

Technical Solutions for the 3G Long-Term Evolution

Hannes Ekström, Anders Furuskär, Jonas Karlsson, Michael Meyer, Stefan Parkvall,
Johan Torsner, and Mattias Wahlqvist, Ericsson

ABSTRACT

Work has started in the 3GPP to define a long-term evolution for 3G, sometimes referred to as Super-3G, which will stretch the performance of 3G technology, thereby meeting user expectations in a 10-year perspective and beyond. The fundamental targets of this evolution — to further reduce user and operator costs and to improve service provisioning — will be met through improved coverage and system capacity as well as increased data rates and reduced latency. This article presents promising technologies to fulfill these targets, including OFDM, multi-antenna solutions, evolved QoS and link layer concepts, and an evolved architecture. Furthermore, the results of a performance evaluation are presented, indicating that the requirements can indeed be reached using the proposed technologies.

BACKGROUND AND TARGETS FOR 3G EVOLUTION

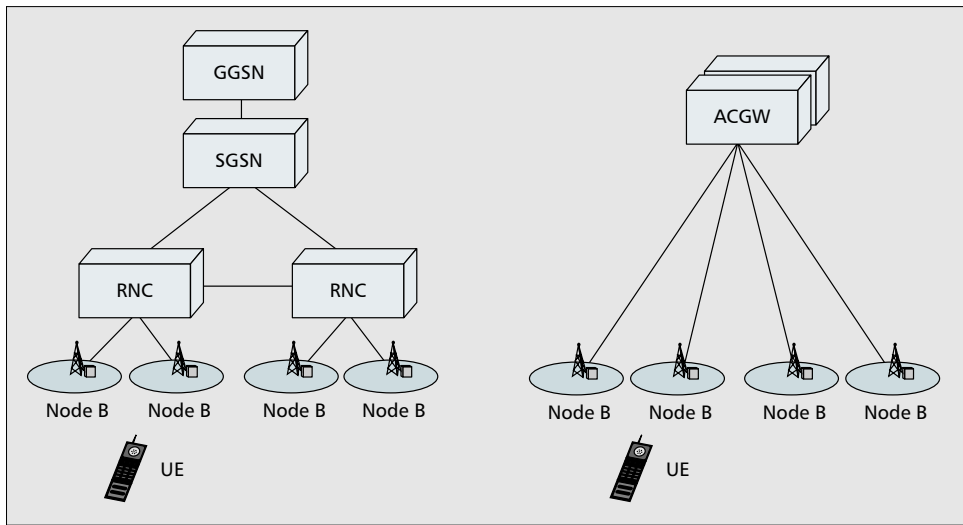
Third-generation (3G) wireless systems, based on wideband code-division multiple access (WCDMA) radio access technology, are now being deployed on a broad scale all over the world. The first step in the evolution of WCDMA has also been taken by the Third Generation Partnership Project (3GPP) through the introduction of high-speed downlink packet access (HSDPA) [1] and enhanced uplink [2]. These technologies provide 3GPP with a radio access technology that will be highly competitive in the mid-term future.

However, user and operator requirements and expectations are continuously evolving, and competing radio access technologies are emerging. Thus, it is important for 3GPP to start considering the next steps in 3G evolution, in order to ensure 3G competitiveness in a 10-year perspective and beyond. As a consequence, 3GPP has launched the study item *Evolved UTRA and UTRAN*, the aim of which is to study means to achieve further substantial leaps in terms of service provisioning and cost reduction. The overall

target of this long-term evolution (LTE) of 3G, sometimes also referred to as Super-3G, is to arrive at an evolved radio access technology that can provide service performance on a par with or even exceeding that of current fixed line accesses, at substantially reduced cost compared to current radio access technologies. As it is generally assumed that there will be a convergence toward the use of Internet Protocol (IP)-based protocols (i.e., all services in the future will be carried on top of IP), the focus of this evolution should be on enhancements for packet-based services. 3GPP aims to conclude on the evolved 3G radio access technology in 2007, with subsequent initial deployment in the 2009–2010 time-frame. At this point, it is important to emphasize that this evolved radio access network (RAN) is an evolution of current 3G networks, building on already made investments.

Among others, the targets of long-term 3G evolution are [3]:

- The possibility to provide significantly higher data rates than do the current steps of 3G evolution (HSDPA and enhanced uplink), with target peak data rates up to 100 Mb/s for the downlink and up to 50 Mb/s for the uplink.
- The capability to provide three to four times higher average throughput and two to three times higher cell-edge throughput (measured at the 5th percentile) when compared to 3GPP Release 6 (Rel-6) systems (i.e., systems based on HSDPA and enhanced uplink).
- Improved spectrum efficiency, targeting an improvement on the order of a factor of 3 compared to current standards.
- Significantly reduced control and user plane latency, with a target of less than 10 ms user plane RAN round-trip time (RTT) and less than 100 ms channel setup delay
- Reduced cost for operator and end user.
- Spectrum flexibility, enabling deployment in different spectrum allocations. This involves a smooth migration into other frequency bands, including those currently used for second-generation (2G) cellular technologies such as GSM and IS-95.



■ **Figure 1.** The current 3GPP Release 6 architecture (left) and one possible evolved 3G architecture reducing the number of nodes along the data path from 4 to 2 (right).

Based on the requirements of reduced latency and cost, it is natural to consider system architectures that contain a reduced number of network nodes along the data path. This would reduce both the overall protocol-related processing as well as the number of interfaces, which in turn reduces the cost of interoperability testing.

One additional requirement is the possibility for smooth introduction of technical solutions that fulfill these targets. Thus, any new or evolved radio access technology must be able to coexist with current 3G radio access technologies and radio network architectures and vice versa.

To achieve the above-mentioned targets, 3GPP needs to consider new radio transmission technologies as well as updates and modifications to the existing radio network architecture. Many such technologies have been proposed in the context of new fourth-generation (4G) mobile systems research [4–7]. However, in order to protect operator and vendor investments, the performance gain of any proposed update to or evolution of the 3G radio access or RAN must always be traded off against its impact on already made investments.

In this article candidate building blocks of a possible long-term 3G evolution are described. These building blocks are: an evolved system architecture, evolved quality of service (QoS) and link layer concepts, the use of orthogonal frequency-division multiplexing (OFDM) as a new access technology enabling frequency domain adaptation, and finally, the possibility of employing multi-antenna solutions. It should be noted that the standardization of 3G long-term evolution is currently ongoing. It is therefore uncertain to what extent these technical solutions will be included in the standard. An initial performance evaluation of some of these building blocks is also provided, and finally, conclusions are drawn and presented.

TECHNICAL SOLUTIONS

This section presents candidate technical solutions for the evolved radio access and RAN. A top-down approach is followed, beginning with architecture aspects and ending with physical layer issues.

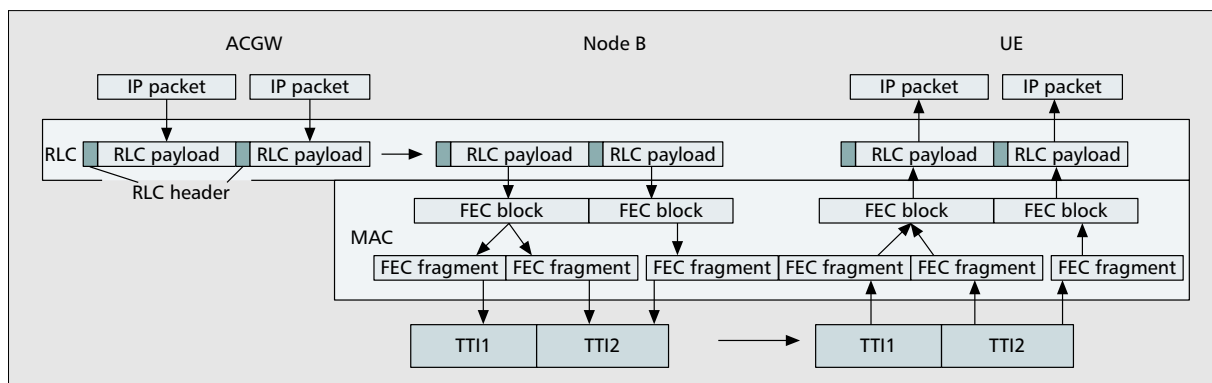
ARCHITECTURE EVOLUTION

Based on the requirements of reduced latency and cost, it is natural to consider system architectures that contain a reduced number of network nodes along the data path. This would reduce both the overall protocol-related processing as well as the number of interfaces, which in turn reduces the cost of interoperability testing. A reduction of the number of nodes also makes it possible to reduce call setup times, as fewer nodes will be involved in the call setup procedure. Such a reduction also gives greater possibilities to merge control plane protocols, thereby potentially further reducing call setup times.

Figure 1 illustrates the current Rel-6 architecture and a possible path for an architecture evolution.

In Rel-6, the Node B handles the lower layers of the wireless access, as this is the node with the antenna. The radio network controller (RNC) handles radio resource management, mobility management (locally), call control, and transport network optimization. It further acts as a termination point for the radio protocols. The gateway General Packet Radio Service (GPRS) support node (GGSN) acts as an anchor node in the home network. The serving GPRS support node (SGSN) acts as an anchor node in the visiting network and handles both mobility management and session management. Typically all traffic is routed back to the home network so that a consistent service environment can be maintained while also allowing the operator to filter traffic and provide security to the end user (e.g., by means of firewalls).

In the proposed LTE architecture, the Rel-6 nodes GGSN, SGSN, and RNC are merged into a single central node, the access core gateway (ACGW) as shown in Fig. 1. The ACGW terminates the control and user planes for the user equipment (UE), and handles the core network functions provided by the GGSN and SGSN in Rel-6. The control plane protocol for the UE



■ Figure 2. Schematic data flow through the RLC and MAC layer for downlink traffic.

will be similar to radio resource control (RRC) in Rel-6, for example, handling control of mobility and radio bearer configuration. In the user plane the ACGW will handle functions like header compression, ciphering, integrity protection, and automatic repeat request (ARQ).

The proposed architecture has the following merits:

- User-plane latency is reduced, as there are fewer nodes, and less protocol packing/unpacking.
- Call/bearer setup time is reduced, as there are fewer nodes involved in the setup procedure.
- Complexity is reduced, as there are fewer interfaces to implement and test. The amount of interoperability testing required will therefore also be reduced.
- Placing an ARQ protocol in the ACGW will provide both robustness against lower-layer losses and a simple way to provide lossless mobility.
- Performing ciphering and integrity protection of control and user plane data in the ACGW allows for a security solution at least as strong as in Rel-6.
- Support for macrodiversity can be provided with centralized radio control handling. This has been shown to give significant coverage and capacity gains.
- There is no need for a direct Node B–Node B interface for mobility. Such an interface would increase the operational burden for the operator (through additional configuration and planning) and also impose a new security threat to the network.
- A new function in the proposed architecture compared to Rel-6 is support for ACGW pooling. This allows for network redundancy solutions that increase the reliability of the network.

QUALITY OF SERVICE

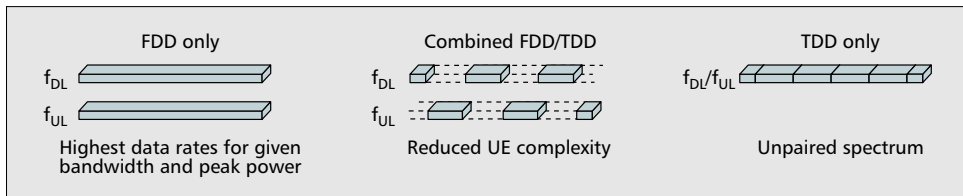
The key driver behind the QoS concept described in this section is to provide operators with effective and simple means to provide service differentiation over networks that employ high-speed shared channels. Two components of the QoS concept are described: service differentiation and simplified bearer realizations. It should be noted that the concept presented

below should be seen as one possible evolution of the current 3GPP QoS concept.

Service differentiation is enabled by classification and marking of each packet at the network edge (i.e., ACGW for downlink traffic and UE for uplink traffic). The edge node classifies each incoming packet into different predefined service classes, such as Internet access and voice over IP (VoIP). This classification could, for example, be done on the basis of information contained in the protocol headers. Following classification, the packet is marked. An explicit form of packet marking is the use of IP layer differentiated services (DiffServ) code points, while an implicit form of packet marking is the mapping of packets to “marked bearers,” such as a “marked” packet data protocol (PDP) context or radio access bearer. This marking is then used by each subsequent node to identify the service class to which the packet belongs. The edge node further performs rate policing and/or admission control to ensure that flows do not exceed a specified maximum bit rate. For some service classes (e.g., Internet access), this maximum bit rate may be specified on a subscription basis, whereas for others (e.g., VoIP) it may be specified on a session basis during the session setup phase.

Once all incoming packets have been marked and policed, each node in the data path uses the markings to carry out appropriate queuing and policy-based scheduling. The queuing in the nodes may be service class dependent; that is, the size and dropping strategies of the queue may differ depending on the characteristics of the traffic belonging to the service class. Policy-based scheduling denotes the process of scheduling according to predefined policies. Such policies can, for example, govern the distribution of bandwidth between different service classes. It is foreseen that such policies can be modified dynamically depending on the expected usage of particular services. It should be possible for the operator to push new policies to the relevant nodes through the network management system. For a more detailed discussion of service differentiation and scheduling, see [8].

Another key component of the QoS concept is the use of a reduced number of radio bearer realizations. For example, it is believed that due to the significant latency reduction in the evolved



■ **Figure 3.** Duplex schemes.

RAN, even real-time services like VoIP could be supported over a reliable link layer (acknowledged mode). The bearers will, however, differ in the scheduling policies that are assigned to them. It is believed that this reduction in radio bearer realizations will help to reduce the time to market when introducing new services, since no service-specific bearer will need to be defined and tested prior to introduction of a new service.

LINK LAYER SOLUTIONS

While the Rel-6 link layer protocols support the peak data rates of HSDPA and enhanced uplink effectively, the requirements on the evolved RAN demand enhanced link layer concepts. A fixed radio link control (RLC) protocol data unit (PDU) size is regarded as too inflexible to operate over a wide range of data rates. Small PDUs lead to too large header overhead, while large PDUs would introduce too much padding overhead for small packets like VoIP frames or TCP acknowledgments. Therefore, another solution, called the packet-centric link layer, is outlined here.

The concept foresees two layer 2 ARQ protocols as in Rel-6. The RLC protocol, which contains ARQ functionality, operates between ACGW and UE, while the hybrid ARQ (HARQ) protocol is embedded in the medium access control (MAC) layer, and operates between Node B and UE. The RLC protocol is needed to provide a reliable mobility and ciphering anchor point, and cope with congestion losses on the Iub interface, while radio interface transmission errors are typically not handled by the RLC, but by HARQ.

The key characteristic of the packet-centric link layer is to map packets (i.e., either IP packets or RRC messages) one-to-one to RLC PDUs, thereby making the size of these PDUs variable, as depicted in Fig. 2. This concept deems segmentation and concatenation at the RLC layer obsolete, thereby eliminating padding overhead. An additional field, specifying the PDU size, is required in the protocol header. However, despite this added overhead, an overall gain in terms of overhead is typically achieved since padding is avoided.

In addition, the concept has the advantage that IP packets become implicitly visible in the Node B, because each RLC PDU corresponds to exactly one IP packet. This fact can be exploited by the scheduler in the MAC layer, which now sees complete IP packets as opposed to segments thereof. This is expected to allow for more efficient scheduling decisions.

A potential problem of the packet-centric concept is that one RLC PDU may be too large to be transmitted in one frame (e.g., when the

receiver is experiencing bad signal quality). In this case, segmentation is required in the Node B. However, instead of segmenting an RLC PDU into multiple pieces,¹ it is proposed to first encode the RLC PDU into forward error correction (FEC) blocks and then use rate matching to form FEC fragments, which fit into the available radio resources. If the RLC PDU is large, this may result in a very high initial code rate, in some cases even higher than one, making it highly unlikely that such a transmission can be decoded correctly. Therefore, in combination with incremental redundancy HARQ, so-called autonomous retransmission is performed, whereby more data from this PDU is transmitted without waiting for a negative acknowledgment. This is repeated until the probability of successful reception has exceeded a certain threshold. Subsequently, conventional HARQ feedback is used to request further retransmissions if needed.

The RAN transport is expected to remain an expensive part of the network, and overdimensioning of these links cannot generally be assumed. Therefore, packet losses due to congestion in the transport network will occur despite deployment of enhanced flow control mechanisms. A further enhancement deals with this problem. In such scenarios the Node B can act as an RLC relay node and send negative acknowledgments back to the ACGW to request local retransmissions. This avoids time-consuming ARQ operations over the radio interface. Furthermore, since the sequence number is visible in the Node B, head-of-line blocking for such retransmissions can be avoided by ordering the PDUs according to their sequence number in the Node B transmission queue.

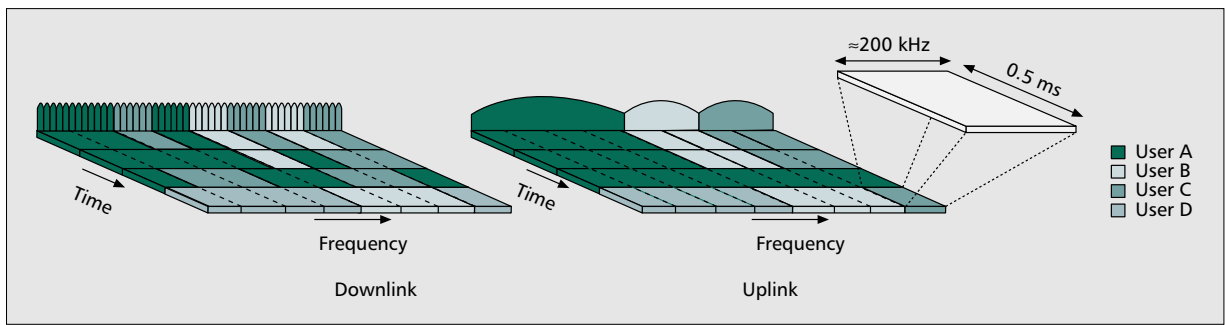
PHYSICAL LAYER AND RADIO RESOURCE MANAGEMENT

OFDM [9] is an attractive choice to meet requirements for high data rates, with correspondingly large transmission bandwidths, and flexible spectrum allocation. OFDM also allows for a smooth migration from earlier radio access technologies and is known for high performance in frequency-selective channels. It further enables frequency domain adaptation, provides benefits in broadcast scenarios, and is well suited for multiple-input multiple-output (MIMO) processing.

The possibility to operate in vastly different spectrum allocations is essential. Different bandwidths are realized by varying the number of subcarriers used for transmission, while the subcarrier spacing remains unchanged. In this way operation in spectrum allocations of 1.25, 2.5, 5, 10, 15, and 20 MHz can be supported. Due to

The RAN transport is expected to remain an expensive part of the network and over-dimensioning of these links cannot generally be assumed. Therefore, packet losses due to congestion in the transport network will occur despite deployment of enhanced flow control mechanisms.

¹ The segmentation of RLC PDUs at the MAC layer is another viable alternative. Since it is the traditional approach, further discussion is omitted here.



■ **Figure 4.** Time-frequency structure for downlink (left) and uplink (right).

the fine frequency granularity offered by OFDM, a smooth migration of, for example, 2G spectrum is possible. A 2G GSM operator can in principle migrate on a 200 kHz GSM carrier-by-carrier basis by using only a fraction of the available OFDM subcarriers. Frequency-division duplex (FDD), time-division duplex (TDD), and combined FDD/TDD, as illustrated in Fig. 3, are supported to allow for operation in paired as well as unpaired spectrum.

Downlink: OFDM with Frequency Domain Adaptation — The basic time-frequency structure of the OFDM downlink is illustrated on the left of Fig. 4. A subcarrier spacing of 15 kHz is adopted, allowing for simple implementation of dual mode Rel-6/LTE terminals as the same clock frequencies can be used. To minimize delays, the subframe duration is selected as short as 0.5 ms, corresponding to seven OFDM symbols. The cyclic prefix length of 4.7 μ s is sufficient for handling the delay spread for most unicast scenarios, while only adding modest overhead. Very large cells, up to and exceeding 120 km cell radius, with large amounts of time dispersion are handled by reducing the number of OFDM symbols in a subframe by one in order to extend the cyclic prefix to 16.7 μ s. Broadcast services are supported by transmitting the same information from multiple (synchronized) base stations. To the terminal, the received signal from all base stations will appear as multipath propagation and thus implicitly be exploited by the OFDM receiver.

Exploiting channel variations in the time domain through link adaptation and channel-dependent scheduling, as is done in current 3G systems such as WCDMA and HSDPA, has been shown to provide a substantial increase in spectral efficiency. With the evolved radio access, this is taken one step further by adapting the transmission parameters not only in the time domain, but also in the frequency domain. Frequency domain adaptation is made possible through the use of OFDM and can achieve large performance gains in cases where the channel varies significantly over the system bandwidth. Thus, frequency domain adaptation becomes increasingly important with an increasing system bandwidth. Information about the downlink channel quality, obtained through feedback from the terminals, is provided to the scheduler. The scheduler determines which downlink chunks to

allocate to which user and dynamically selects an appropriate data rate for each chunk by varying the output power level, the channel coding rate, and/or the modulation scheme. Quadrature phase shift keying (QPSK), 16-quadrature amplitude modulation (16-QAM), and 64-QAM modulation schemes are supported in the downlink.

Uplink: Single-Carrier FDMA with Dynamic Bandwidth — For uplink transmission, an important requirement is to allow for power-efficient user-terminal transmission to maximize coverage. Single-carrier frequency-domain multiple access (FDMA) with dynamic bandwidth, illustrated on the right of Fig. 4, is therefore preferred. For each time interval, the base station scheduler assigns a unique time-frequency interval to a terminal for the transmission of user data, thereby ensuring intracell orthogonality. Primarily time domain scheduling is used to separate users, but for terminals with limitations in either transmission power or the amount of data awaiting transmission, frequency domain scheduling is also used. Note that a terminal is only assigned chunks contiguous in the frequency domain to maintain the single-carrier properties and thereby ensure power-efficient transmission. Frequency domain adaptation is typically not used in the uplink due to lack of channel knowledge, as each terminal cannot continuously transmit a pilot signal covering the whole frequency domain. Slow power control, compensating for path loss and shadow fading, is sufficient as no near-far problem is present due to the orthogonal uplink transmissions.

Multipath propagation is handled by frequency domain equalization at the base station, aided by the insertion of a cyclic prefix in the transmitted signal. Transmission parameters, coding, and modulation are similar to the downlink transmission.

MULTI-ANTENNA SOLUTIONS

In order to fulfill the requirements on coverage, capacity, and high data rates, various multi-antenna schemes need to be supported as part of the long-term 3G evolution. For example, beamforming can be used to increase coverage and/or capacity, and spatial multiplexing, sometimes referred to as MIMO, can be used to increase data rates by transmitting multiple parallel streams to a single user. However, conventional multi-antenna diversity techniques, at both

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.