for generic packet data transport and both CP signalling and UP data are ultimately packet data. Only the volumes are different.

References

[1] Harri Holma and Antti Toskala, *LTE for UMTS: OFDMA and SC-FDMA Based Radio Access* J.Wiley & Sons, 2009

[2] 3GPP TS 23.401, *General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access (Release 8).*

[3] M. Poikselk. et al., *The IMS: IP Multimedia Concepts and Services*, 2nd edition, Wiley, 2006.

[4] Evolved Packet System (EPS): An Overview of 3GPP's Network Evolution, Qualcomm white paper, 2007.