The LTE Link-Layer Design

Anna Larmo, Magnus Lindström, Michael Meyer, Ghyslain Pelletier, Johan Torsner, and Henning Wiemann, Ericsson Research

ABSTRACT

The LTE radio interface for 3GPP Release 8 was specified recently. This article describes the LTE link-layer protocols, which abstract the physical layer and adapt its characteristics to match the requirements of higher layer protocols. The LTE link-layer protocols are optimized for low delay and low overhead and are simpler than their counterparts in UTRAN. The state-of-the-art LTE protocol design is the result of a careful crosslayer approach where the protocols interact with each other efficiently. This article provides a thorough overview of this protocol stack, including the sub-layers and corresponding interactions in between them, in a manner that is more intuitive than in the respective 3GPP specifications.

INTRODUCTION

Following the successful standardization of high-speed packet access (HSPA), the 3rd Generation Partnership Project (3GPP) recently specified the universal mobile telecommunications system (UMTS) terrestrial radio-access network — or UTRAN — long term evolution (LTE) to meet the increasing performance requirements of mobile broadband. The result includes a flexible and spectrally efficient radio link protocol design with low overhead, which meets the challenging targets [1] that were set to ensure good service performance in varying deployments. The data rate can vary from more than 300 Mb/s in the downlink and 75 Mb/s in the uplink — for terminals in favorable radio conditions — to a few tens of kb/s at the cell edge, depending on the deployment scenario. The one-way latency target is set to be less than 5 ms between terminal and base station, and the handover mechanism supports real-time applications such as voice. The LTE architecture also should contribute to reducing the cost of network deployment, as discussed in the following section.

The high peak rates and the large range of possible data rates of the LTE physical layer, in combination with the strict latency requirements and the new, simplified architecture were the main challenges when designing the link-layer protocols.

In this article, we explain in detail the link-layer protocols that handle the data flow over the LTE radio interface, as well as crosslayer interactions that are required to realize the required functionality efficiently. Finally, we discuss the low header-overhead design of the LTE link-layer protocols and provide simulation results that illustrate the retransmission protocol design for high transmission control protocol (TCP) performance.

EVOLVED UTRAN ARCHITECTURE

The result of the 3GPP standardization effort is the evolved packet system (EPS) that consists of the core network part, the evolved packet core (EPC) and the radio network evolution part, the evolved UTRAN (E-UTRAN), also known as LTE. The EPC also can be connected to other 3GPP and non-3GPP radio-access networks. As illustrated in Fig. 1, the EPC consists of one control-plane node, called a mobility management entity (MME), and two user-plane nodes, called serving gateway (S-GW) and packet-data network gateway (P-GW). The LTE radio-access network consists of the base stations, denoted as enhanced NodeB (eNB), that are connected to each other through the X2 interface and to the EPC through the S1 interface. The mobile terminal is denoted as user equipment (UE).

The architecture in EPC/LTE, with only two user-plane nodes (eNB and S/P-GW),[1] is simpler than in UTRAN Release 6 with four nodes (NodeB, radio network controller [RNC], serving general packet radio service [GPRS] support node [SGSN], and gateway GPRS support node [GGSN]) and reduces the user-plane latency. One consequence is that some functionality performed by the RNC in UTRAN, such as ciphering and header compression, is performed by the eNBs in LTE. Further, handovers between eNBs are handled through packet forwarding over the X2 interface rather than by means of a central automatic repeat reQuest (ARQ) entity in the RNC as in UTRAN.

PHYSICAL LAYER CHARACTERISTICS

The properties of the physical layer determine to a large extent the characteristics of a cellular system regarding peak data rates, latencies, and coverage. Although the main focus of this article is on the link-layer protocols, an introduction to the main physical-layer properties is provided. A more detailed description can be found in [2, 3].

The LTE downlink uses conventional orthogonal frequency division multiplex (OFDM) due to the inherent robustness to time dispersion of

[1] P- and S-gateway are expected to be implemented as a common node divided into logical entities. The entities cannot share the same hardware only in the case of roaming.

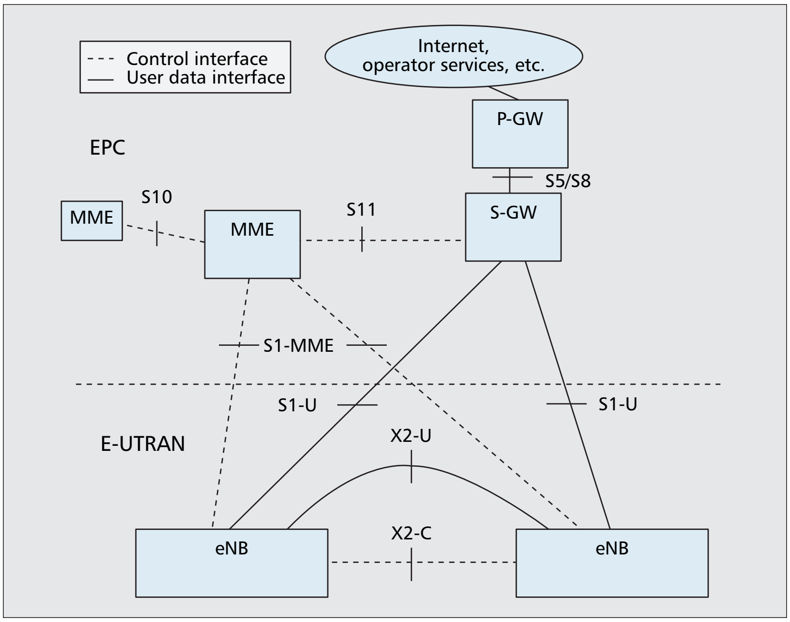
 0163-6804/09/$25.00 © 2009 IEEE

the radio channel. In addition to its advantages for a low-complexity receiver design, the multi-carrier concept enables the operation of LTE in various system bandwidths up to 20 MHz by adapting the number of subcarriers used to the allocated system bandwidth. Finally, OFDM supports multi-user access because within a transmission interval, subcarriers can be allocated to different users.

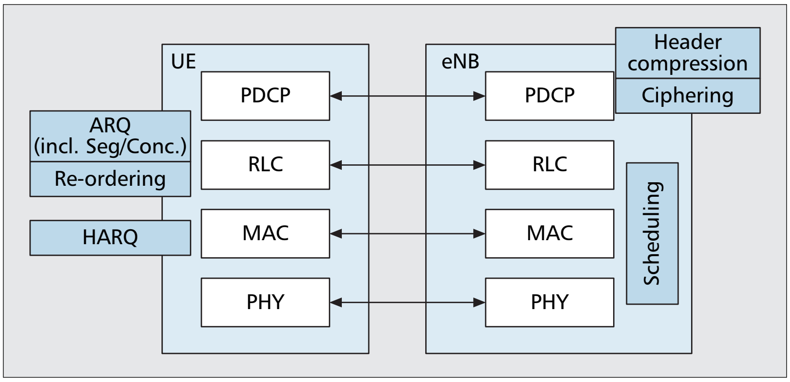
The LTE uplink employs a discrete Fourier transform (DFT)-spread OFDM (also denoted as single-carrier frequency division multiple access [SC-FDMA]). Compared to conventional OFDM, this OFDM variant provides an improved peak-to-average power ratio that enables more power-efficient terminals.

The basic LTE radio resource that is addressable for data transmission in the two-dimensional time-frequency grid is called a resource block. This type of resource block assembles 12 subcarriers and has a bandwidth of 180 kHz. In the time domain, the resource block has a subframe duration of only 1 ms. Such a short subframe enables the exploitation of channel variations by scheduling users depending on their current channel quality. At the same time, a short hybrid ARQ (HARQ) round-trip time of only 8 ms can be obtained.

By allocating a variable number of resource blocks to a certain user and selecting a modulation and coding scheme to meet the current channel conditions, widely scalable transport block sizes are possible, resulting in a wide range of user-data rates. In addition, it is possible to aggregate up to two streams by utilizing multiple-input multiple-output (MIMO) transmissions to increase the data rate even further under favorable radio conditions.



■ **Figure 1.** *Overview of the EPC/LTE architecture.*



■ **Figure 2.** *User plane protocol stack.*

USER PLANE PROTOCOL STACK

While the physical layer essentially provides a bit pipe, protected by turbo-coding and a cyclic redundancy check (CRC), the link-layer protocols enhance the service to upper layers by increased reliability, security, and integrity. In addition, the link layer is responsible for the multi-user medium access and scheduling.

One of the main challenges for the LTE link-layer design is to provide the required reliability levels and delays for Internet Protocol (IP) data flows with their wide range of different services and data rates. In particular, the protocol overhead must scale. For example, it is widely assumed that voice over IP (VoIP) flows can tolerate delays on the order of 100 ms and packet losses of up to 1 percent. On the other hand, it is well-known that TCP file downloads perform better over links with low bandwidth-delay products. Consequently, downloads at very high data rates (e.g., 100 Mb/s) require even lower delays and in addition, are more sensitive to IP packet losses than VoIP traffic.

Overall, this led to the following design of the LTE link layer: it consists of three sublayers (Fig. 2) that are partly intertwined. The Packet Data Convergence Protocol (PDCP) sublayer [4] is responsible mainly for IP header compression and ciphering. In addition, it supports lossless mobility in case of inter-eNB handovers and provides integrity protection to higher layer-control protocols. The radio link control (RLC) sublayer [5] comprises mainly ARQ functionality and supports data segmentation and concatenation. The latter two minimize the protocol overhead independent of the data rate, as is explained in more detail below. Finally, the medium access control (MAC) sublayer [6] provides HARQ and is responsible for the functionality that is required for medium access, such as scheduling operation and random access.

Figure 3 depicts the data flow of an IP packet through the link-layer protocols down to the physical layer. The figure shows that each protocol sublayer adds its own protocol header to the data units.

This figure is used in subsequent sections as a reference when relevant functionality is introduced.

RETRANSMISSION HANDLING

As in any communication system, there are occasional data transmission errors, for example, due to noise, interference, and/or fading. Link-layer, network-layer (IP), and transport-layer protocols are not prepared to cope with bit errors in head-

The two-layer ARQ design achieves low latency and low overhead without sacrificing reliability. Most errors are captured and corrected by the lightweight HARQ protocol. Only residual HARQ errors are detected and resolved by the more expensive (in terms of latency and overhead) ARQ retransmissions.

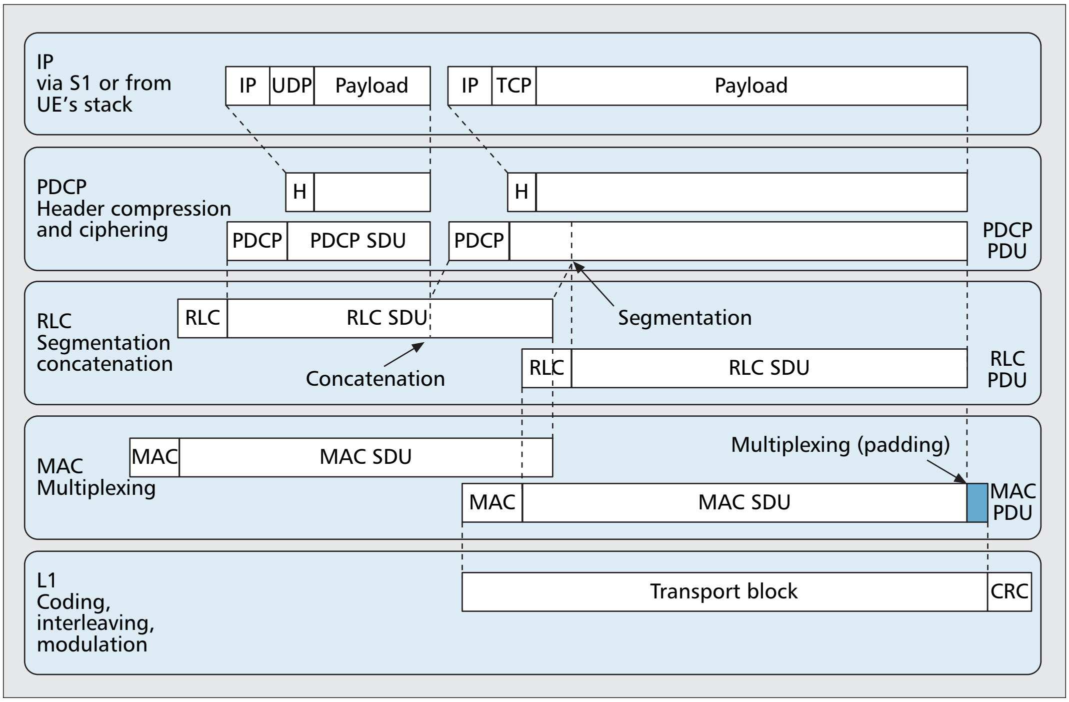


Figure 3. *Illustration of data flow through L2 protocol stack.*

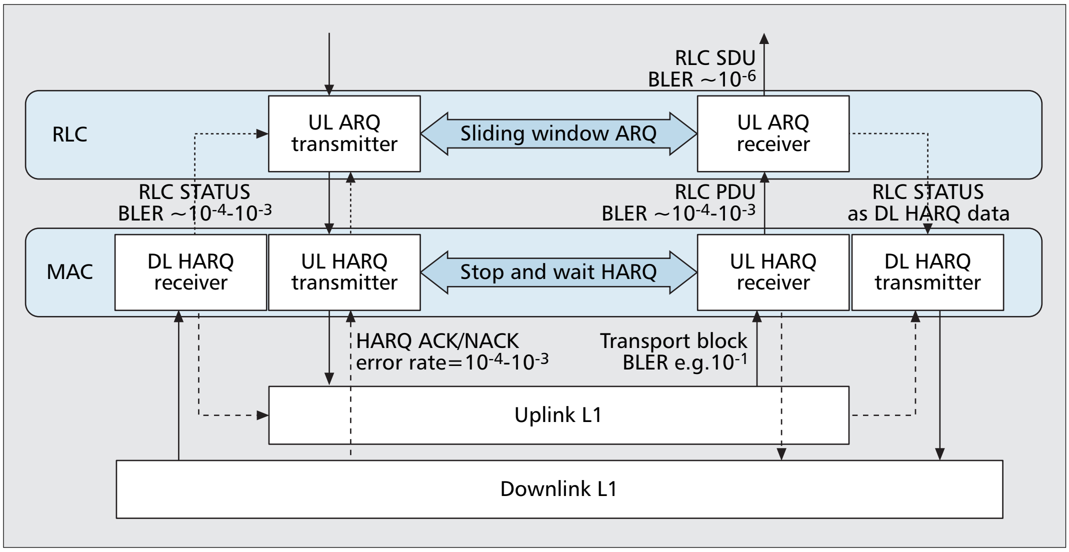
ers, and the majority of the protocols are not capable of handling errors in the payload either. Therefore, a fundamental design choice for LTE has been not to propagate any bit errors to higher layers but rather to drop or retransmit the entire data unit containing bit errors. As illustrated in Fig. 3, the physical layer attaches a 24-bit CRC checksum to the data units, thus allowing the receiver to detect bit errors and to forward only error-free packets to the IP layer.

Most TCP/IP protocols are designed to cope only with rather low packet-loss rates. State-of-the-art voice codecs perform well at error rates up to at most 10^{-2}. High-speed, TCP-based file downloads requires loss rates on the order of 10^{-5} to 10^{-6} [7]. The HARQ scheme on the MAC sublayer performs retransmissions of corrupted transport blocks and thereby corrects the majority of all transmission errors. The HARQ mechanism is very similar to the solution adopted for HSDPA [2], that is, the protocol uses multiple stop-and-wait HARQ processes. The functionality and performance is comparable to that of a window-based selective repeat protocol. In particular, it allows continuous transmission, which cannot be achieved with a single stop-and-wait scheme. Instead of a status message containing a sequence number, a single-bit HARQ feedback acknowledgment/negative acknowledgment (ACK/NACK), with a fixed-timing relation to the corresponding transmission attempt, provides information about the successful reception of the HARQ process. Thereby, it gains in terms of delay, simplicity, and control overhead compared to a window-based selective repeat protocol.

It is important that the HARQ protocol is fast yet consumes as few radio resources as possible. The single-bit HARQ feedback fulfils these requirements, but the probability for misinterpreting a negative acknowledgment as a positive acknowledgment, and thereby causing a residual packet loss, is in the order of 10^{-4} to 10^{-3}. It would be expensive, in terms of transmit power, to reduce the feedback error rate further and thereby, to ensure the desired very low residual loss rates required by TCP for achieving high-data rates. Furthermore, certain errors in other control signaling, such as scheduling information, result in HARQ failures. When such failures are detected by the receiver, the HARQ process typically has been re-used for new data, and the single-bit HARQ feedback is not a valid reference for the desired retransmission.

Due to the error cases mentioned above, the fast HARQ protocol with low-overhead, ACK/NACK feedback and retransmissions with incremental redundancy is complemented by a highly reliable window-based selective repeat-ARQ protocol that resides in the RLC sublayer as depicted in Fig. 4.

When the CRC check is successful, the MAC HARQ receiver delivers RLC protocol data units (PDUs) to the corresponding RLC entity. If the RLC receiver detects a gap in the sequence of received PDUs based on the RLC sequence number, it starts a reordering timer assuming that the missing packet still is being retransmitted in the HARQ protocol. Note that the reordering functionality required on top of a multi-process stop-and-wait mechanism reuses the same RLC sequence numbers as the ARQ mechanism, saving additional overhead compared to a sequence, number-based re-ordering mechanism in MAC, like in HSPA. In the rare case that the reordering timer expires, an RLC acknowledged-mode (AM) receiver sends a status message comprising the sequence number of the missing RLC PDU(s) to its transmitting peer entity. The MAC layer treats the RLC status message as any other data, meaning that it also applies the same HARQ operation and CRC to



■ **Figure 4** *HARQ and ARQ retransmissions on MAC and RLC layer.*

this message. Consequently, errors or loss of the ARQ feedback can be detected and recovered by sending another RLC status. Upon reception of the RLC status message, the ARQ transmitter triggers a retransmission of the corresponding RLC PDU(s). The HARQ layer does not attempt to combine the RLC retransmission with the previous transmission but treats it as new data.

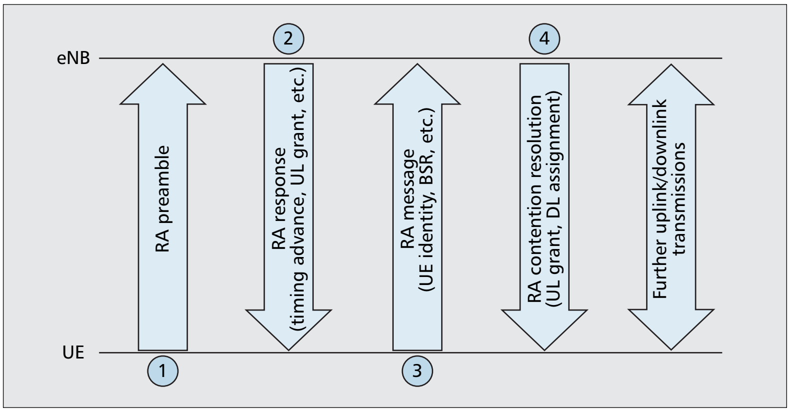
The fact that RLC and MAC are terminated in the same nodes enables a tighter interconnection of the two protocol sublayers. In LTE, the HARQ transmitter can indicate a so-called local NACK to the ARQ transmitter, if it knows or assumes that a HARQ transmission failed. The most prominent example is that the maximum number of HARQ transmission attempts was reached. This might happen, for example, if the chosen modulation and coding scheme is too weak for the given channel quality. The main benefit of the local NACK is the shorter detection delay, resulting in improved performance compared to the gap detection at the ARQ receiver.

The two-layer ARQ design achieves low latency and low overhead without sacrificing reliability. Most errors are captured and corrected by the lightweight HARQ protocol. Only residual HARQ errors are detected and resolved by the more expensive (in terms of latency and overhead) ARQ retransmissions.

Finally, it should be noted that services that can sustain error rates on the order of 10^{-3} to 10^{-2}, while gaining from reduced delays, can be mapped to a radio bearer running the RLC protocol in unacknowledged mode (UM), that is, without the second ARQ layer. In that case, residual errors on the MAC layer are not recovered, but packet losses propagate to higher layers. RLC UM generally is assumed to be used for VoIP and real-time gaming traffic.

Random Access

To keep transmissions from different UEs orthogonal, uplink transmissions in LTE are aligned with the frame timing at the eNB. When timing is not aligned yet or alignment was lost due to a period of inactivity during which time alignment was not maintained by the eNB, a random access (RA) procedure is performed to acquire time alignment.



■ **Figure 5.** *Contention-based random access procedure.*

The RA procedure establishes uplink-time alignment by means of a four-phase contention-based procedure outlined in the following and shown in Fig. 5:

1. RA Preamble — The UE randomly selects an RA preamble sequence from the set of sequences available in the cell and transmits it on an RA channel. A guard period is applied to the RA preamble transmission to avoid creating interference in adjacent subframes. To minimize non-orthogonal transmissions and thereby improve resource efficiency, unsynchronized and unscheduled transmissions, like the first step in the RA procedure, do not carry data.
2. RA Response — The eNB detects the preamble transmission, estimates the uplink transmission timing of the UE, and responds with an RA response providing the UE with the correct timing-advance value to be used for subsequent transmissions and with a first grant for an uplink

The SR mechanism is one of two types: dedicated SR (D-SR), where the SR is conveyed on a dedicated resource on the physical uplink-control channel (PUCCH), and random access-based SR (RA-SR), where the SR is indicated by performing an RA procedure.

transmission. For efficiency, RA responses pertaining to different RA preamble sequences can be multiplexed.

3. RA Message — Because the randomly selected RA preamble does not enable unique identification of the UE, and it is possible that multiple UEs attempted RA with the same RA preamble sequence on the same RA channel, the UE provides its identity to the eNB with the first scheduled uplink transmission. Space remaining in the transport block after including UE identification is used for data.
4. RA Contention Resolution — The eNB receives the RA message transmitted in phase 3; only one RA message is typically received even if two or more were transmitted by contending UEs. The eNB resolves the (potential) contention by echoing the received UE identity back. The UE, seeing its own identity echoed back, concludes that the RA was successful and proceeds with time-aligned operation.

UEs that do not receive an RA response or do not receive their own identity in the contention resolution, must repeat the RA procedure. In the case of congestion, the eNB can provide a back-off indicator to instruct UEs that did not succeed with their RA attempt to apply a back-off procedure. The back-off indicator is multiplexed with the RA responses.

Note that for cases where an RA is anticipated by the network, that is, at handover completion and eNB-triggered uplink re-alignment, LTE also provides a faster two-phase contention-free RA procedure. In this case the eNB assigns a dedicated preamble to be used by the UE. Because the UE that corresponds to the received dedicated preamble is known, phase 3 and 4 are not required.

Scheduling a Request and Buffer Status Report

To allow the UE to request uplink-transmission resources from the eNB, LTE provides a scheduling request (SR) mechanism. The SR conveys a single bit of information, indicating that the UE has new data to transmit. The SR mechanism is one of two types: dedicated SR (D-SR), where the SR is conveyed on a dedicated resource on the physical uplink-control channel (PUCCH), and random access-based SR (RA-SR), where the SR is indicated by performing an RA procedure. The D-SR is simpler than the RA-SR but assumes that the uplink of the UE already is time aligned. If the uplink of the UE is not time aligned, RA-SR must be used to (re-)establish time alignment. RA-SR also is used, regardless of the uplink-timing state, when no PUCCH resources for D-SR were assigned to the UE.

Because the SR procedure conveys little detail about the UE resource requirement, a buffer status report (BSR) with more detailed information about the amount of data waiting in the UE is attached to the first uplink transmission following the SR procedure. In fact, the requirement to transmit a BSR triggers the SR.

Scheduling

Because LTE is based on OFDM, it is possible to distribute available transmission resources in the frequency domain to different UEs. This allocation can be changed dynamically once per subframe, that is, once per millisecond. The MAC scheduler in the eNB is in charge of assigning both uplink and downlink radio resources. The scheduling decision covers not only the resource-block assignment but also which modulation and coding scheme to use and whether or not to apply MIMO or beamforming.

A particular challenge for the schedulers is to provide the desired quality of service (QoS) on a shared channel. Traditional mobile communication systems, such as the UMTS and the global system for mobile communications (GSM), provide guaranteed bit rates by pre-allocating radio resources statically to dedicated channels. LTE does not provide dedicated channels but only two shared channels, one in the uplink and one in the downlink. A number of default QoS characteristics, for example, suitable for VoIP, signaling traffic, and Internet access, have been standardized for EPS [8]. However, it is up to the eNB implementation and consequently, the responsibility of the scheduler to assign radio resources in a way that the terminals and radio bearers obtain the QoS characteristics assigned by the EPC.

Depending on the implementation, the scheduler can base its scheduling decision on the QoS class and the queuing delay of the available data, on the instantaneous channel conditions, or on fairness indicators. The channel conditions in a wideband system vary not only over time but also can differ in the frequency domain. If the UE provides sufficiently detailed channel-quality information to the eNB, the scheduler can perform channel-dependent scheduling in the time and frequency domain and thereby improve the cell and system capacity. Also, the physical downlink-control channel (PDCCH) that carries the scheduling decisions to the affected UE and the PUCCH that carries HARQ feedback and channel quality information to the eNB have a finite capacity and thus, may constrain the scheduler in its freedom of how many users to address in a subframe.

Finally, the scheduler must ensure that HARQ retransmissions are performed on a timely basis. In the uplink direction, the HARQ retransmission must occur exactly one round-trip time (i.e., 8 ms for frequency-division duplex [FDD]) after the previous transmission attempt, whereas the scheduler can postpone downlink retransmissions in favor of higher priority transmissions.

For the downlink, the scheduler selects not only the appropriate user but also decides which radio bearer to serve. In contrast, uplink scheduling grants are dedicated to particular UE but do not comprise instructions about which radio bearers to serve. This additional information would increase the size of the uplink grants and thereby limit the capacity of the PDCCH and consequently, the number of UE units that could be addressed in a subframe. Rather, the UE makes this decision autonomously in the logical channel prioritization function, which is preconfigured by the eNB. Furthermore, the UE sends BSRs for active radio bearers. Based on these reports, the eNB can ensure that users with high priority data are prioritized and obtain the assigned QoS characteristics. Not only user data but also control information, namely, MAC control elements such as BSR, and discontinuous reception (DRX) and timing advance messages can be chosen for transmission.

When generating the transport block, the MAC layer typically incorporates those MAC control elements first. Secondly, it triggers the scheduled RLC entities, which send either new RLC PDUs, or in the case of RLC AM, retransmissions. The size of a new RLC PDU is variable so that an RLC entity generates, at most, one new RLC PDU per transport block. This minimizes segmentation and multiplexing of data packets and consequently, protocol overhead. It also reduces the required RLC-sequence-number space and makes it independent of the data rate and future proof. The RLC re-segmentation function allows changing the size of RLC retransmissions if the scheduler did not provide sufficient resources to transmit the requested RLC PDU at once. Finally, if the size of the computed transport block does not exactly match the chosen transport format, the remaining bytes are filled with padding before handing the block to the physical layer for transmission.

DISCONTINUOUS RECEPTION

LTE supports DRX to enable UE power savings by turning off some or all of its radio circuitry, thereby increasing the battery lifetime of the UE. The DRX function is configured and controlled by the network. The UE behavior is based on a set of rules that define when the UE must monitor the PDCCH for scheduling assignments.

When the UE does not have an established radio-resource control (RRC) connection, that is, no radio bearers configured for data transmission, it wakes up and monitors the paging channel every DRX cycle. When the UE has an RRC connection, the DRX function is characterized by a DRX cycle, an on-duration period, and an inactivity timer. The UE wakes up and monitors the PDCCH at the beginning of every DRX cycle for the entire on-duration period. If no scheduling assignment is received, the UE falls asleep again. Whenever the UE receives an assignment from the network, it starts (or restarts) the inactivity timer and continues to monitor the PDCCH until the timer expires. Note that the HARQ operation overrides the DRX function. Thus, the UE wakes up for possible HARQ feedback, as well as for possible retransmissions during a configurable amount of time as soon as a retransmission can be expected.

Optionally, the network may configure the UE with two DRX cycles of different lengths, in which case, the UE moves to the longer cycle after a given period without receiving a scheduling assignment.

HANDOVER SUPPORT

In LTE, the UE performs measurements when radio conditions reach a certain configured threshold and provides measurement reports to the eNB it is connected to. The involved eNBs must negotiate through the X2 interface and decide whether or not to handover a UE to another cell or eNB. The EPC is not involved in the preparation signaling unless the change of serving cell involves an S1 handover as well.

In the case of inter-eNB handover, the source eNB prepares neighboring eNBs over the X2 interface and then transmits a handover command to the UE with the required information to perform the handover to one of the prepared eNBs.

The source eNB can forward data to the target eNB. For RLC AM data bearers, the source eNB forwards unacknowledged downlink PDCP PDUs with their sequence number (SN) and not-yet-transmitted IP packets received over the S1 interface to the target eNB. It also can forward the uplink-PDCP-service-data units (SDUs) received out-of-sequence to the target eNB. For RLC UM data bearers in the downlink, only the not-yet-transmitted IP packets received from the S1 interface are forwarded.

The PDCP layer ensures that no data is lost at handover for RLC AM bearers by retransmitting missing data. In the UE, duplicate removal and in-sequence delivery of the PDCP SDUs received from the source eNB and from the target eNB also are handled by the PDCP, based on the PDCP SN. For RLC UM, no data is retransmitted by the PDCP.

PERFORMANCE

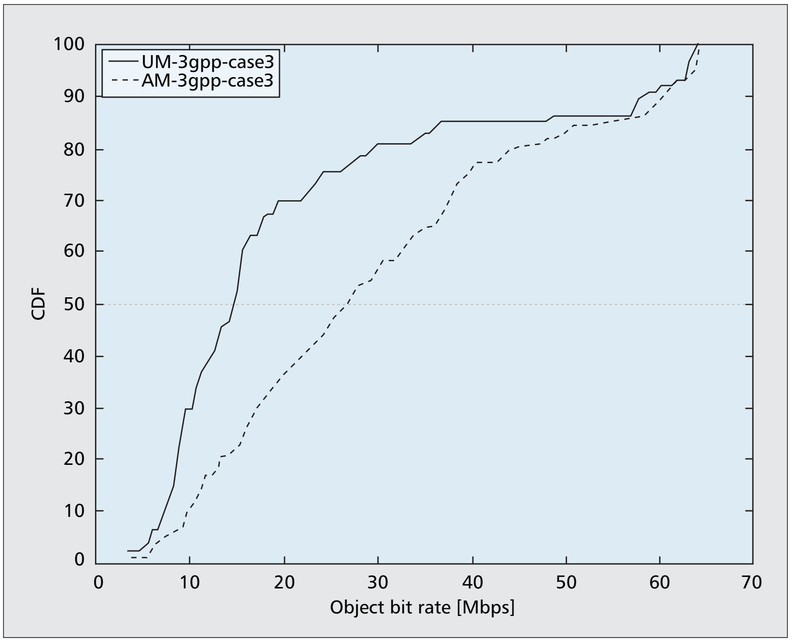
PROTOCOL OVERHEAD

Purely packet-switched access technologies, with dynamically scheduled access to a shared transmission resource, can offer exceptional performance for services with varying or adaptive demands in terms of throughput and delay. The crosslayer protocol design of LTE aims at low protocol overhead, while offering the desired flexibility. In this section, we provide an overview of the headers added on different LTE protocol layers (Fig. 3) and provide a comparison of the relative overhead for TCP and VoIP services.

Robust header compression (ROHC) is applied to an IP packet arriving at the transmitting PDCP entity. ROHC compresses the UDP, Real-time Transport Protocol (RTP), and IP headers to as little as 3–4 bytes; it can compress the TCP and IP headers to 8 bytes, for both IPv4 (40 bytes) and IPv6 (60 bytes). Assuming IPv4, for TCP services with typical packet sizes of 1500 bytes, this means a reduction in packet size by ~2.5 percent to about 1468 bytes. VoIP packets, using the wideband adaptive multirate (WB-AMR) codec mode of 12.65 kb/s, and TCP acknowledgments, for example, experience a more significant size reduction from 73 to 35 bytes (52 percent) and from 40 to 3 bytes (93 percent), respectively.

The PDCP and the RLC protocol each offer two SN lengths to optimize the header overhead for certain services. The lower values are used for low-rate services such as VoIP, whereas TCP/IP services require a larger SN space and make heavy use of concatenation, thus requiring more information in the headers. The relative overhead, however, is significantly lower for the

Purely packet-switched access technologies, with dynamically scheduled access to a shared transmission resource, can offer exceptional performance for services with varying or adaptive demands in terms of throughput and delay.



■ **Figure 6.** *Performance samples of consecutive file downloads with RLC AM and RLC UM.*

latter group of services as depicted in Table 1.

RLC AM bearers employ a two-layer ARQ functionality intended to provide low delays and very low residual-error rates while causing reasonably low overhead. In this section, we present simulation results showing the benefit of the second ARQ layer in terms of file transfer performance. We simulate a large number of file transfers of 100-MB files toward a UE connected to an LTE cell in a-10 MHz deployment with low load. The simulation settings are according to the 3GPP case 3 scenario [9], and a simplified 2×2 MIMO scheme is used. The theoretical peak-bit rate for this case is around 70 Mb/s. The HARQ block error-rate target is set to 10 percent. HARQ failures appear if the maximum number of five transmission attempts (used here) are exceeded or due to NACK-to-ACK errors. The residual loss rate above the MAC layer is on the order of 10^{-4}. With RLC UM, these errors propagate to higher layers and must be handled by TCP; whereas with RLC AM, they are recovered by the RLC protocol.

Figure 6 shows the performance of the object-bit rate (OBR), that is, the file size divided by the file transfer time, of the download for both AM and UM radio bearers as a cumulative distribution function (CDF). As expected, the AM radio bearer using RLC ARQ almost always achieves better performance than the UM radio bearer. Whereas the best 10 percent of the file transfers for RLC UM perform well, the majority suffers significantly from residual HARQ failures that trigger the TCP congestion control and thereby increase the transmission delay. The 50th percentile of the OBR decreases by 44 percent to 15 Mb/s. The actual performance of a file transfer over RLC UM depends on the number of losses seen by TCP and on the instantaneous state of the TCP congestion control upon occurrence of the loss.

The cost for the second ARQ layer is marginal due to the few residual HARQ errors requiring ARQ retransmissions. Also, the cost for regular RLC status reports is negligible in this case because they are typically multiplexed with TCP acknowledgments that are transmitted in the reverse direction in any case.

The motivation for offering an unacknowledged mode for RLC operation may not be

Protocol Layer	TCP/IP with RLC AM		VoIP with RLC UM AMR-WB codec 12.65 kb/s
	TCP segment	TCP ACK	
Application-, Transport-, and IP Layer	1460 + 40 bytes	40 bytes	33 + 40 bytes
PDCP ROHC	40 bytes to ~8 bytes	40 bytes to ~8 bytes	40 bytes to ~3 bytes
PDCP SDU	~1468 bytes	~8 bytes	~36 bytes
PDCP Header	2 bytes 12 bits SN		1 byte 7 bits SN
RLC Header	4 bytes 12 bits SN and framing sub-header for concatenation		1 byte 5 bits SN
MAC Header	1 byte		1 byte
L1 CRC	3 bytes		3 bytes
Total Overhead L2 + L1 on Shared Channel	10 bytes		6 bytes
Net overhead reduction including ROHC	22 of 1500 bytes (–1.5%)	22 of 40 bytes (–55%)	31 of 73 (–42%)

■ **Table 1.** *Protocol overhead in LTE assuming IPv4.*

straightforward when looking at these results. From the viewpoint of a VoIP optimization, the motivation is clear; running VoIP over an RLC AM bearer increases the header overhead, and more significantly, generates RLC status messages for almost every VoIP segment. Although a UE is typically not transmitting and receiving VoIP packets simultaneously, these control messages would lead to a bi-directional transmission causing additional control overhead, cell load, and inter-cell interference.

CONCLUSION

Within the 3GPP LTE specification process, a state-of-the-art link-layer protocol stack has been standardized. This article provides a comprehensive description of these LTE protocols, as well as the rationale for certain design decisions. A key characteristic of the LTE link layer is the tight interaction of the MAC and RLC protocols with a two-layer ARQ functionality and interactions between scheduling in MAC and segmentation in RLC. This close interworking resulted in a low overhead protocol-header design. Other highlights are the advanced sleep-mode feature (DRX) for the UE and the fast and lossless handover mechanism between base stations over a dedicated interface between eNBs.

The LTE link layer, as well as the entire LTE design, was optimized to meet the challenges and requirements from IP-based services ranging from low-rate real-time applications like VoIP to high-speed broadband access by providing high data rates and low delays combined with high reliability when required, for example, for TCP.

REFERENCES

[1] 3GPP TR 25.913, "Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN)."

[2] E. Dahlman *et al.*, *3G Evolution: HSPA and LTE for Mobile Broadband*, 2nd ed., Academic Press, 2008.

[3] D. Astély *et al.*, "LTE — The Evolution of Mobile Broadband," *IEEE Commun*. Mag., Apr. 2009.

[4] 3GPP TS 36.323, "E-UTRA Packet Data Convergence Protocol (PDCP) Specification," Rel. 8.

[5] 3GPP TS 36.322, "E-UTRA Radio Link Control (RLC) Protocol Specification," Rel. 8.

[6] 3GPP TS 36.321, "E-UTRA Medium Access Control (MAC) Protocol Specification," Rel. 8.

[7] M. Meyer *et al.*, "ARQ Concept for the UMTS Long-Term Evolution," *VTC Fall 2006*, Montreal, Canada, Sept. 2006.

[8] 3GPP TS 23.203, "Policy and Charging Control Architecture," Rel. 8.

[9] 3GPP TS 25.814, "Physical Layer Aspects for Evolved Universal Terrestrial Radio-Access (UTRA)," Rel. 7.

BIOGRAPHIES

ANNA LARMO (Anna.Larmo@ericsson.com) holds an M.S. degree in telecommunications from the Helsinki University of Technology (2005). She joined Ericsson to work in wireless access network research in 2004. Her research interests include radio-interface protocols and radio-system simulation-tools development for the WCDMA and LTE systems.

MAGNUS LINDSTRÖM (Magnus.Q.Lindstrom@ericsson.com) received an M.S. degree in electrical engineering from the Royal Institute of Technology, Stockholm, in 1997. He received his Lic.Eng. and Ph.D. degrees in radio communication systems from the Royal Institute of Technology in 2003 and 2005, respectively. He has been with Ericsson Research since 2005, involved in research on radio resource management and radio protocols. He is active in concept development and 3GPP standardization of LTE and future wireless technologies.

GHYSLAIN PELLETIER (Ghyslain.Pelletier@ericsson.com) holds an engineering degree in electrical engineering and telecommunication electronics from École Polytechnique de Montréal, Canada. He joined Ericsson Research in 2001 and has worked in research and standardization in the area of wireless access networks. He has been working with the development of algorithms and protocols to optimize the transport of IP packets. His research interests include robust IP header compression and the design of link-layer protocols in wireless systems. His previous positions include being a senior research engineer in the department of computer sciences and electrical engineering at the Luleå University of Technology, Sweden (1999–2001).

MICHAEL MEYER (Michael.Meyer@ericsson.com) received his diploma and doctoral degree in electrical engineering from the University of Paderborn, Germany, in 1991 and 1996, respectively. He has been the manager of the radio protocol research group at Ericsson Research in Aachen since 2008. He joined Ericsson in 1996. From 2000 to 2008, he held the position of a senior specialist for wireless protocol interactions. He is actively involved in the concept development and standardization of LTE.

JOHAN TORSNER (Johan.Torsner@ericsson.com) holds an M.S. in electrical engineering from the Royal Institute of Technology, Stockholm (1996). In 2009, he became manager of the Ericsson research activities in Finland. He worked as the manager of the Ericsson wireless access network research branch in Finland starting in 2004. He joined Ericsson in Sweden in 1998 to work on the development of the HIPERLAN/2 technology. In 2000, he moved to Ericsson Finland and since then has been active in 3GPP standardization and concept development for WCDMA, HSPA, LTE, and lately, LTE advanced.

HENNING WIEMANN (Henning.Wiemann@ericsson.com) received his diploma degree in electrical engineering from the University of Technology in Aachen, Germany in 2000. He currently contributes to the development and standardization of link-layer protocols for LTE at Ericsson. His other research interests cover quality of service and advanced scheduling algorithms in mobile broadband systems, and he contributes to the development of corresponding simulation tools. He joined the wireless access networks branch of Ericsson Research in 2000.

A key characteristic of the LTE link layer is the tight interaction of the MAC and RLC protocols with a two-layer ARQ functionality and interactions between scheduling in MAC and segmentation in RLC. This close interworking resulted in a low overhead protocol-header design.