RADIO PROTOCOLS FOR LTE AND LTE-ADVANCED

RADIO PROTOCOLS FOR LTE AND LTE-ADVANCED

SeungJune Yi
SungDuck Chun
YoungDae Lee
SungJun Park
SungHoon Jung
LG Electronics, South Korea

WILEY

This edition first published 2012

Registered office
John Wiley & Sons Singapore Pte. Ltd., 1 Fusionopolis Walk, #07-01 Solaris South Tower, Singapore 138628

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

Library of Congress Cataloging-in-Publication Data

Radio protocols for LTE and LTE-advanced / SeungJune Yi ... [et al.].
p. cm.
Includes bibliographical references and index.
ISBN 978-1-118-18853-8 (cloth)
1. Long-Term Evolution (Telecommunications) 2. Mobile communication systems–Technological innovations. 3. Wireless communication systems–Technological innovations. I. Yi, SeungJune.
TK5103.48325.R33 2012
621.3845′6–dc23

2012020587

ISBN: 9781118188538

Set in 10/12 pt Times by Thomson Digital, Noida, India

Contents

Foreword

Back in 2004, 3G mobile communication systems providing high-speed Internet access were available, such as HSPA. At that time, the introduction of HSPA was considered a key enabler for 3G systems to meet market demand and remain competitive against other systems for a number of years. Looking forward into the future, however, multimedia and ubiquitous traffic were expected to grow rapidly. To support future traffic growth and maintain competitiveness, many operators shared the strong intention to evolve the 3G system with a long-term perspective. Driven by these motivations, 3GPP started specification work for Long-Term Evolution based on the 3G system and completed the work in Release 8 specifications.

The advent of high-performance smartphones opened a new era of opportunities for service providers and customers in the mobile industry. The rapid adoption of smartphones motivated customers towards the experiences of rich multimedia services including broadband mobile Internet. As a result, mobile traffic has increased explosively in an unprecedented manner, often approaching the network capacity of the existing mobile network, which has given mobile operators a strong incentive to pursue more aggressive strategies for evolution of the mobile network.

With this in mind, LTE is the central focus of the mobile industry, as LTE is the most promising technology on the market. For mobile operators running GSM/UMTS systems, LTE is a natural choice for their future networks. Even operators deploying 2G/3G CDMA systems are choosing the LTE system as their future network, because LTE is the only viable option for their survival. In this sense, LTE is a truly global mobile communication system that will cover almost all of the world. Currently, more than two hundred mobile operators are planning or are already deploying an LTE system.

The unique value of this book is that readers can gain in-depth understanding, especially on the radio protocols of the LTE system. While most other books on LTE focus on Physical layer aspects, this book distinctively focuses on providing sufficient knowledge of the radio protocols of the LTE system. The worth of this book is also found in the intensive treatment given to the technologies of LTE-Advanced. Most up-to-date issues and technologies in terms of LTE radio protocols are explained with fruitful figures and examples to aid readers' understanding.

All the authors have been actively involved in the standardization of radio protocols in 3GPP. The expertise of the authors has already been proven by their active contribution to the development of UMTS, HSPA, LTE, and LTE-Advanced systems. These experts are in a perfect position to take readers forward by explaining to them not only how the LTE system works but also why the system is designed in that way. Furthermore, the authors are currently participating in LTE-Advanced standardization; thus, readers can also ascertain the main issues, direction of enhancement, and the details of LTE-Advanced covered by this book.

The publication of this book is fortunate for those who want to understand how the LTE system really works, because the well-organized contents of this book will guide readers to reach a thorough understanding of the various aspects of LTE radio protocols such as design principles, architectures, and functions. If readers are already familiar with the Physical layer or the core network of LTE, this book will also provide a chance to bridge their knowledge of one part with knowledge of the other parts, such that those readers can reach a complete understanding of the LTE system.

Takehiro Nakamura
3GPP TSG RAN Chairman
NTT DOCOMO Inc.

Preface

It was only a few years ago that "ubiquitous connectivity" was recognized as the future of wireless communication systems. In the era of ubiquitous connectivity, it was expected that the broadband mobile Internet experience would be pervasive, and seamless connectivity on a global scale would be no surprise at all. The quality of service would be guaranteed no matter when/where/what the users wanted with the connectivity. Connectivity would even be extended to object-to-object communication, where no human intervention was required. All objects would become capable of autonomous communication.

Looking around at life today, those expectations are no longer what we imagine for the future, but what we already experience – ubiquitous connectivity is becoming reality. We are living in the era of this ongoing revolution of connectivity.

The revolution has been accelerated by the recent monumental enhancement of mobile communication technologies. At the center of the enhancement is the LTE system. It is well known that the development of the initial LTE specifications was completed within an unprecedented tight time schedule. During the course of development of the specifications, all the mobile business sectors including global mobile operators, network vendors, and terminal/chipset vendors have made tremendous efforts with regard to the standardization working process. The success of the LTE system as a truly global mobile communication standard is being proven by the mobile ecosystem whereby the LTE system is proliferating rapidly throughout the world. The success of the LTE ecosystem on a grand scale is fueling the rollout of the revolution towards ubiquitous connectivity.

The possibility of ubiquitous connectivity presents new business opportunities in terms of services. Keen-sighted readers will already recognize that the LTE system should be considered an "enabler" of the services we wish to enjoy, rather than simply a high-performance communication system. For key players in the mobile sector and associated industry, it is very important to understand the LTE system as an enabling technology.

The current design of the LTE system is the result of the 3GPP standardization working process, which comprises long courses of discussion and decision making. Often, not even a single decision meets with unanimous agreement among the companies participating in the standardization process. Many of the decisions are the result of discussions and compromises between the gains achievable and the costs and complexity incurred.

In principle, since specifications are collections of these decisions, they should be written in a neutral and ambiguity-free manner – that is, specifications only specify "minimum requirements." Specifications do not explain more than the minimum required to avoid violating neutrality. For this reason, readers often feel that specifications are rather closed and unrevealing.

A couple of books on LTE and LTE-Advanced are available on the market, offering readers information complementary to the specifications. However, when readers flick through such books, they will almost invariably notice that most of the pages in them are allocated to explaining the operations of the Physical layer. Only a small portion, if anything at all, is dedicated to giving descriptions of the radio protocols that work beyond the Physical layer. Such a treatment is insufficient for readers thirsty for a complete understanding of LTE radio protocols.

From an end-to-end communication point of view, the radio protocols contribute to the essential parts of operations and signaling that enable the user service with the LTE system. For example, connection control, mobility management, radio resource management, and user data transfer between the mobile terminal and the network are all controlled by radio protocols. Therefore, for the reader who is eager to view the complete picture of the LTE system, it is essential to obtain precise knowledge of the LTE radio protocols.

The main motivation behind this book is to quench the thirst of those readers by focusing on the LTE radio protocols. The details of each LTE Release 8 radio protocol are given in the first half of the book, by explaining how the terminal and the network interact over the radio interface in terms of both user data transfer and control signaling. Readers' understanding of LTE radio protocols can then be solidified when the enhanced technologies and services introduced in LTE Releases 9/10 are presented in the second half of the book, with useful figures to aid readers' understanding. It should be noted that this book only gives a guide to understanding the LTE radio protocols based on the authors' knowledge. Readers who want to have definitive information should also refer to the specifications published by 3GPP.

It is the authors' desire that this book will help to guide readers to a more complete understanding of the LTE system by combining the core of LTE radio protocols presented in this book with readers' knowledge of the Physical layer and the core network obtained elsewhere. As contributors to the LTE specifications, the authors are pleased with the current LTE system, which is already enabling services of which we only ever dreamed. Bearing in mind that the LTE system is still on its evolutionary path, we strongly believe that the timely publication of this book will make it possible for its readers to become the next-step enablers of services and technologies of which they only ever dreamed and that we never imagined.

SeungJune Yi
SungDuck Chun
YoungDae Lee
SungJun Park
SungHoon Jung

About the Authors

SeungJune Yi holds an MS degree in Electronics Engineering from Seoul National University. He joined LG Electronics in 1999, and has been actively participating in 3GPP TSG RAN Working Group 2 since 2000, contributing hundreds of papers for UMTS, HSDPA, HSUPA, MBMS, LTE, and LTE-Advanced. He has also participated in 3GPP TSG RAN Plenary as a leader of the Radio Protocol Standard team in LG Electronics. He has been the Rapporteur of the LTE PDCP specification (TS36.323) from 2008, and he has been a Vice Chairman of RAN WG2 from August 2011. He is the inventor of hundreds of patents, and won the "Inventor of the Year in LG Electronics" award in 2000, 2002, and 2008. From the Korean Intellectual Property Office, he won the "Patent Technology Award" in 2009, and was awarded "Inventor of the Year in Korea" in 2010. From 2008, he participated in the book *LTE: The UMTS Long Term Evolution* as a section author of "User Plane Protocols."

SungDuck Chun holds an MBA degree from Sloan School of Management at the Massachusetts Institute of Technology in MA, USA, and a BS degree in Electric Engineering from Seoul National University in Seoul. He studied at the Mobile Communication Laboratory at SNU for an MS degree. After joining LG Electronics in 2000, he is now Chief Research Engineer, and has been actively participating in the standardization of HSDPA, HSUPA, MBMS, LTE, and LTE-Advanced, with hundreds of contributions to 3GPP. After successfully leading the standardization activities of LG Electronics, he has been awarded "High Performance Individual" with career development opportunities. He also led several projects on the software and hardware development of mobile handsets for both the UMTS system and the CDMA system, and participated and consulted in the Bench Mark Test with telecom operators. He participated in the book *LTE: The UMTS Long Term Evolution* as a section author of "User Plane Protocols."

YoungDae Lee received an MS degree in Electronics Engineering from the Seoul National University in 1998. He joined LG Electronics (Anyang, South Korea) in 1998, and has been actively participating in 3GPP since 1999, contributing hundreds of technical documents for the standardization of W-CDMA, HSPA, MBMS, LTE, and LTE-Advanced in 3GPP TSG RAN Working Groups 1, 2, and 3, and 3GPP TSG RAN Plenary. He led the Radio Protocol Standard team in LG Electronics from 2004 to 2006, and worked in LG Electronics Mobilecomm France (Villepinte, France) as a Senior Standard Manager from 2007 to 2010. He was the Rapporteur of the "feasibility study on improvement of the MBMS in UTRAN (TS25.905)" and participated in the specification of the LTE Stage 2 as an editor of the

MBMS section (TS36.300 section 15). He was awarded "Inventor of the Year in LG Electronics" in 2000. From 2008, he participated in the book *LTE: The UMTS Long Term Evolution* as a section author of "User Plane Protocols."

SungJun Park graduated from Hanyang University, Korea in 2003 and received an MS degree in Computer Science Engineering from the same university in 2005. He is currently a Senior Research Engineer at LG Electronics and has been participating in standardization efforts for 3GPP since 2005. He has made numerous contributions to HSUPA, LTE, and LTE- Advanced.

SungHoon Jung holds an MS degree in Electronics Engineering from the Korean Advanced Institute of Science and Technology (KAIST) and a BS degree from Yonsei University in Seoul. After joining LG Electronics in 2007, he is now a Senior Research Engineer. He has been actively participating in the standardization of radio protocols for LTE and LTE-Advanced in 3GPP TSG RAN Working Group 2. His research focus spans a variety of issues such as radio resource control, mobility management, network performance optimization, femto cells, Voice over LTE, and various technologies for heterogeneous networks.

1

Introduction

In recent years, the world has seen many changes in terms of how we experience wireless communication. A few years ago, when we talked about wireless cellular communication, we usually meant voice calls through a small handset device or text exchange called Short Message Service (SMS). However, with the wide penetration of smartphones and tablet PCs, people are starting to see possibilities beyond simple voice and text calls. Any time, anywhere, and any place, people can now connect to the Internet to receive up-to-date information such as the latest news, weather information, stock price quotes, traffic information, map data, and so on. The number of applications accessing the Internet is growing rapidly and the bandwidth that these applications require is also growing rapidly. As the use of the Internet through wireless connectivity increases explosively, wireless communication systems are required to provide not just simple connectivity to the Internet but also broadband experiences.

1.1 3GPP

3GPP is an acronym for 3rd Generation Partnership Project [1], which has its roots in the Global System for Mobile Communications (GSM). As the name suggests, the original work scope of 3GPP was to produce technical specifications for the 3G wireless system. The intended 3G wireless system was an evolutionary step from the legacy GSM system, and many functional and network elements of the intended system were to be adopted from the legacy GSM system. In the end, the scope of 3GPP was expanded to include the maintenance and development of the technical specifications for GSM.

3GPP is an entity comprising multiple Organization Partners (OPs) around the world, the goal of whom is to develop a globally adopted standard of wireless communication. GSM, which is one of the 2nd generation wireless systems, is also a collaborative outcome of multiple companies from multiple countries. However, because standardization and development activity for GSM were confined originally to Europe, GSM cannot be recognized as a truly global standard. In fact, in the timeframe of development of 2nd generation communication systems, there were several efforts to standardize and develop a wireless communication system in each geographical region, resulting in the development of multiple standards such as IS-95, GSM, PDC, and so on. To develop a wireless communication standard that would be

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

adopted in most countries, it was deemed necessary to include all regional standardization entities under one roof. The motivation was that one global telecommunication standard would make it easier for people to travel anywhere with continued availability of their wireless service. It would also facilitate easy and low-cost development of wireless handsets and equipment. With these intentions in mind, 3GPP was formed to develop one unified communication standard by including as many regional Standards Development Organizations (SDOs) as possible.

The following SDOs participate in 3GPP:

- **Association of Radio Industries and Business (ARIB):** This public service corporation was chartered by the Ministry of Posts and Telecommunications of Japan in 1995. Its main role is to conduct investigation, research, and development of radio systems.
- **Alliance for Telecommunications Industry Solution (ATIS):** This standardization organization consists of more than 200 companies which have operations in North America. ATIS is accredited by the American National Standards Institute (ANSI).
- **China Communications Standard Association (CCSA):** This standardization organization was established with the approval of the Ministry of Information Industry of China and registered with the Ministry of Civil Affairs of China in 2002.
- **European Telecommunications Standards Institute (ETSI):** ETSI is a standardization organization for telecommunication industries in Europe, with more than 700 members. ETSI was created by the European Conference of Postal and Telecommunications Administrations (CEPT) in 1977 and is officially recognized by the European Commission.
- **Telecommunications Technology Association (TTA):** This organization in Korea was established in 1988 by civil law. Its major activities are planning, establishment, and certification of standards.
- **Telecommunication Technology Committee (TTC):** This standardization organization was established by the Ministry of Internal Affairs and Communication of Japan to research and develop standards for telecommunication.

In terms of structure, 3GPP consists of a Project Co-ordination Group (PCG) and Technical Specification Groups (TSGs), as shown in Figure 1.1.

The PCG is the highest decision-making body within 3GPP and its main responsibility is to ensure that the 3GPP specification is produced in a timely manner according to market requirements. In addition, the PCG approves final adoption of each TSG's work items, ratifies the election results of subordinate groups, and allocates the human and financial resources to TSGs.

The TSGs are in charge of the development of technical specifications by preparing, approving, and maintaining the specifications within each TSG's work area. Currently, there are four TSGs:

- **TSG Service and System Aspects (SA):** This TSG defines the overall architecture, requirements, and service capabilities of the systems of 3GPP. This TSG leads other TSGs. For example, this TSG decides whether support for a new service is needed or not, what is the requirement for the service, which working groups should be involved in the standardization, and what should be decided by each working group.

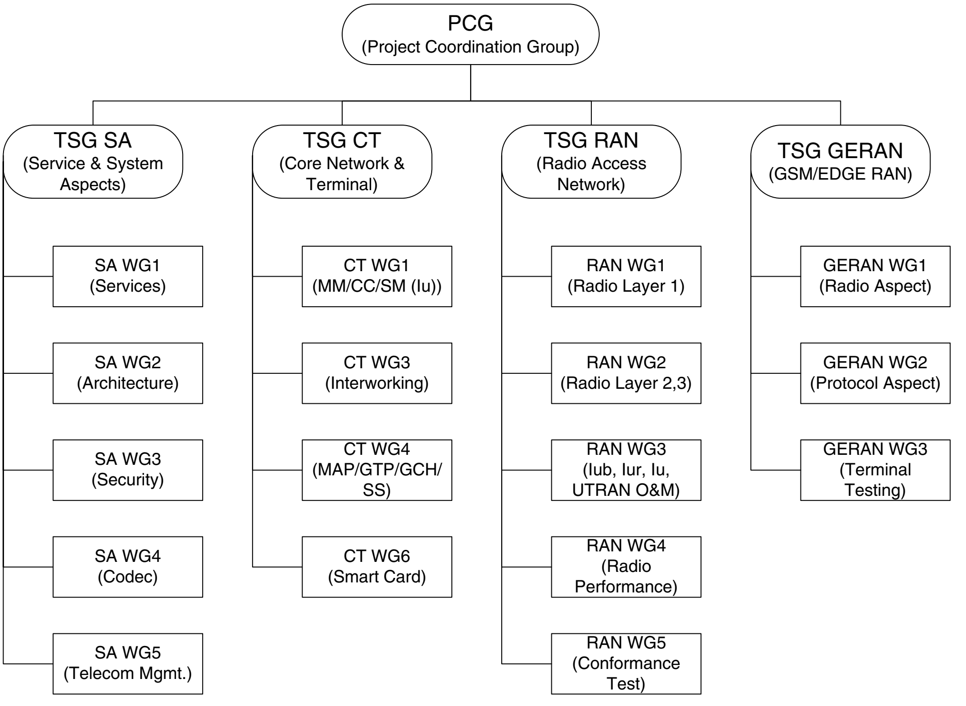

Figure 1.1 Organization chart for 3GPP [2]

- **TSG Core Network and Terminals (CT):** This TSG defines terminal interfaces, terminal capabilities, and the core network part of 3GPP. The former TSG Core Network (CN) and TSG Terminal (T) were merged into TSG CT.
- **TSG Radio Access Network (RAN):** This TSG defines functions, requirements, and interfaces of the UTRAN and the E-UTRAN in FDD and TDD modes. For example, the interface between the UE and the eNB, the interface between the eNBs, and the interface between the eNB and the MME are defined by this TSG.
- **TSG GSM/EDGE Radio Access Network (GERAN):** This TSG defines the specification of the radio access part of GSM/EDGE networks. The specification for the core network part of GSM/EDGE networks is managed within the TSG CT.

Normally, standardization work begins by creating study items or work items in the TSG plenary meeting, which is held once per quarter. Detailed work for work items or study items created in the TSG plenary meeting is undertaken by each working group of the TSG. Each working group discusses proposals submitted by individual members and tries to reach agreement on a consensus basis. For issues which fall outside each working group's expertise, the working group can consult other working groups by sending a Liaison Statement (LS). The agreements made by working groups are then sent to the TSG plenary meeting for final approval. If the agreements are approved in the TSG plenary meeting, they are included in the relevant specifications.

Table 1.1 3GPP specification series

Series	Subject
21	Requirement
22	Service Aspects (Stage 1)
23	Technical Realization (Stage 2)
24	Signaling Protocols (Stage 3) – UE to Network
25	Radio Aspect
26	CODECs
27	Data
28	Signaling Protocols (Stage 3) – RSS-CN and OAM&P and Charging
29	Signaling Protocols (Stage 3) – Intra-Fixed Network
30	Programme Management
31	SIM/USIM, IC Cards, Test Specs
32	OAM&P and Charging
33	Security Aspects
34	UE and USIM Test Specs
35	Security Algorithm
36	LTE and LTE-A Radio Technology
37	Multiple Radio Access Technology Aspects

The Mobile Competence Center (MCC) provides relevant support for making 3GPP specifications. The MCC also provides support for 3GPP meetings, such as planning meeting schedules, providing logistics, managing the IT environment of the meeting, and so on.

Each specification of 3GPP is assigned a specification number consisting of five digits. The first two digits of the specification number, called the *series*, indicate the area to which the specification belongs. A list of specification series is given in Table 1.1.

The specifications for 3GPP systems are developed in releases. Each release includes new features on top of previous releases of that specification. For example, Release 10 of the LTE specification includes support for Carrier Aggregation in addition to Release 8 of the LTE specification. Thus, though 3GPP systems are advanced with new features, backward compatibility is ensured to support UEs of old releases. Backward compatibility means that a UE of older releases can operate on a network of a later release and vice versa.

1.2 Evolutionary Path of 3GPP Systems

3GPP systems are wireless communication systems whose specifications were developed and managed by 3GPP. Examples of 3GPP systems are GSM, GPRS, EDGE, UMTS, HSPA, LTE, and LTE-Advanced. Figure 1.2 shows the evolutionary path of the 3GPP systems. Each of the 3GPP systems is explained briefly in the following sections.

1.2.1 GSM

GSM was developed originally by ETSI from 1982, and the first GSM specification was published in 1990. GSM was put into commercial use from 1991. The GSM system is one of the 2nd generation wireless communication systems and uses digital technologies (TDMA

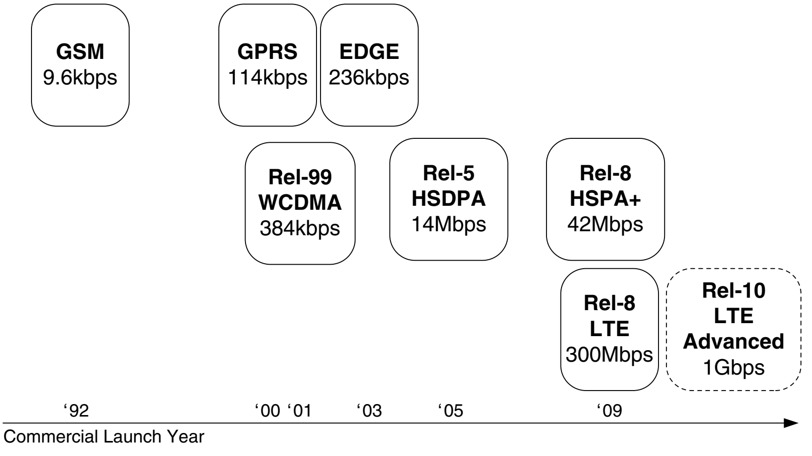

Figure 1.2 Evolution of 3GPP systems

technologies as a baseline; Figure 1.3 shows the overall network architecture of GSM), unlike the 1st generation wireless communication systems, which were based on analog technologies. By including most countries in Europe from the development phase onward, GSM could be adopted widely across Europe and eventually around the world. Now, GSM subscribers make up over 80% of the world population.

The User Equipment (UE) can be regarded as a mobile phone. The UE is composed of the Terminal Equipment (TE), which includes applications and interface to a user, the Mobile Termination (MT), which includes a module for wireless communication, and the Subscriber Identity Module (SIM), which carries information to identify each subscriber. The International Mobile Subscriber Identity (IMSI) is a unique identity that identifies a user, and is

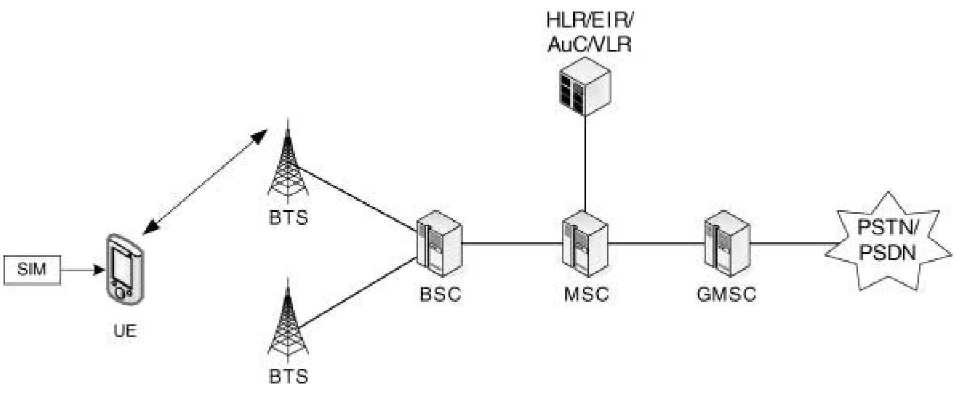

Figure 1.3 Architecture of GSM. Reproduced by permission of 3GPP, © 1998. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

stored in the SIM card. The International Mobile Equipment Identity (IMEI) is a unique identity that distinguishes each TE from other TEs.

The Base Transceiver Station (BTS) is a network node that wirelessly connects wire-line networks of GSM with a mobile terminal. Thus, the BTS is the border of the radio access network and is in charge of transmitting/receiving radio signals to/from UEs. The BTS performs multiplexing/de-multiplexing, modulation/demodulation, coding/decoding, and so on. The radio interface between the UE and the BTS is called the *Um interface*.

The Base Station Controller (BSC) is a network node that controls the BTS and manages radio resource. The BSC allocates radio channels and controls the mobility of UEs. The interface between the BSC and the BTS is called the *A-bis interface*.

The Mobile Switching Center (MSC) is a network node that controls setup and release of overall connections of services such as voice calls or SMSs. It processes requests for call setup from mobile terminals as well as requests from landlines. The interface between the MSC and the BSC is called the *A interface*.

The Gateway MSC (GMSC) is the MSC that connects the MSC to the Public Switched Telephone Network (PSTN).

The Home Location Register (HLR) manages the information of its subscribers, such as MSISDN, IMSI information, the current location of the UE, and the lists of services that each UE has subscribed to.

The Equipment Identity Register (EIR) manages information related to the status of mobile phones, especially for the information related to IMEI. This status information is used to block the use of unauthorized or stolen mobile phones within the network.

The Authentication Center (AuC) manages the authentication of mobile terminals and the encryption of communication. Based on the agreed parameter between the UE and the network, the AuC generates security keys that will be used to validate the SIM card attached to the UE. The UE will be able to use services only when authentication is completed successfully.

The Visitor Location Register (VLR) manages lists of roaming users within the network. When a UE moves into another operator's network to which it does not subscribe, information about the UE is stored temporarily in the VLR of the visited operator to provide roaming services.

1.2.2 GPRS/EDGE

In terms of resource usage, a circuit-switched service requires a continued allocation of communication resource while the call is ongoing. Thus, the service over a circuit-switched network guarantees a certain level of quality but has limitations in terms of the number of concurrent users. Voice calls are one of the typical services over a circuit-switched network. On the other hand, a packet-switched network uses communication resource per demand, enabling flexible allocation and de-allocation of resource and thereby maximizing the utilization of the resource. Thus, resource to transport data may sometimes not be immediately available, depending on the amount of data generated by users. Internet packets are a typical example of data over a packet-switched network.

The first version of GSM supported only circuit-switched services such as voice calls and SMS. To support packet-switched services such as Internet services, GSM was enhanced to include the General Packet Radio Service (GPRS). Figure 1.4 shows the overall network architecture of the GPRS.

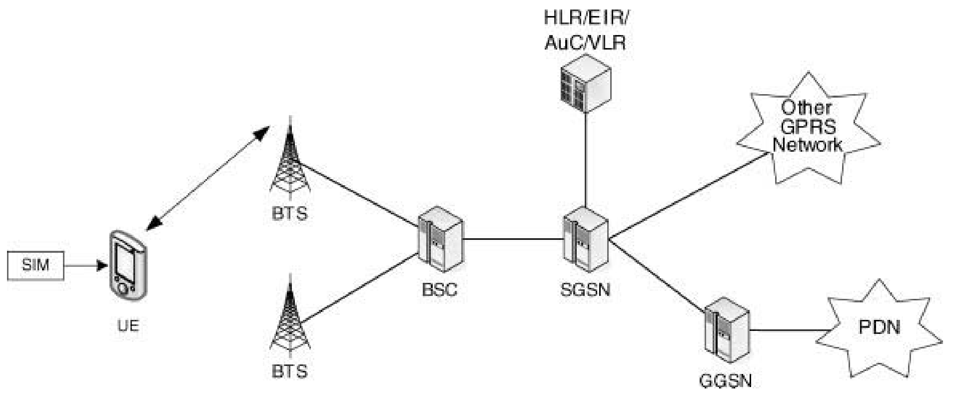

Figure 1.4 Architecture of GPRS. Reproduced by permission of 3GPP, © 2002. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

To support GPRS, the GSM core network further includes the Serving GPRS Support Node (SGSN) and the Gateway GPRS Support Node (GGSN). The SGSN is in charge of packet routing, mobility management, authentication, and so on. The GGSN is in charge of connecting the GPRS network to the external network. When packets generated outside of the GPRS networks arrive at the GGSN, the GGSN captures the packets and delivers them to the UE. When packets are generated by the UE, they are delivered to the GGSN via the BSC and the SGSN, and the GGSN eventually transfers the packets outside the PDN.

Enhanced Data rates for GSM Evolution (EDGE) is an extension of the GSM/GPRS network to provide an even higher data rate service. While EDGE uses the same core network elements as GPRS, it provides a higher data rate with support for higher modulation and coding schemes enabled by upgrades in the radio access network.

1.2.3 UMTS

The Universal Mobile Telecommunication System (UMTS) is one of the 3rd generation wireless communication systems. While GSM is based on TDMA, the radio access part of UMTS is based on Wideband CDMA (WCDMA), which uses Direct Sequence CDMA (DS-CDMA) technologies. The radio access network of UMTS is called the UMTS Terrestrial Radio Access Network (UTRAN). Because UMTS shares the same core network as GSM/GPRS, major changes compared to GSM/GPRS have been made on the radio access network. In particular, the Base Station Controller (BSC) and the Base Transceiver Station (BTS) of the GSM network cannot be used in UMTS, and thus Node B and the Radio Network Controller (RNC) are deployed instead in UMTS. The architecture of UMTS is shown in Figure 1.5.

The first version of the UMTS specification, called Release 99, was published in March 2000. The commercial service of UMTS was launched in 2002, and it has been deployed successfully all around the world subsequently.

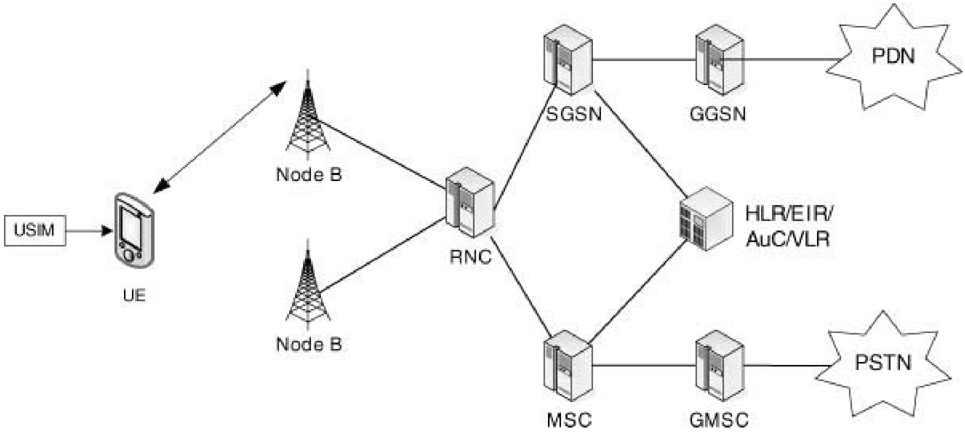

Figure 1.5 Architecture of UMTS. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

The RNC is in charge of controlling radio resource and mobility of the UE. Like GSM/GPRS, the RNC is connected to the MSC for circuit-switched services, and to the SGSN for packet-switched services. The RNC is also connected to Node B to control it. Node B is in charge of transmission and reception of radio signals to/from the UE.

1.2.4 HSPA

UMTS Release 99 – that is, the first release of UMTS – provided a maximum bit rate of 384 Kbps to a single user. To support a higher data rate, High Speed Downlink Packet Access (HSDPA) was introduced in Release 5. In Release 5, a new channel called the High Speed Downlink Shared Channel (HS-DSCH) was introduced together with various techniques such as adaptive modulation, 16 QAM, and HARQ to provide a higher data rate up to 14 Mbps. In Release 6, not only the downlink direction, but also the uplink direction was enhanced via High Speed Uplink Packet Access (HSUPA). A new channel called the Enhanced DCH (E-DCH) was introduced in the uplink to support a peak data rate up to 7 Mbps. Recently, HSDPA and HSUPA have been merged into one term, High Speed Packet Access (HSPA).

From Release 5, downlink and uplink throughputs have been enhanced one by one for each release. The enhanced version of HSPA is called HSPA+. In HSPA+, further enhanced techniques such as 64 QAM, Multiple-Input Multiple-Output (MIMO), and multiple carriers support an increased data rate up to 84 Mbps.

Figure 1.6 shows the difference in protocol termination between Release 99 DCH and Release 5 HS-DSCH. Compared to Release 99 DCH, Release 5 HS-DSCH is better for fast scheduling in that a new MAC entity, called MAC-hs, is located in Node B to perform dynamic scheduling of UEs.

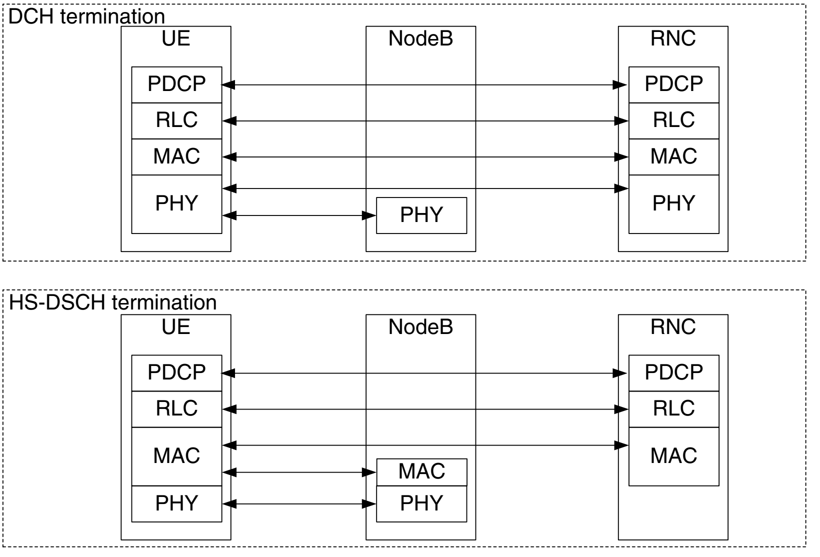

Figure 1.6 Protocol termination of DCH (Release 99) and HS-DSCH (Release 5). Reproduced by permission of 3GPP, © 2010. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

1.2.5 LTE

LTE, the Long-Term Evolution of UMTS, is the latest extension in the evolutionary path of 3GPP systems [3]. The study of LTE started in 2004, at which time the name LTE indicated a work item focusing on the potential evolution of UMTS. However, the simple sound of "LTE" fascinated many people, and hence it became the brand name of a new wireless communication system.

The radio access network in LTE is called Evolved UTRAN (E-UTRAN) compared with the UTRAN in UMTS. When the radio access network of 3GPP was upgraded from GSM to UMTS, there were not many changes in the core network. However, when the radio access network was enhanced from UMTS to LTE, the core network was also enhanced. The standardization work targeting enhancement of the core network architecture is called System Architecture Evolution (SAE), and the evolved core network is called the Evolved Packet Core (EPC). SAE is based on an all-IP network and supports not only 3GPP radio access networks but also non-3GPP radio access networks such as WIMAX and CDMA2000. The EPC's support for non-3GPP radio access networks enables operators with previous non-3GPP radio access networks to adopt LTE as their future radio access network. The term Evolved Packet System (EPS) is used to refer to the combination of E-UTRAN and EPC.

The LTE specification is defined from Release 8. LTE Release 8 is the first version of the LTE standard including the basic functionality that is essential to perform as a wireless system. A major addition in LTE Release 9 is the Multimedia Broadcast/Multicast Service (MBMS), which is used to provide broadcast and multicast services in LTE. Major new features in LTE Release 10 are the support of Carrier Aggregation (CA), which involves using multiple LTE carriers to provide a higher data rate; Relay, which is to increase LTE coverage; and Machine-Type Communication (MTC), which is to support access from many machine-type devices. As of January 2012, standardization for Release 11 of LTE is ongoing. Major works in LTE Release 11 are the enhancement of CA and MBMS.

1.3 Market Trend

In Korea, from mid-2009 to mid-2010, mobile data traffic over three telecommunication operators increased by 344, 232, and 114%. One of these operators forecasts that there will be a 49-fold increase in mobile data traffic. Similarly, Japanese operators have also predicted huge growth in mobile data traffic – a 15-fold increase by 2015, or 60% annual growth is expected. This is not a regional trend but also appears in other parts of the world. Thus, for today's operators, the time-critical issue is how to cope with the explosively increasing demand for mobile data traffic.

Recent research shows that smartphones represent only 13% of total global handsets, but these smartphones constitute over 78% of total global handset traffic. From this simple fact, it can be foreseen that when all existing handsets are replaced by smartphones, the demand for mobile traffic will be huge. Furthermore, as more applications are developed and become popular, the average amount of mobile traffic per capita will also increase in addition to the penetration of smartphones. The popular use of video streaming and video telephony adds another dimension to the increase in mobile traffic.

To accommodate the increased demand for mobile traffic, more dense deployment of base stations using the concept of femto or pico cells is being considered. The use of Wi-Fi to

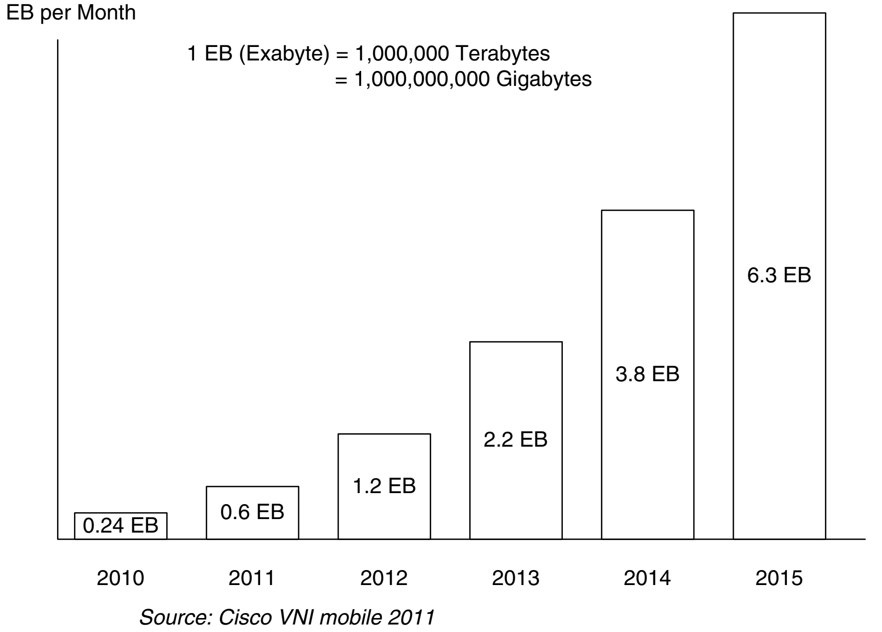

Figure 1.7 Forecast of mobile traffic [4]. Reproduced with permission from Cisco VNI Global Mobile Data Forecast, 2011–2016

offload mobile traffic of cellular networks will also be widespread. However, as humans basically tend to move around, the evolution of the cellular network is the most essential issue.

Figure 1.7 shows a simple estimation of mobile traffic in the coming years. As of today, the typical devices for wireless connectivity are either smartphones or tablets. However, in the future, traffic consumed by Machine-to-Machine type communication will increase. For example, telemetric devices for utilities, remote healthcare equipment, and assistance devices in a car will find more usage. While the advent of new types of devices and business domains like Machine-to-Machine offers new opportunities for expanding revenue, frequent access and the generated traffic from these devices will pose a challenge to the mobile communication industry.

1.4 Requirement of LTE

Back in 2004, when LTE standardization work started, the demand for a mobile data service was low. At that time, there was no killer application requiring a data rate up to 384 Kbps (which was supported by UMTS). Nonetheless, the mobile industry saw a need to develop a new wireless system because the basic consensus at that time was that UMTS was oriented toward a voice service and was not suitable for a high data rate service. For example, it was argued that the OFDM technique is better suited than the CDMA technique to a high data rate service requiring larger bandwidth. To combat the innate nature of a fast-fading channel environment, OFDM was preferred to CDMA.

In addition, as the hardware capability of mobile terminals advanced and wired Internet usage increased dramatically, a next generation wireless system optimized for a packet-oriented service was deemed necessary. Because the UMTS system had to support both a circuit-switched service and a packet-switched service from the beginning, latency, coverage, network architecture, and capacity could not be optimized for a packet-switched service. Rather than upgrading based on the existing architecture, defining a totally new system was inevitable. In fact, the justification of the LTE study item states that the LTE system should ensure that the new system should be competitive for the next ten years.

When LTE standardization work started, the following were agreed to be major requirements of the LTE system:

Support of high data rate

- downlink peak data rate of 100 Mbps for 20 MHz downlink bandwidth;
- uplink peak data rate of 50 Mbps for 20 MHz uplink bandwidth.

The peak data rate means the maximum bit rate that one single user in a cell can be provided with. A UE in a good channel condition can be served with 100 Mbps if all the resource in a cell is allocated to the UE. 100 Mbps is the requirement for a cell with 20 MHz bandwidth, and the requirement for a cell with smaller bandwidth is proportionally reduced.

Low latency

- packet delivery latency of less than 5 ms;
- connection establishment latency of less than 100 ms.

In general, low latency is desirable for user experiences. However, low latency requires more power consumption at the terminal side. For example, let's assume that there is a piece of information that the network wants to send to the UE. If the network wants to reduce the amount of delay in delivering the packet from the network to the UE, it would be beneficial for the UE to monitor the network command for a longer time period. However, this increased monitoring time means increased UE wake-up time and more battery usage. Thus, a low latency value does not always guarantee the best performance.

The 5 ms requirement for packet delivery latency of LTE is the achievable value in a good channel scenario. That is, 5 ms packet delivery latency is the value that should be met when the processing time at the network and the UE is minimized, the transmission is always successful at the first transmission, and radio resource is allocated as soon as data are available. Of course, the UE is assumed to be already connected to the network and in active mode.

The 100 ms requirement for connection establishment is the value taken when a UE transits from RRC_IDLE to RRC_CONNECTED. In other words, this is the transition time from the state where a UE cannot immediately transmit any packet to the state where a UE is ready for immediate transmission and reception of packets. While it would be beneficial for a UE to stay in RRC_CONNECTED as much as possible to reduce latency, this requires additional measurement, signaling exchange, and network management of the UE. Thus, how long a UE is kept in RRC_CONNECTED also requires a lot of consideration.

High capacity

- support for at least 200 active users per cell of bandwidth equal to 5 MHz.

LTE is expected to have more capacity than UMTS; it is also expected to support more simultaneous voice calls with given radio resource than UMTS. In addition, one of the characteristics of a packet-switched service is that it does not require continuous allocation of radio resource. For example, in the case of UMTS, each active user using a voice call with a circuit-switched connection should be allocated one channel code during the lifetime of the voice call. Thus, the number of simultaneous voice call users in UMTS is limited by the number of channel codes available in the cell. However, in LTE, due to dynamic allocation of radio resource based on each user's demand, there is no need for an active user to be constantly allocated radio resource during the lifetime of the service. This means that the LTE system can bring more users into the active state.

Spectral efficiency (bps/Hz/site)

- to 4 times downlink spectral efficiency of UMTS Release 6 (HSDPA);
- to 3 times uplink spectral efficiency of UMTS Release 6 (HSUPA).

Spectral efficiency is a key metric that measures how much data can be delivered with a given bandwidth. The increase in spectral efficiency of commercial mobile networks indicates the advance of digital signal processing technologies. In the early stage of wireless communication, digital signal processing was hard to achieve because the capability of hardware and software was limited. As hardware produces more computing power with less energy, more digital processing becomes possible and more data can be delivered with limited radio resource. LTE has come closer to the maximal spectrum efficiency that can be achieved theoretically. Through high-order modulation and the MIMO technique, LTE is required to produce 2 to 4 times better results than UMTS Release 6.

Mobility support

- high performance support up to 120 km/h;
- connection maintenance up to 500 km/h depending on frequency band;
- voice and real-time support with similar QoS as UTRAN at the supported speed;
- interruption time less than 300 ms for voice and real-time service during switching between UTRAN/GERAN and E-UTRAN.

In the case where a user is moving in a car, the user's velocity may reach 120 km/h, hence it is reasonable that LTE should provide high performance up to that speed. However, this does not mean that LTE does not support speeds above 120 km/h. Because high-speed trains are available in many parts of the world, support for mobility up to 300 km/h is necessary, though the service quality is not as good as that achieved at lower speeds.

One difference between UMTS and LTE is that handover in LTE is hard handover while handover in UMTS is soft handover. Here, hard handover means that the connection to the source cell is disconnected before the UE establishes the connection to the target cell during handover. On the other hand, for a soft handover, a UE makes the connection to the target cell before disconnecting from the source cell. Thus, one may assume that UMTS provides better quality for a voice and real-time service because there is no disconnection during handover. However, the voice codec generates a voice packet every 20 ms and voice packet delay up to 200 ms does not have much impact on human experience. Thus, though the LTE system uses the hard handover method, the impact on the user experience is marginal as long as the hard handover is performed quickly enough.

However, in the case of inter-RAT handover, where a voice call is forwarded from LTE to GSM/UMTS or vice versa, the required interruption time is longer compared to intra-LTE handover.

Coverage

- performance should not be degraded up to 5 km from the cell center;
- degradation is acceptable up to 30 km.

From an efficiency point of view, it would be good for a cell to be able to cover a large area, because the operator could then install fewer base stations. However, as a cell gets larger in size, the cell needs to support more users and this could result in a low data rate service to the users. In addition, as a UE stays farther from the cell center, the UE has to use more power for its transmitted signal to reach the base station, leading to more power consumption. On the other hand, deployment of smaller cells will lead to more handover, leading to more UE measurement, more data transfer interruption, and more signaling exchange between network nodes. Thus, the typical or reasonable assumption on the cell radius is important when a system is designed. Considering both rural and urban environments, it was assumed that the LTE system should be optimized for a cell with a radius of 5 km. However, as stated above, taking into account that high-speed trains can reach 300 km/h, a larger cell size up to 30 km should not be excluded.

Deployment

- support for different spectrum bandwidth from 1.25 to 20 MHz.

For each country or region, there is a regulation that governs the allocation of radio spectrum. Accordingly, while many countries use the same frequencies for the LTE service, it is to be expected that some countries will use different frequencies for it. In addition, due to other commercial, experimental, military, or amateur purposes, the amount of bandwidth allocated to the LTE service is different in each country. Thus, to accommodate different bandwidths and frequencies for LTE use, one of the requirements of LTE is to support different sizes of spectrum bandwidth. In addition, the needs of existing operators who want to upgrade their legacy networks to LTE should be considered. For example, LTE's support for bandwidth of 1.25 MHz is for operators who want to upgrade their 1.25 MHz CDMA2000 networks to LTE.

Service

- packet-based architecture.

This is one of the critical requirements of LTE because LTE aims to provide the best Internet experience. When a system has to support a circuit-switched service, the functionality of static resource reservation is required in the system. This eventually makes optimization for a packet-switched service difficult. Because the Internet has become the main stream of information distribution, LTE has chosen to support only packet-switched services.

One of the basic services that a wireless system should provide is a voice service. In fact, the typical service scenario of a circuit-switched system is a voice service. Thus, while the principle of supporting only a packet-switched service in LTE makes sense, support for a voice service in LTE is also important for the success of the LTE system. Thus, LTE should include relevant support for a Voice over IP (VoIP) service. In particular, a VoIP service has special characteristics such that the service generates voice packets at regular intervals and the generated packets have huge overhead due to IP or UDP headers. This leads to the inclusion of special functionality in the protocol stack of LTE.

1.5 Overview of LTE Architecture

As stated earlier, LTE is the term used to describe the radio access network, and EPC is the term used to describe the core network. The radio access network manages radio connections to the UE, and the core network manages overall services. For example, the radio access network transmits towards and receives from the UE, performs handover, and so on. The core network performs voice call coding, charging, and interworking with outer networks. In this section, both SAE and LTE architectures are analyzed.

1.5.1 Network Architecture

The EPC is comprised of the PDN Gateway (P-GW), the Serving Gateway (S-GW), the Home Subscriber Server (HSS), the Policy Control and Charging Rules Function (PCRF), and the Mobility Management Entity (MME), as shown in Figure 1.8.

The P-GW is the outermost entity of the EPC that bridges the inner nodes of the EPC and the external Internet. It allocates IP addresses to the UE and performs filtering of IP packets according to the guidance set by the PCRF. After filtering the IP packets for a specific user into several streams called EPS bearers, each stream is treated according to the specific QoS treatment set for the stream. For example, by filtering at the P-GW, packets of Internet browsing and packets for voice traffic are mapped to different bearers within 3GPP networks.

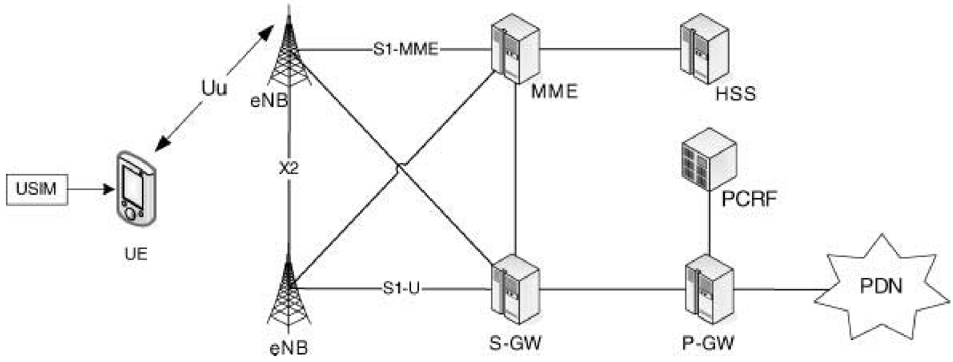

Figure 1.8 Architecture of LTE/SAE. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

The S-GW is a mobility anchor for a UE. A mobile anchor means that the data for a specific UE pass through the S-GW regardless of the serving eNB to which the UE is connected. In addition, when the UE is in RRC_IDLE, downlink packets that arrive from the outside Internet are buffered temporarily at the S-GW until the location of the UE is identified and the connection to the UE is established.

The HSS stores and handles user-specific information such as subscription information, authentication information, and security information.

The MME is a central node of the core network that exchanges signaling with the UE. The MME performs bearer management such as establishment, modification, and release. In addition, the MME also performs a mobility-related function such that it tracks RRC_IDLE UEs' location for paging. Regarding security, the MME performs the authentication procedures to check the validity of the UE, and allocates temporary identities to the UE to provide user identity protection. To enable the eNB to perform ciphering and an integrity check, the MME further provides security keys to the eNB.

The radio access part of the network consists only of eNBs. Because there were two network nodes – Node B and the RNC – in the UTRAN, performance could not be optimized. For example, an RLC PDU generated at the RNC has to pass through Node B to reach the UE. As a result, the round trip time in the RLC layer is long, resulting in delayed retransmission and a lower data rate. In LTE, by minimizing the number of network elements in the path of data delivery, more efficient use of radio resource becomes possible.

Basically, the eNB is responsible for the management of radio resource for a UE. It establishes and manages radio bearers toward a UE, provides relevant QoS over the radio interface, and performs mobility-related functions such as measurement and handover procedures. In addition, it performs dynamic allocation and de-allocation of radio resource to deliver user data.

The eNB is connected to the MME and the S-GW via the S1 interface. From a single UE's perspective, the signaling message of the UE is delivered to the MME via the eNB, and user-generated traffic is transported to the S-GW via the eNB. Here, the signaling message is a control message that is used to provide Internet connectivity to the UE. The signaling message that is used by user application itself is transported via the S-GW. For example, signaling messages that a UE sends to update its current location are delivered to the MME. However, the log-in-related signaling messages for the instant messaging service are delivered to the S-GW as an IP packet which is transparent to the eNB.

The S1 interface can be classified into S1-MME and S1-U. The interface between the eNB and the MME is called S1-MME, and the interface between the eNB and the S-GW is called S1-U.

Major functions of the S1 interface are:

- **S1 UE context management:** This function includes establishment, management, and release of UE context to support individual signaling for a given UE.
- **Management of E-RAB:** This function includes establishment, management, and release of E-UTRAN resources for the delivery of user data.
- **Mobility support:** This function includes intra-LTE handover, inter-RAT handover, and so on to support UEs' mobility.

Figure 1.9 shows the kinds of functions performed by each node in the network. A brief description of each function is given below.

eNB functions:

- **Inter-cell RRM:** Inter-cell RRM is about coordination among eNBs and within an eNB regarding how to use radio resource optimally in each cell within the network. Because the same frequency is typically used by a neighboring cell in LTE, a useful radio signal of a cell is at the same time interference to neighboring cells. Interference coordination with resource allocation and power adjustment is an example of this role.
- **RB control:** The radio bearer is the unit of differentiation of service provision over the radio interface. The eNB establishes, modifies, and releases radio bearers based on the service requirement of each radio bearer.
- **Connection mobility control:** When a UE enters RRC_CONNECTED, the mobility of the UE is controlled by the eNB with assistance information delivered from the UE. The

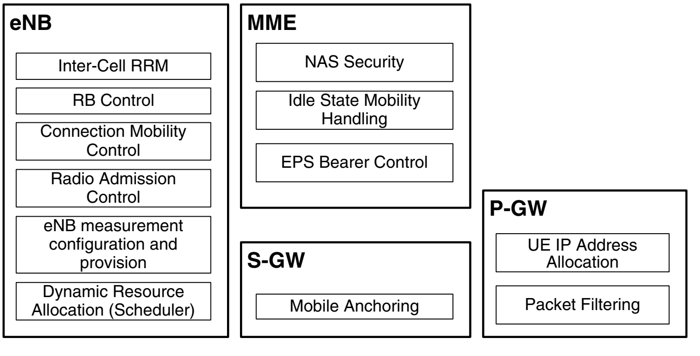

Figure 1.9 Functional split of LTE and SAE. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

eNB decides which cell should be used for the connection to the UE and makes an appropriate handover decision. The eNB decides the measurement configuration for a UE considering the UE's capability and the topologies of the neighboring cell.
- **Radio admission control:** Based on the resource usage of a cell and the resource requirement of services, the eNB decides whether to accept a new RRC connection and a new radio bearer. In addition, when there is resource shortage in a cell, the eNB further decides which RRC connection or radio bearer should be released to mitigate congestion in the network.
- **eNB measurement configuration and provision:** To optimize radio resource usage and network configuration, the eNB performs several measurements and provides the results to operators or control entities. For example, the statistics of network resource usage or error rate information over the radio interface can be delivered to the OAM (Operation, Administration, and Maintenance) entity to identify the problematic area or nodes within the network and change-related parameters.
- **Dynamic resource allocation:** To maximize radio efficiency, radio resource is allocated dynamically based on the needs of each UE. Radio channel condition is also considered in deciding resource distribution.

MME functions

- **NAS security:** Over the radio interface, the PDCP layer provides ciphering and integrity protection functionality to perform E-UTRAN level security. In addition to this, the MME provides additional security to protect signaling messages of the Non-Access Stratum (NAS) layer which is located above the RRC layer.
- **Idle state mobility handling:** For a UE in RRC_IDLE, the eNB does not have any context information for the UE, but the MME maintains the UE's context. The MME keeps track of the UE's location at Tracking Area (TA) level, and initiates the paging procedure for an RRC_IDLE UE when there is an incoming packet for the UE. The TA consists of at least one cell, and is defined by operators. Whenever a change of TA is detected, the UE should update its location to the MME.
- **EPS bearer control:** While QoS control and management of radio bearers is performed at the eNB, overall QoS control and management of EPS bearers is performed at the MME.

S-GW functions

- **Mobile anchoring:** When a UE moves around within the E-UTRAN, there should exist a pivot point within a network through which user data are delivered or forwarded to the outside network. The S-GW serves as a mobility anchor where the user data pass through regardless of a change in cell or eNB within the E-UTRAN.

P-GW functions

- **UE IP address allocation:** For a UE which does not have a static IP address, the UE should be allocated a dynamic IP address. The P-GW assigns the dynamic IP address.
- **Packet filtering:** When multiple EPS bearers are established for a UE, the P-GW classifies user packets into appropriate EPS bearers according to configured rules.

1.5.2 QoS Architecture

Quality of Service (QoS) is about all related characteristics of a connection such as delay, delivery time, throughput, signal-to-noise ratio, error rate, interruption, clarity, jitter, and so on. It is a set of parameters that are related closely to how a user might feel about the service or how a connection should provide service to satisfy a user's expectations. The required level for each parameter is different between services, and the network should be able to select the most suitable values for the parameters. A simple example of QoS differentiation is the mix of VoIP and Internet-browsing services. For a VoIP service, a small loss of packet is not a problem but delayed transfer of voice data is a big problem with regard to the user experience. On the other hand, for an Internet-browsing service, a little delay is not a problem but loss of even one packet is a problem. To satisfy these different requirements of services, separate streams are used to transport data of different services. To support different QoSs, each stream of different characteristics is mapped to different bearers. Thus, a bearer is a unit of QoS control, and one bearer is used to fulfill one set of QoS.

Typical parameters that constitute QoS settings are:

- **Guaranteed Bit Rate (GBR):** This is a bit rate guaranteed by the LTE network. The UE will be served with this bit rate even when there is congestion in the network.
- **Maximum Bit Rate (MBR):** This is the maximal bit rate that can be provided to a bearer. When there are a lot of users or there is ongoing congestion in the network, this bit rate cannot be achieved.
- **Jitter:** Even if packets originate from the same source and are delivered to the same destination, the packets will arrive with different levels of delay. This parameter indicates the variation in transfer delays of the packets.

When a UE initially registers itself to the network, the network and the UE establish one default EPS bearer for the packet-switched service. At this time, all data for the UE are transported using the default EPS bearer. Later, if the core network identifies a need to apply separate treatment for specific user data streams among the data transported over the default bearer, the EPC establishes additional bearers that are used to meet the specific requirements of the specific streams. After that, each node within the network is configured to provide the required level of treatment for the bearers.

Figure 1.10 shows the overall picture of QoS provision.

The End-to-End service bearer is the connection between the application at the UE and the application at the other side of the Internet. The data over this virtual bearer are delivered through the EPS bearer that provides a delivery service between the UE and the P-GW, and the external bearer that is the connection between the P-GW and the server at the other side of the Internet. The P-GW performs mapping of the data that arrive at the P-GW from the Internet to the appropriate EPS bearer. The EPS bearer is further composed of the E-UTRAN Radio Access Bearer (E-RAB) and the S5/S8 bearer. The E-RAB is a bearer between the S-GW and the UE while the S5/S8 bearer is a bearer between the S-GW and the P-GW. The E-RAB can be decomposed into the radio bearer and the S1 bearer. The radio bearer is a bearer between the eNB and the UE while the S1 bearer is a bearer between the eNB and the S-GW. Among these bearers, the radio bearer is established over the radio interface while the other bearers are typically established in the wired network.

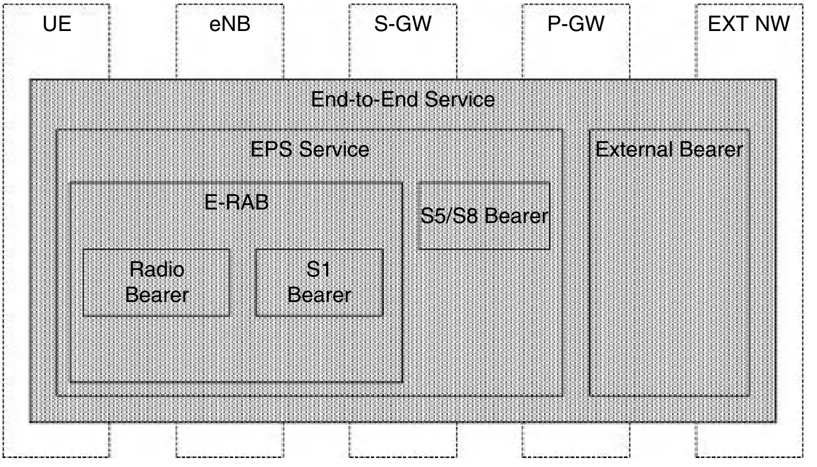

Figure 1.10 QoS architecture of LTE and SAE. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

For all the different types of bearer listed above, there is a one-to-one mapping relationship. In other words, there is a unique match between an EPS bearer and an E-RAB, between an E-RAB and a radio bearer, and between a radio bearer and an S1 bearer. From the UE's perspective, there is no need to differentiate an EPS bearer from an E-RAB bearer. Thus, only the mapping information between EPS bearer and radio bearer is provided to the UE. That is, when a UE receives data from an eNB through a radio bearer, the UE only needs to know to which EPS bearer the data should be forwarded, without knowing the related E-RAB.

In LTE, a parameter called the QoS Class Identifier (QCI) is used. The QCI defines a unique expected treatment of a bearer and is intended to provide similar handling of bearers of the same QCI even if the network nodes were developed by different manufacturers. From an operator's point of view, regardless of the eNB to which a UE is connected, the same level of experience should be provided for the same type of service. To ensure that all network elements within an operator's network provide the same level of experience for the same service, the QCI is used to inform the network nodes of the kind of treatment expected for a given bearer. Based on the received QCI value, each network node knows how it should treat the given bearer. The QCI value is specified within 3GPP so that each vendor knows the expected characteristics for a bearer with a given QCI value. The characteristics include priority level, delay budget, packet loss rate, and so on.

The following parameters are used to define the QCI:

- **Priority Level:** The priority level value is used to decide the data stream to be prioritized. This parameter is used to differentiate flows within a UE as well as to differentiate flows among different UEs. This parameter is used when it becomes impossible to meet the PDB requirement of all data that a certain network node processes. Data of lower priority level are treated after the PDB requirement for data of higher priority is satisfied. Value 1 indicates the highest priority.

- **Packet Delay Budget (PDB):** This parameter specifies the upper limit of time that is allowed to transport the data block of the bearer. Thus, network nodes should try to deliver a packet of data within the time specified by the Packet Delay Budget. However, this cannot always be satisfied in cases where the network is congested.
- **Packet Error Loss Rate (PELR):** This parameter defines the amount of data that may be lost during transfer. When this parameter is set, the involved entity should try to ensure that the amount of data which are not delivered successfully to the receiving side is not more than the amount defined by the PELR parameter. In E-UTRAN, appropriate setting of the RLC layer and HARQ parameter is required to meet the PELR requirement. While PDB and priority level are brought into play when there is congestion on the network, the PELR parameter is used even for a non-congested situation.
- **Resource Type:** This parameter indicates whether a certain level of long-term allocation of resource is required or not. If this field is set to "GBR", it means that the concerned bearer must be provided with at least a certain level of data rate. Typically, GBR bearers are used to transport service data of voice services or video services, which generally require semi-static guaranteed allocation of resource. In general, as long as a service using a GBR bearer does not generate more packets than the configured amount of data rate, it is expected that the PELR or PDB requirement will be met for the data of the GBR bearer. On the other hand, for non-GBR bearers, there is no guarantee that a certain level of data rate will be provided.

Table 1.2 shows the QCI list currently defined in 3GPP [5].

While the QCI is about the treatment of data for a certain bearer carrying user data, the allocation and retention priority (ARP) defines a relative priority between bearers. Because

Table 1.2 QCI list. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

QCI	Resource Type	Priority	Packet Delay Budget	Packet Error Loss Rate	Example Services
1	GBR	2	100 ms	10^{-2}	Conversational Voice
2	GBR	4	150 ms	10^{-3}	Conversational Video (Live Streaming)
3	GBR	3	50 ms	10^{-3}	Real-Time Gaming
4	GBR	5	300 ms	10^{-6}	Non-Conversational Video (Buffered Streaming)
5	NON-GBR	1	100 ms	10^{-6}	IMS Signaling
6	NON-GBR	6	300 ms	10^{-6}	Video (Buffered Streaming) TCP-based Services
7	NON-GBR	7	100 ms	10^{-3}	Voice, Video (Live Streaming) Interactive Gaming
8	NON-GBR	8	300 ms	10^{-6}	Video (Buffered Streaming) TCP-based Services
9	NON-GBR	9	300 ms	10^{-6}	Video (Buffered Streaming) TCP-based Services

radio resource and network capacity are limited, all bearers requested either by the network or by the UE cannot be established or maintained. Thus, at times of congestion, the ARP setting is used to decide whether to accept a new bearer request, or which bearer should be released. ARP is composed of the following parameters:

- **Priority level:** This parameter indicates the relative importance of a resource request. Based on this value, the network node decides whether to accept or reject a bearer establishment request. A value of 1 is the highest priority and a value of 15 is the lowest priority.
- **Pre-emption capability:** This parameter defines whether a data flow of this bearer can use resource of other data flows of lower priority.
- **Pre-emption vulnerability:** This parameter defines whether resource of a data flow of this bearer can be used for other data flows of higher priority. If this parameter is set to "yes", the data over this bearer may not be delivered when resource congestion problems occur.

1.5.3 Radio Protocol Architecture of LTE

The E-UTRAN is composed of the eNB and the UE. The interface between the eNB and the UE is called the *Uu interface*. The radio protocol of LTE includes only the Uu interface. While control plane architecture is used to deliver and exchange signaling messages that are critical to manage UEs' connectivity, user plane architecture is used to deliver and exchange data packets that are consumed by users or applications.

The network nodes involved in the control plane are the UE, the eNB, and the MME. Protocol entities that constitute the control plane are shown in Figure 1.11. For the Uu interface, the control plane is composed of the Physical layer (PHY), the MAC layer, the RLC layer, the PDCP layer, and the RRC layer. For the S1-MME interface, the Stream Control Transmission Protocol (SCTP) and the S1 Application Protocol (S1AP) are used for the control plane. The SCTP is used to support signaling exchange for the S1AP and is defined by RFC4960.

Protocol entities that constitute the user plane are shown in Figure 1.12. For the Uu interface, the user plane is composed of the Physical layer, the MAC layer, the RLC layer, and the PDCP layer.

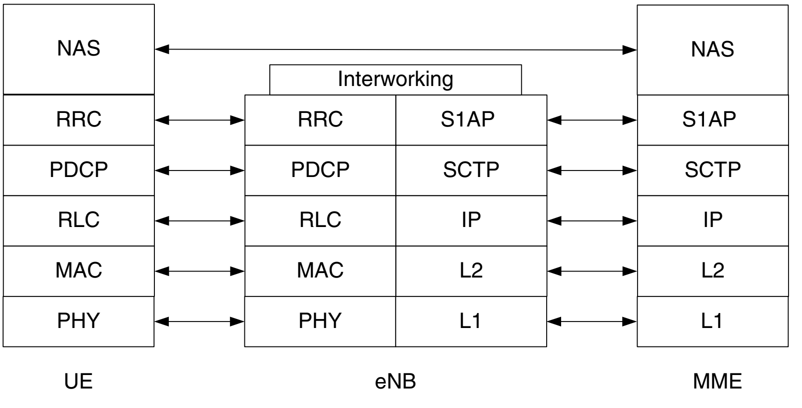

Figure 1.11 C-plane architecture. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

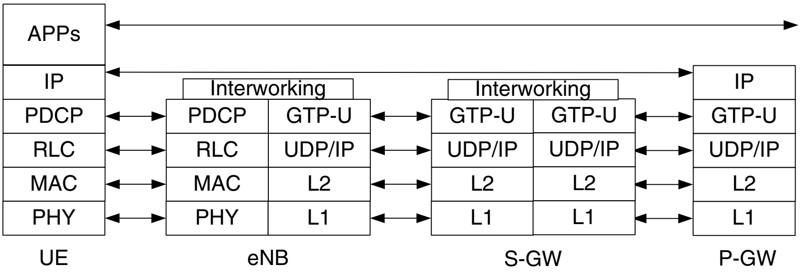

Figure 1.12 U-plane architecture. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

Figures 1.13 and 1.14 respectively show the functions performed by the MAC layer, RLC layer, and PDCP layer at the eNB and the UE. Brief descriptions of each function are given below, including some of the Physical layer functionality.

RRC layer

The RRC layer is in charge of radio resource control over the radio interface. The RRC layer in the eNB and the UE exchanges signaling messages that are essential to managing the connection between the eNB and the UE. The following functions are performed by the RRC:

- **Connection management:** The RRC layer manages RRC connection between the eNB and the UE. This includes establishment, modification, and release of an RRC connection. Without RRC connection, the eNB and the UE cannot exchange any signaling messages that are necessary for the UE's proper operation within the wireless network.
- **Radio bearer management:** The RRC layer manages radio bearers between the eNB and the UE. This includes establishment, modification, and release of a radio bearer. The radio

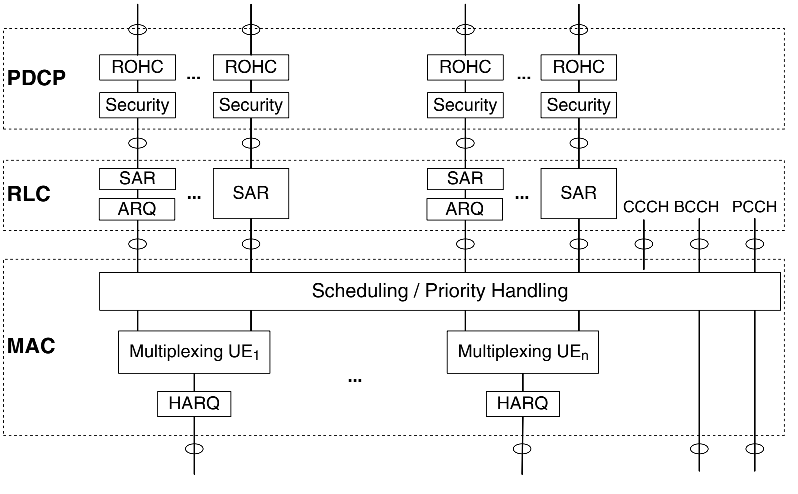

Figure 1.13 Overall radio functions at the eNB. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

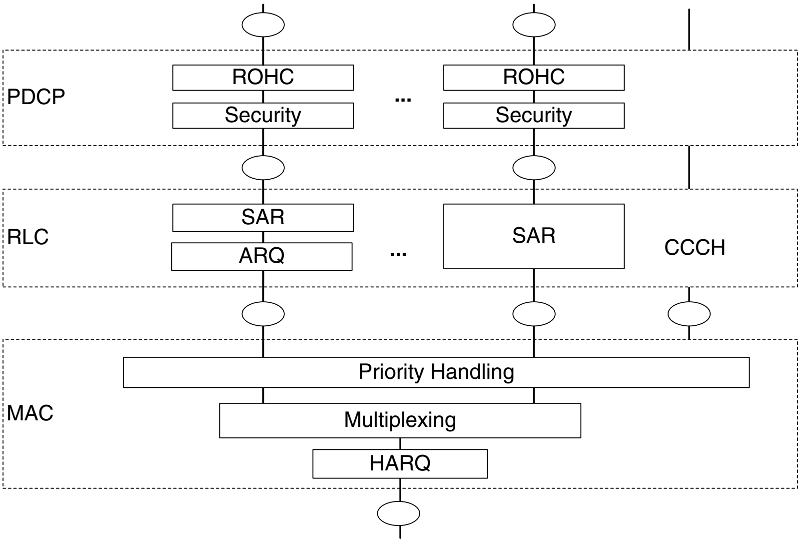

Figure 1.14 Overall radio functions at the UE. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

bearer is the last leg of the bearers used to transport user data. When the RRC layer configures and re-configures a radio bearer, it uses the QoS information of the concerned radio bearer to meet the expected service requirement for the data transported over the radio bearer.

- **Mobility management:** The RRC layer controls mobility of the UE. After taking into consideration cell load, radio link quality, and location, the RRC layer makes a decision on the best cell to serve the UE, and makes an appropriate adjustment. To support this function, measurement configuration and measurement reporting are performed.
- **Signaling connection:** The RRC layer provides signaling bearer services to transport the upper layer's signaling message. Upper layer signaling is the exchange of control messages between the UE's upper layer and the EPC, which includes the MME.

PDCP layer

The PDCP layer is the layer located below the RRC layer or the IP layer and above the RLC layer. The role of the PDCP layer is two-fold: to perform a security function and to carry out header compression. The security function provides ciphering to prevent non-authorized parties from looking at data and provides integrity protection to prevent data from being tampered with by non-authorized parties. Header compression is used to maximize radio resource efficiency by avoiding transmission of redundant information over the radio interface. Robust Header Compression (ROHC) defined by IETF is used for header compression.

RLC layer

The RLC layer is the layer located below the PDCP layer and above the MAC layer. The main role of the RLC layer is to provide Segmentation and Re-assembly (SAR) and to provide error-free transmission. Because the amount of data that can be transmitted over the radio interface at a given point in time is limited, the SAR function is about formatting upper layer data units into low layer data units or vice versa. Because data units transmitted over the radio interface can be lost due to hostile channel environments, the Automatic Repeat reQuest (ARQ) is used for error-free transmission. When a missing data unit is detected, retransmission by the ARQ function is used to guarantee lossless transmission over the radio interface. In addition, due to the use of a multi-channel HARQ function in the lower layer, a reordering function in the RLC layer is used to guarantee in-order delivery of data units.

MAC layer

The MAC layer is located below the RLC layer and above the Physical layer. The main role of the MAC layer is to control the upper layer's access to radio resource. The MAC layer decides how much radio resource should be allocated to data from which logical channel of which UE. To achieve this, the MAC layer performs buffer status reporting to indicate the amount of data in each logical channel, scheduling to allocate radio resource to each UE, logical channel prioritization to allocate radio resource to each logical channel, multiplexing and de-multiplexing to combine and disassemble upper layer data units into/out of lower layer data units, and HARQ to transfer data units over the radio interface. To reduce power consumption and to meet timing requirements, Discontinuous Reception (DRX) and Time Alignment (TA) functions are also performed.

Physical layer

After processing in the above layers, the data units are finally processed and transmitted over the radio interface by the Physical layer. The typical unit of time is called a subframe or Transmission Time Interval (TTI). One subframe is equal to 1 ms, and the transmission opportunity for a MAC PDU is given in the units of subframes.

Several physical channels are defined for actual transmission/reception at the Physical layer: the Physical Downlink Control Channel (PDCCH) is used to deliver control signaling, such as resource allocation information, HARQ ACK/NACK information, and so on; the Physical Downlink Shared Channel (PDSCH) is used to deliver upper layer data units in the downlink direction; and the Physical Uplink Shared Channel (PUSCH) is used to deliver upper layer data units in the uplink direction.

1.6 UE Capabilities

Although the LTE system provides up to 100 Mbps theoretically, this does not mean that all LTE terminals should be able to support 100 Mbps. Depending on the marketing purpose or user segments, the radio access capability of each LTE terminal may differ. For proper configuration of radio parameters, the eNB needs to know the radio capability of each UE. For example, if a UE can process 50 Mbps, allocating radio resource equivalent to 80 Mbps will lead to overflow in the UE's buffer or a waste of radio resource.

In the market, there will be thousands of combinations of UE capability. If information regarding this different capability is delivered to the network as it is, it will incur huge signaling overhead. To reduce the complexity of UE capability signaling, a limited set of UE categories is defined (Table 1.3). The values in the table are the minimum requirement that a UE in each category should meet. When a UE indicates its category to the network, its capability should be at least the value specified for the category. Thus, for example, a UE in category 3 should be able to process at least 1 237 238 soft channel bits.

Table 1.3 UE categories

UE Category	Category 1	Category 2	Category 3	Category 4	Category 5
Maximum number of DL-SCH transport block bits received within a TTI	10 296 (10.3 Mbps)	51 024 (51 Mbps)	102 048 (102 Mbps)	150 752 (150 Mbps)	299 552 (300 Mbps)
Maximum number of bits of a DL-SCH transport block received within a TTI	10 296	51 024	75 376	75 376	149 776
Total number of soft channel bits	250 368	1 237 248	1 237 248	1 827 072	3 667 200
Maximum number of supported layers for spatial multiplexing in DL	1	2	2	2	4
Maximum number of bits of a UL-SCH transport block transmitted within a TTI	5160 (5.1 Mbps)	25 456 (25 Mbps)	51 024 (51 Mbps)	51 024 (51 Mbps)	75 376 (75 Mbps)
Support for 64 QAM in UL	NO	NO	NO	NO	YES
Total layer 2 buffer size (bytes)	150 000	700 000	1 400 000	1 900 000	3 500 000

References

1. 3rd Generation Partnership Project, “About 3GPP”, available at http://www.3gpp.org/About-3GPP, accessed on 14 March 2012.
2. 3rd Generation Partnership Project, “Specification Groups”, available at http://www.3gpp.org/Specification-Groups, accessed on 14 March 2012.
3. 3rd Generation Partnership Project, “LTE”, available at http://www.3gpp.org/LTE, accessed on 14 March 2012.
4. Cisco, “Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2011–2016”, available at http://www.cisco.com, accessed on 14 March 2012.
5. 3GPP Technical Specification 23.203, “Policy and charging control architecture (Release 10)”, www.3gpp.org, accessed on 14 March 2012.

2

Idle Mode Procedure

RRC_IDLE is a UE AS state in which the UE is switched on but does not have any established RRC connection. No RRC connection means that the presence of the UE is, in general, not known to the network at the cell level because eNBs do not have any context for the UE. The location of the UE in RRC_IDLE is known to the network at the level of tracking areas, which consist of cells. The UE AS behaviors applied in RRC_IDLE are defined in [1].

The motivation behind using RRC_IDLE is to reduce UE power consumption by utilizing, for example, discontinuous reception. The number of processes that the UE is required to perform in RRC_IDLE is significantly smaller than in RRC_CONNECTED. For example, the UE in RRC_IDLE is required sometimes to wake up for a short time to perform a limited number of processes, for example, monitoring incoming calls and measurements for mobility, but the UE is soon allowed to sleep again and stays inactive most of the time. Such a long inactive time in RRC_IDLE significantly reduces battery consumption in the UE.

The term "camp on" is often used when UE behaviors in RRC_IDLE are described. "Camp on" is the UE state in which the UE stays on a cell and is ready to initiate a potential dedicated service or to receive an ongoing broadcast service. Camping on a cell by the UE in RRC_IDLE has several purposes:

- reception of system information for the camped cell (see Section 3.2);
- initiation of an RRC Connection Establishment on the camped cell when necessary (see Section 3.4);
- reception of *paging* messages for mobile terminating calls on the camped cell (see Sections 2.9 and 3.3);
- reception of Public Warning System (PWS) notifications (see Section 10.4).

2.1 Idle Mode Functions

In this section, we look briefly at the processes performed in RRC_IDLE before going through the details of each process in the subsequent sections.

Once a UE has been switched on, the UE performs the PLMN selection process [2]. The selected PLMN may be associated with Radio Access Technology. Section 2.4 covers the PLMN selection process.

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

Once the PLMN has been selected, the cell selection process follows in order to find a suitable cell to camp on. The cell selection process is outlined in Section 2.6.

After successful cell selection, the UE attempts to register with an MME by performing the Attach procedure at the NAS layer. If registration is successful, the UE can be reached by the network even in RRC_IDLE. The location registration process is explained in Section 2.5.

While the UE is in RRC_IDLE, it may select another cell due to its mobility or a change of network policy. The details of the cell reselection process are given in Section 2.7.

When a UE attempts to camp on a cell, it has to evaluate whether access to the cell is allowed. The access verification process is described in Section 2.8.

The UE in RRC_IDLE is required to monitor *Paging* messages to receive incoming calls. The paging process is described in Section 2.9.

When camping on a cell, the UE is required to monitor changes in system information and it obtains new system information to ensure that this is kept up to date. The handling of system information is described in Section 3.2.

A UE may support Closed Subscriber Group (CSG)-related functionalities. The details of CSG-related features are given in Section 9.

The UE may also receive PWS messages in RRC_IDLE. The PWS is described in Section 10.

The UE may also receive Multimedia Broadcast/Multicast Services (MBMSs) in RRC_IDLE. The MBMSs over LTE are described in Section 11.

The UE may be configured with Logged MDT to assist network performance optimization. Logged MDT is explained in Section 14.2.

The processes performed in RRC_IDLE are summarized in Table 2.1. The main functions among these processes are: (1) PLMN selection; (2) cell selection and cell reselection; (3) location registration; (4) CSG support, as shown in Figure 2.1. Other processes provide assistance to the main functions or supplementary services.

Table 2.1 RRC_IDLE processes

RRC_IDLE Processes	Purpose
PLMN selection	PLMN is selected for normal service.
Registration	UE location is registered to network.
Cell selection	UE finds a suitable cell to camp on.
Cell reselection	UE finds a better cell to camp on.
Access verification	UE is barred from access to unallowable cell.
Paging reception	UE receives mobile-terminating call.
System information acquisition	UE acquires common information that should be configured to obtain service on a cell.
CSG support	UE gets preferential service from CSG cell when a CSG member.
PWS support	UE receives PWS messages.
MBMS support	UE receives MBMSs.
Logged MDT support	UE collects measurements and reports to assist in network performance optimization.

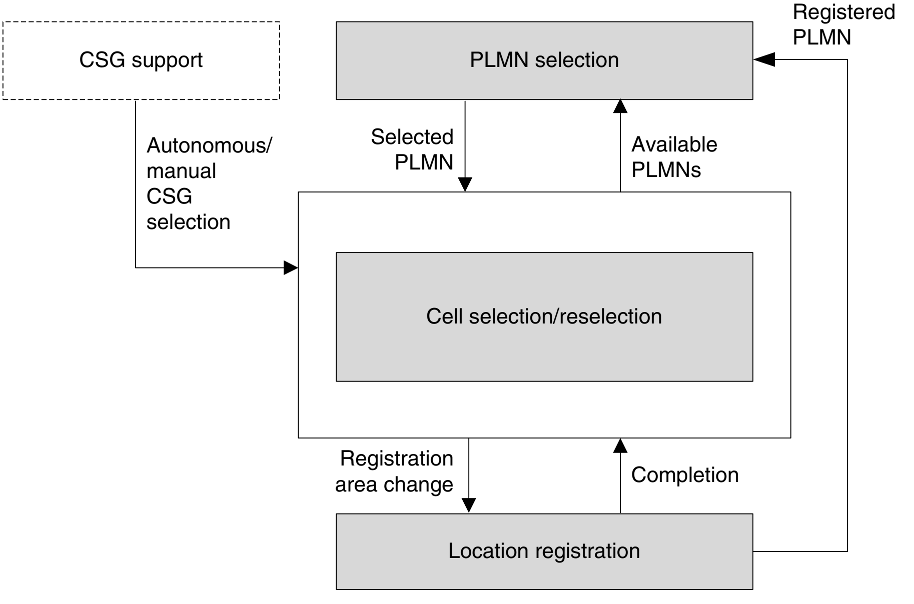

Figure 2.1 Main RRC_IDLE processes. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

2.2 Services and Cell Categorization

The services provided to UEs can be classified as follows:

- **Normal Service:** The public service in the normal case;
- **Limited Service:** The emergency service (emergency calls and PWS messages);
- **Operator Service:** The service for UEs with special access rights.

Cells are also categorized, depending on the service supported on the cell, as follows:

- **Suitable cell:** Normal services are available to UEs;
- **Acceptable cell:** Only limited services can be provided;
- **Barred cell:** No service is available, access to the cell is not permitted;
- **Reserved cell:** Normal services are provided to UEs with special access rights.

Combining the service categories and cell types described above, Figure 2.2 summarizes the association between each service category and cell type.

Depending on which requirements have been satisfied, the cell type seen by a specific UE is decided. Accordingly, the same cell can be of different cell types to different UEs.

In Table 2.2, the requirements for suitable cells and acceptable cells are described. The mark "O" indicates that the corresponding requirement should be fulfilled. The requirements for barred cells and reserved cells are not shown in the table, as the cell is barred/reserved if the system information indicates that to be the case, as described in Section 2.8.

Table 2.2 Requirements for different cell categories

Requirement Category	Requirement Description	Suitable Cell	Acceptable Cell
USIM	The UE should have a valid USIM	○	
PLMN	The cell is part of (1) a selected PLMN by PLMN selection, or (2) a registered PLMN, or (3) a PLMN of the equivalent PLMN (EPLMN)	○	
Tracking Area (TA)	The cell is part of at least one TA, which (1) is not in the "forbidden TA for roaming", (2) belongs to a PLMN fulfilling the PLMN requirement above	○	
Cell selection	The cell selection criterion is satisfied (see Section 2.6)	○	○
Barring	The cell is not barred. (see Section 2.8 for the condition "not barred")	○	○
CSG	The UE is a CSG member of the cell. Note that this requirement applies only for CSG cells (see Section 2.9)	○	

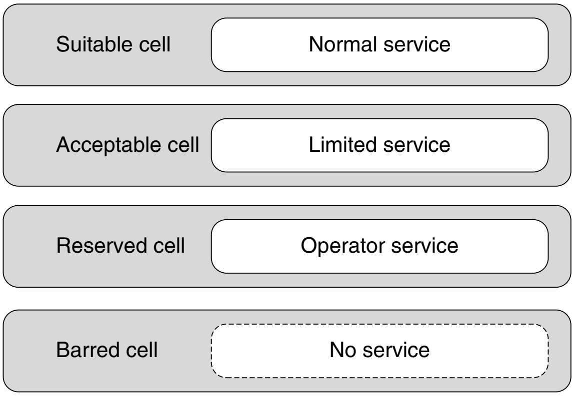

Figure 2.2 Association between cell categories and service types

2.3 UE States and State Transitions

There are three states applicable to UEs in RRC_IDLE: the *CampedNormally* state, the *CampedOnAnyCell* state, and the *AnyCellSelection* state. One state is differentiated from the other states by the cell type of the current serving cell or by whether the UE is camping on a cell or not. The UE transits from one state to another by performing idle mode process(es). Figure 2.3 shows the state transition diagram among these states.

- **CampedNormally:** is the UE state in which the UE in RRC_IDLE is camped on a suitable cell.
- **CampedOnAnyCell:** is the UE state in which the UE in RRC_IDLE is camped on an acceptable cell.

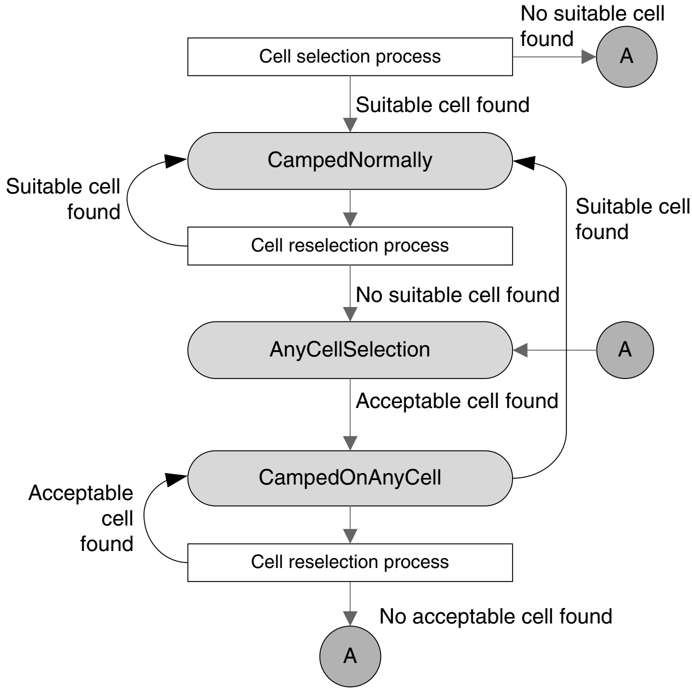

Figure 2.3 State transition in idle mode. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

- **AnyCellSelection:** is the UE state in which the UE in RRC_IDLE attempts to find an acceptable cell to camp on.

If the UE fails to find any suitable cell during cell selection or whilst in the *CampedNormally* state, it enters the *AnyCellSelection* state. The UE also enters the *AnyCellSelection* state if it cannot find any acceptable cell whilst in the *CampedOnAnyCell* state.

Whilst the UE is in the *AnyCellSelection* state, it attempts to find an acceptable cell. Since the requirements for acceptable cells only include fulfillment of the cell selection criterion and "not barred" status, as shown in Table 2.2, the UE searches for an acceptable cell through any PLMN and all RATs supported by the UE. Once the UE finds and selects an acceptable cell, it enters the *CampedOnAnyCell* state.

Whilst the UE is in the *CampedOnAnyCell* state, it regularly searches for a suitable cell through all frequencies of all RATs supported by the UE. Once the UE finds and selects a suitable cell, it enters the *CampedNormally* state.

2.4 PLMN Selection

A Public Land Mobile Network (PLMN) is a network deployed and operated by a mobile network operator (or operators). Each mobile network operator runs one or more PLMNs.

Each PLMN can be identified by a Mobile Country Code (MCC) and a Mobile Network Code (MNC). The PLMN information of a cell is broadcast in the *PLMN-Identity* included in the *SystemInformationBlockType1* [3].

For PLMN selection, cell selection, and cell reselection, several types of PLMN are considered by the UE:

- **Home PLMN (HPLMN):** The PLMN whose MCC and MNC match the MCC and the MNC of the UE's IMSI.
- **Equivalent HPLMN (EHPLMN):** Any PLMN that is equivalent to the HPLMN.
- **Registered PLMN (RPLMN):** The PLMN for which location registration is successful.
- **Equivalent PLMN (EPLMN):** Any PLMN that is equivalent to the RPLMN.

Each mobile service subscriber has a subscription with an HPLMN. When the normal service is provided to a UE by the HPLMN or the EHPLMN, the UE is not in a roaming state. On the other hand, when the service is provided to the UE by a PLMN other than the HPLMN/EPHPLN, the UE is in a roaming state, and the PLMN is called the Visited PLMN (VPLMN).

2.4.1 Triggering of PLMN Selection

When the UE is switched on with a valid USIM, the UE normally attempts to register to the previously registered PLMN to save the time otherwise required to perform PLMN selection. If the registration fails, the UE NAS initiates PLMN selection. PLMN selection may also take place when the UE is lacking service coverage for the selected PLMN. A user may also trigger PLMN selection manually to choose another PLMN.

2.4.2 Search of Available PLMNs

If PLMN selection is triggered, the UE AS performs a search for available PLMNs. During the search, the UE scans all frequencies of E-UTRAN bands supported by the UE. The UE is required to synchronize with the strongest cell on each frequency and read concerned system information to identify the PLMN(s) of the cell.

If the measured RSRP value of the strongest cell is equal to or exceeds $-110\,\text{dBm}$, the PLMN information of the cell is reported to the NAS as a "high quality PLMN". Otherwise, the PLMN information is reported with the measured RSRP value to the UE NAS. The search of available PLMNs can be optimized by utilizing stored information (for example, frequencies and other cell parameters) known to the UE during, for example, previous measurements. The available PLMNs found by the UE AS are reported to the UE NAS. The UE AS may also report the RAT(s) associated with the reported PLMNs.

2.4.3 PLMN Selection

Upon receiving the PLMN information reported by the UE AS, the UE NAS evaluates and performs PLMN selection from among the reported PLMNs. During the PLMN selection, the UE considers the information of the forbidden PLMNs and the equivalent HPLMN, if signaled. The UE also makes use of available information regarding the priority of the PLMN and the associated RAT, stored in its USIM. The information can be provided in an

Table 2.3 Example of information in the Selector data file

Entry #	Description (listed in priority order)
1a	1st PLMN (the highest priority)
1b	1st PLMN Access Technology Identifier
2a	2nd PLMN
2b	2nd PLMN Access Technology Identifier
:	:

"HPLMN Selector with Access Technology" data file, an "Operator controlled PLMN Selector with Access Technology" data file, or a "User Controlled PLMN Selector with Access Technology" data file. The general structure of such a data file is described in Table 2.3. The details of each type of data file can be found in [4].

The entries in Table 2.3 are listed in decreasing priority order. Note that in the data file, each PLMN entry can be listed with the associated RAT. The exact information structure can be found in [2].

There are two modes of PLMN selection:

- **Automatic PLMN selection:** The UE autonomously selects an available PLMN from the reported PLMN list based on priority order.
- **Manual PLMN selection:** A user manually selects a PLMN from the list of PLMNs provided by the UE AS.

In automatic PLMN selection, the UE selects and attempts registration on the PLMN/RAT in the predetermined order as specified in [2] until the registration is successful. The order of the selected PLMN/RAT is as follows:

- PLMNs listed in the HPLMN or the EHPLMN list;
- PLMNs listed in the "User Controlled PLMN Selector with Access Technology";
- PLMNs listed in the "Operator Controlled PLMN Selector with Access Technology";
- PLMNs reported by the AS as high quality PLMNs;
- PLMNs reported by the AS in decreasing order of signal quality.

In manual PLMN selection, the UE displays a list of available PLMN/RATs to a user. When displayed, the order of the PLMN/RATs is almost the same as that used in automatic PLMN selection. After selecting the PLMN, the UE attempts registration for the selected PLMN. The details can be found in [2].

The information relating to the selected PLMN/RAT is indicated to the UE AS. The UE AS then performs cell selection to camp on a suitable cell that belongs to the selected PLMN/RAT. The details of cell selection are given in Section 2.6.

If a suitable cell is found, the UE attempts location registration on the selected PLMN. If the location registration is successful, information on the selected PLMN is presented to a user and the PLMN selection process ends. With the success of location registration, the selected PLMN becomes the RPLMN. During the location registration process, the UE may be provided with information on a list of EPLMNs. If provided, the UE stores the EPLMN information. The UE considers all the PLMNs in the EPLMN list equivalent to the current RPLMN for PLMN selection and mobility; that is, cell selection/reselection and handover.

2.5 Location Registration

Registration to a network is a prerequisite for a UE to obtain services, and this is performed through NAS layer signaling. After successful registration, the UE is ready for a service from its RPLMN or the EPLMN. A UE that has registered with a network should be reachable by the network at any time. If the UE is in ECM-CONNECTED (equivalently RRC_CONNECTED), the network is aware of the cell the UE is being served on. However, while the UE is in ECM-IDLE (equivalently RRC_IDLE), the context of the UE is not available at the eNB but is stored in the MME. In this case, the location of the UE in ECM-IDLE is only known to the MME at the granularity of a list of Tracking Areas (TAs). A single TA is identified by the Tracking Area Identity (TAI), which consists of the PLMN Identity the tracking area belongs to and the Tracking Area Code (TAC) that uniquely represents the TA in the PLMN.

3GPP Release 8 utilizes the concept of the TAI list for location management for the UE in ECM-IDLE to reduce the signaling load from frequent occurrences of Tracking Area Update (TAU). The UE initiates a TAU for the location registration only if the UE enters a cell belonging to a tracking area that is not listed in the TAI list. The tracking area of a cell is broadcast in *SystemInformationBlockType1*. The TAI list is signaled to the UE at location registration.

2.6 Cell Selection

The cell selection process finds and selects a suitable cell to camp on. Cell selection is normally triggered when the UE is not camped on any cell. For this reason, cell selection should be completed as quickly as possible.

Cell selection can be triggered by the PLMN selection, where a suitable cell should be selected for the selected PLMN. Cell selection is also triggered when the UE leaves RRC_CONNECTED, as explained in Section 2.6.2.

The RAT(s) that is (are) subject to the cell selection can be controlled by the NAS. For example, the UE considers cells of the RAT associated with the PLMN chosen by the PLMN selection process described in Section 2.4 to be the cell selection candidates.

In general, there are two types of cell selection, depending on whether the UE utilizes stored information for cell selection:

- **Initial Cell Selection:** The UE performs the cell selection without using any stored information of the E-UTRAN frequencies.
- **Cell Selection with Stored Information:** The cell selection is assisted by stored information on carrier frequencies and, optionally, cell parameters, by which cell selection can be done more promptly.

When cell selection is triggered, cell selection with stored information is triggered first if stored information is available. If cell selection with stored information fails to find a suitable cell, or if cell selection is triggered when no stored information is available, the UE performs the initial cell selection process.

In initial cell selection, the UE scans all the frequencies of E-UTRAN bands supported by the UE and is required to find the strongest cell on each frequency as the cell selection candidate.

If a cell satisfying the cell selection criterion (Section 2.6.1 below) is found during cell selection, the UE selects the cell and evaluates the suitability of the cell, as described in Section 2.2. If the UE fails to find any suitable cell, the UE enters the *AnyCellSelection* state, as described in Section 2.3.

2.6.1 Cell Selection Criterion

The cell selection criterion, referred to as the *S criterion*, evaluates whether the signal level of the cell exceeds the corresponding threshold. The cell selection criterion is one of the conditions that should be satisfied for a cell to be deemed suitable.

For cell selection in Release 8, the Reference Signal Received Power (RSRP) metric is used in the S criterion. The S criterion is extended in Release 9 to further consider the Reference Signal Received Quality (RSRQ) metric. The RSRQ is defined as the ratio of RSRP to the E-UTRAN carrier Received Signal Strength Indicator (RSSI). Although RSRP gives essential information about the "signal strength" of the measured cell, it cannot reflect well the "signal quality" of the cell. RSRQ is a metric that properly reflects the quality of the measured cell, since the RSSI included in the RSRQ calculation reflects the interference level. The use of the RSRQ metric in Release 9 helps a UE to avoid selecting a cell that shows a high level of RSRP but a low level of RSRQ due to interference.

When the RSRQ metric for the S criterion is applicable (i.e., if the UE supports the RSRQ metric for cell selection and the network advertises the relevant thresholds), the UE selects a cell if both S_{rxlev} and S_{qual} exceed 0 dB for the cell, where S_{rxlev} and S_{qual} are defined as:

- $S_{rxlev} = Q_{rxlevmeas} - (Q_{rxlevmin} + Q_{rxlevminoffset}) - P_{compensation}$;
- $S_{qual} = Q_{qualmeas} - (Q_{qualmin} + Q_{qualminoffset})$.

In the S criterion, S_{rxlev} is an RSRP metric and S_{qual} is an RSRQ metric. Assuming that *X* denotes "rxlev" and "qual" for RSRP evaluation and for RSRQ evaluation respectively within this section, the criterion S_X further consists of Q_{Xmeas}, Q_{Xmin}, $Q_{Xminoffset}$, and $P_{compensation}$:

- Q_{Xmeas} is the measured level for the evaluated cell;
- Q_{Xmin} is the required minimum level for cell selection;
- $Q_{Xminoffset}$ is the offset to avoid a ping-pong in selection during the periodic search of higher priority PLMNs at the PLMN boundary;
- $P_{compensation}$ is defined as max($P_{EMAX} - P_{PowerClass}$, 0), where P_{EMAX} is a maximum TX power level at which the UE is allowed to transmit in the cell, and $P_{PowerClass}$ is a maximum RF output power of the UE according to UE power class, as defined in [5]. This offset is to avoid the UE selecting the cell when the transmission power of the UE may not be sufficient.

When the RSRQ metric for the S criterion is not applicable (i.e., if the UE does not support the RSRQ metric for cell selection or the network does not advertise the relevant thresholds), the UE utilizes the S criterion using only S_{rxlev} for cell selection; that is, whether S_{rxlev} exceeds 0 dB.

The parameters related to cell selection described above are broadcast in the *CellSelectionInfo* included in the *SystemInformationBlockType1* message.

Note that the cell selection criterion applies only for E-UTRAN cells. For cell selection for other RATs, a UE should follow the cell selection criterion defined by the specifications of the concerned RAT.

2.6.2 Cell Selection upon RRC Connection Release

When leaving RRC_CONNECTED, the UE performs cell selection to camp on a cell in RRC_IDLE. The UE normally selects the last serving cell where the UE was served whilst in RRC_CONNECTED.

The eNB may include *RedirectedCarrierInfo* in the *RRCConnectionRelease* message in order to redirect the UE to another carrier frequency of the E-UTRAN or another RAT (see Section 3.12). If *RedirectedCarrierInfo* is included, the UE attempts to camp on a suitable cell on the RAT/frequency indicated in the *RedirectedCarrierInfo*. If no suitable cell is found on the indicated RAT/frequency, the UE searches for other carrier frequencies of the indicated RAT. If the UE still fails to find a suitable cell, the UE initiates cell selection with stored information.

Note that when the UE returns to RRC_IDLE after staying in RRC_CONNECTED to which the UE moved from the *CampedOnAnyCell* state, the UE attempts to camp on an acceptable cell.

2.7 Cell Reselection

While the UE is camping on a cell, it attempts to reselect a "better" cell. The criterion to determine a "better" cell is described in this section. Regarding cell reselection, a UE is required to perform the following steps:

- **Step 1:** Measurements of neighboring cells;
- **Step 2:** Evaluation of the measured results of neighboring cells;
- **Step 3:** Cell reselection and access verification.

For step 1, the network provides the UE with the reselection information indicating the carrier frequencies that the UE is required to measure. The cell reselection information is provided in system information through the *SystemInformationBlockType3* to the *SystemInformationBlockType8* for the E-UTRAN and the other RATs, for example, GERAN, the UTRAN, and the CDMA2000 1xRTT/HRPD.

When the carrier frequency for cell reselection is indicated, the network also provides the reselection priority of the indicated frequency/RAT to the UE. The reselection priority is one of the key parameters to control UEs' camping behaviors across frequencies/RATs, where the basic principle of cell reselection is that the UE is required to reselect a suitable cell of the highest priority frequency. For example, a network may concentrate the UEs in RRC_IDLE on a certain carrier/RAT or it may distribute them over several carrier frequencies by configuring the reselection priorities, depending on the operator's policy.

In addition to the broadcast reselection priorities, the reselection priorities can be configured in a UE-specific manner. Dedicated reselection priorities can be signaled to the UE, optionally with the validity time of the dedicated priorities, when the network releases the RRC connection of the UE via the *RRCConnectionRelease* message. After the UE has received the dedicated reselection priorities, it applies the dedicated reselection parameters, rather than the broadcast reselection parameters, for reselection during RRC_IDLE. The dedicated reselection parameters are discarded when the UE enters RRC_CONNECTED, PLMN selection is performed, or the validity timer expires.

2.7.1 Measurement Rules

In RRC_IDLE, the UE should perform measurements on one or more frequencies for which frequency priorities have been provided by the network. The UE does not perform measurements on frequencies for which priorities are not known to the UE.

2.7.1.1 Measurement Requirements

The UE performs measurements according to the measurement requirements specified in [6]. The requirements for different measurement targets are different, as follows:

- Intra-frequency measurement requirements:
  - Serving cell: The UE measures the RSRP and RSRQ level of the serving cell at least once every DRX cycle.
  - Intra-frequency neighboring cell: When the UE is required to perform intra-frequency measurements, it measures RSRP and RSRQ at least every $T_{measure,EUTRAN\_Intra}$.
- Inter-frequency measurement requirements:
  - Neighboring cell of higher frequency priority: For a neighboring cell whose frequency priority is higher than that of the serving cell, the UE periodically searches for cells on the higher priority frequency for cell reselection at least every $T_{higher\_priority\_search} = (60 * N_{layers})$ seconds, where N_{layers} is the total number of configured higher priority carrier frequencies including the E-UTRA, the UTRA FDD, the UTRA TDD, and the CDMA2000 1xRTT/HRPD. If one or more groups of GSM frequencies is also configured as having higher priority, N_{layers} is incremented by one. For higher priority cells that are identified, the UE shall perform measurements at least every $T_{measure,EUTRAN\_Inter}$.
  - Neighboring cell of equal or lower frequency priority: For an E-UTRAN neighboring cell whose frequency priority is equal to or lower than that of the serving cell, the UE shall perform measurements at least every $K_{carrier} * T_{measure,EUTRAN\_Inter}$ where $K_{carrier}$ is the number of E-UTRAN carrier frequencies indicated by the serving cell in the *SystemInformationBlockType5*.

2.7.1.2 Power-efficient Measurements

Typically, one operator may provide more than one carrier frequency for E-UTRAN and/or other RATs. If a UE was always required to measure all carrier frequencies of those RATs, it would result in quite heavy battery consumption for the UE. To avoid such wasteful battery consumption, the network can allow a UE to utilize a more efficient measurement mechanism. The key to such efficient measurement is that the UE skips some measurements that are considered less urgent. For example, if the measured level of the serving cell is quite good enough, the measurements of some neighboring cells can be assumed to be less critical and thus omitted.

For a UE to determine whether it is allowed to apply efficient measurements, only the RSRP metric of the serving cell (S_{rxlev}) is used in Release 8 while both the RSRP metric and RSRQ metric (S_{rxlev} and S_{qual}) are used in Release 9 and onwards. The joint consideration of RSRP and RSRQ in Release 9 is to avoid unnecessary cell reselection to a new cell for which the measured value of RSRP is high but the measured value of RSRQ is low.

Depending on the values of S_{rxlev} and S_{qual}, the efficient measurements are applied differently to intra-frequency measurements and inter-frequency measurements:

- **Intra-frequency measurements:** The UE can skip measurements of intra-frequency neighboring cells if S_{rxlev} exceeds the RSRP threshold ($S_{IntraSearchP}$) and S_{qual} exceeds the RSRQ threshold ($S_{IntraSearchQ}$).
- **Inter-frequency measurements:** The UE can skip measurements of inter-frequency neighboring cells if S_{rxlev} exceeds the RSRP threshold ($S_{nonIntraSearchP}$) and S_{qual} exceeds the RSRQ threshold ($S_{nonIntraSearchQ}$).

The relevant threshold parameters – that is, $S_{IntraSearch}$ and $S_{nonIntraSearch}$ – are optionally provided in *SystemInformationBlockType3*. If absent, the default value of infinity is applied for $S_{IntraSearchP}$ and $S_{nonIntraSearchP}$, and the default value of zero is applied for $S_{IntraSearchQ}$ and $S_{nonIntraSearchQ}$. Note that setting the RSRP threshold to infinity will result in the inequality that S_{rxlev} never exceeds the RSRP threshold, while setting the RSRQ threshold to zero will make S_{qual} always exceed the RSRQ threshold. Figure 2.4 illustrates the required measurements depending on the measured values of S_{rxlev} and S_{qual} of the serving cell, assuming that $S_{IntraSearch}$ is higher than $S_{nonIntraSearch}$.

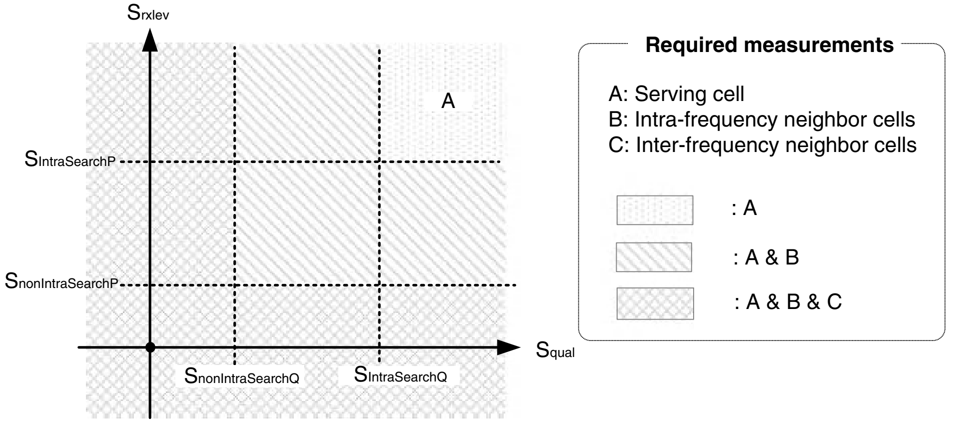

Figure 2.4 Power-efficient measurements

2.7.2 Reselection to a Neighboring Cell

When measurement results are available, the UE evaluates the measurement results for cell reselection. There are two types of reselection criteria: (1) criterion for reselection within equal priority cells, and (2) criterion for reselection across non-equal priority cells. The first criterion applies to reselection to a cell whose frequency priority is the same as that of the serving cell. The second criterion applies to cell reselection to a cell whose frequency priority is different from that of the serving cell.

Note that when reselection criteria are evaluated, as described below, the following are commonly and jointly applied:

- **Reselection Time:** When evaluating a neighboring cell with the S metric (S_{rxlev} and/or S_{qual}) or the R metric (R_n), as defined below, the evaluation criterion can be met only if the value of the evaluated metric keeps exceeding the relevant threshold during the reselection time interval. This time interval is denoted $T_{reselection}$ and can be configured per RAT. $T_{reselection}$ is used to avoid the ping-pong that could happen otherwise, for example, if reselection is triggered towards a cell that meets the criterion only for a time interval shorter than $T_{reselection}$.
- **Minimum Serving Cell Residence Time:** To prevent cell reselections from happening too frequently during a short time, cell reselection can be performed only if the UE has camped on the serving cell for more than 1 second.

2.7.2.1 Reselection to Equal Priority Cells

This is the case when the UE performs reselection evaluation towards a neighboring cell whose frequency priority is the same as that of the serving cell. This corresponds to the reselection to an intra/inter-frequency neighboring cell having the same frequency priority as the serving cell.

For cells of the same frequency priority, the cell-ranking criterion (R criterion) is used to determine which cell the UE should reselect. For the R criterion, the UE calculates the R values of its serving cell (R_s) and neighboring cells (R_n) of the same frequency priority, as defined below:

- $R_s = Q_{meas,s} + Q_{Hyst}$ for serving cell;
- $R_n = Q_{meas,n} - Q_{offset}$ for each neighboring cell.

The R criterion further consists of Q_{meas}, Q_{Hyst}, and Q_{offset}:

- Q_{meas} is the measured value in RSRP;
- Q_{Hyst} is a hysteresis value applied to the serving cell;
- Q_{offset} is an offset applied to the evaluated neighboring cell.

The UE performs ranking of the cells based on the calculated R values. If the neighboring cell is better ranked than the serving cell – that is, $R_n > R_s$ – the UE reselects the neighboring cell, desirably to the highest-ranked neighboring cell.

On the reselected cell, the UE needs to perform access verification by reading the system information of the cell, as explained in Section 2.8. If the highest-ranked cell is verified as being suitable, the cell reselection becomes successful and the UE can camp on the cell.

2.7.2.2 Reselection to Non-equal Priority Cells

This is the case when the UE performs reselection evaluation on neighboring cells whose frequency priority is different from that of the serving cell; that is, reselection to a neighboring cell of higher frequency priority or a neighboring cell of lower frequency priority.

In Release 8, the reselection evaluation towards non-equal frequency priority cells only utilizes the RSRP metric (S_{rxlev}). In Release 9, this is extended to consider the RSRQ metric (S_{qual}), when applicable. If the network broadcasts the RSRQ threshold in *threshold ServingLowQ* included in *SystemInformationBlockType3*, the UE applies the RSRQ criterion – that is, using the RSRQ metric (S_{qual}) and the relevant thresholds – otherwise, the UE applies the RSRP criterion – that is, using the RSRP metric (S_{rxlev}) and the relevant thresholds for reselection evaluation. The thresholds applicable for Release 9 include $Thresh_{Serving,LowQ}$, $Thresh_{RAT,HighQ}$, and $Thresh_{RAT,LowQ}$ for RSRQ, and $Thresh_{Serving,LowP}$, $Thresh_{RAT,HighP}$, and $Thresh_{RAT,LowP}$ for RSRP, respectively. Note that the thresholds $Thresh_{Serving,LowP}$, $Thresh_{RAT,HighP}$, and $Thresh_{RAT,LowP}$ correspond to $Thresh_{Serving,Low}$, $Thresh_{RAT,High}$, and $Thresh_{RAT,Low}$ respectively in the Release 8 specification. Note also that the RSRQ metric is applicable only for E-UTRAN and UTRAN FDD neighboring cells. For UTRAN TDD, GERAN, and CDMA2000 neighboring cells, only the RSRP metric is applicable.

The UE utilizes different evaluation criteria for reselection towards a neighboring cell of different frequency priority, depending on the priority of the evaluated cell:

- **Reselection to a Cell on a Higher Priority Frequency:** If a cell on a higher priority frequency is detected, the UE reselects the cell if the evaluated metric of the cell exceeds the relevant threshold ($Thresh_{RAT,HighQ}$ or $Thresh_{RAT,HighP}$). More specifically,
  - If the UE supports RSRQ-based evaluation (for example, Release 9 UE) and the network advertises the relevant RSRQ threshold, the UE should evaluate the RSRQ criterion; that is, whether S_{qual} of the evaluated cell exceeds $Thresh_{RAT,HighQ}$. If this criterion is met, the UE reselects the evaluated cell.
  - If the UE does not support RSRQ-based evaluation (for example, Release 8 UE) or the network does not advertise the RSRQ threshold, the UE should evaluate the RSRP criterion; that is, whether S_{rxlev} of the evaluated cell exceeds $Thresh_{RAT,HighP}$. If this criterion is met, the UE reselects the evaluated cell.

Figure 2.5 illustrates an example of reselection to a higher priority frequency cell, where the UE reselects a cell (1a) on frequency 1 (F1) whose priority is higher than the serving frequency (F2) even though the measured level of the reselected cell (1a) is lower than that of the current serving cell (2a). When reselection is evaluated towards the higher-priority cell, the serving cell quality is not considered, which allows for reselection to a cell of a higher priority even when the serving cell quality is sufficiently good enough.

- **Reselection to a Cell on a Lower Priority Frequency:** If a cell on a lower priority is detected, the UE reselects the cell if the quality of the serving cell is lower than the relevant threshold ($Thresh_{Serving,LowQ}$ or $Thresh_{Serving,LowP}$) and the quality of the target cell exceeds the relevant threshold ($Thresh_{RAT,LowQ}$ or $Thresh_{RAT,LowP}$). More specifically,
  - If the UE supports RSRQ-based evaluation and the network advertises the relevant RSRQ thresholds, the UE should evaluate the RSRQ criterion; that is, whether S_{qual} of

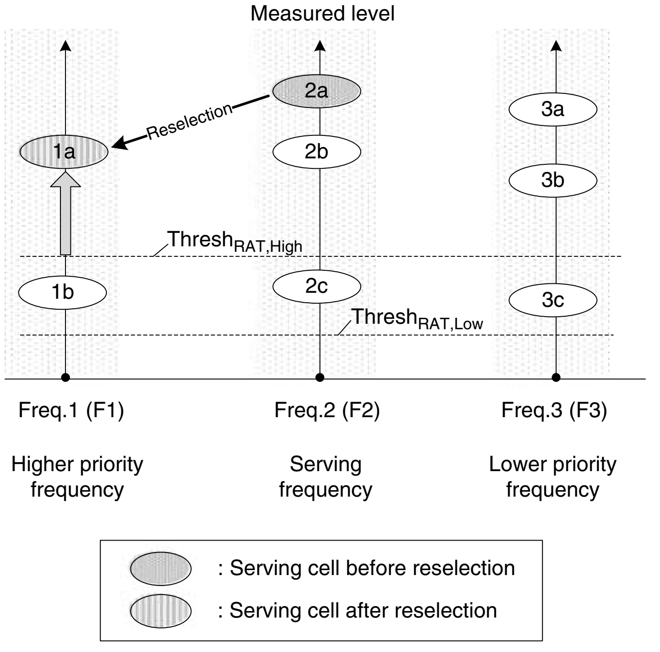

Figure 2.5 Reselection to a higher priority frequency cell

the serving cell is lower than $Thresh_{Serving,LowQ}$ and whether S_{qual} of the evaluated cell exceeds $Thresh_{RAT,LowQ}$. If both criteria are met, the UE reselects the evaluated cell.

– If the UE does not support RSRQ-based evaluation (for example, Release 8 UE) or the network does not advertise the RSRQ thresholds, the UE should evaluate the RSRP criterion; that is, whether S_{rxlev} of the serving cell is lower than $Thresh_{Serving,LowP}$ and whether S_{rxlev} of the evaluated cell exceeds $Thresh_{RAT,LowP}$. If both criteria are met, the UE reselects the evaluated cell.

Figure 2.6 illustrates an example of reselection to a lower priority frequency cell, where the UE reselects a cell (3a) on frequency 3 (F3) whose priority is lower than the serving frequency (F2). When reselection is evaluated towards the lower-priority cell, serving cell quality is taken into account in the evaluation condition, which makes the reselection to a cell of a lower priority frequency take place only when the measured level of the serving cell is not good enough.

2.7.3 Mobility State Dependent Scaling

The optimal mobility parameters may be largely dependent on the UE mobility states. For example, it would be desirable for a high-mobility UE to apply a shorter $T_{reselection}$ value such that reselection to a target cell could take place more quickly, because otherwise the UE may lose coverage of the serving cell before reselection has been initiated or completed. To enable such optimization for the UE in RRC_IDLE, the mobility state of the UE needs to be known to the UE itself.

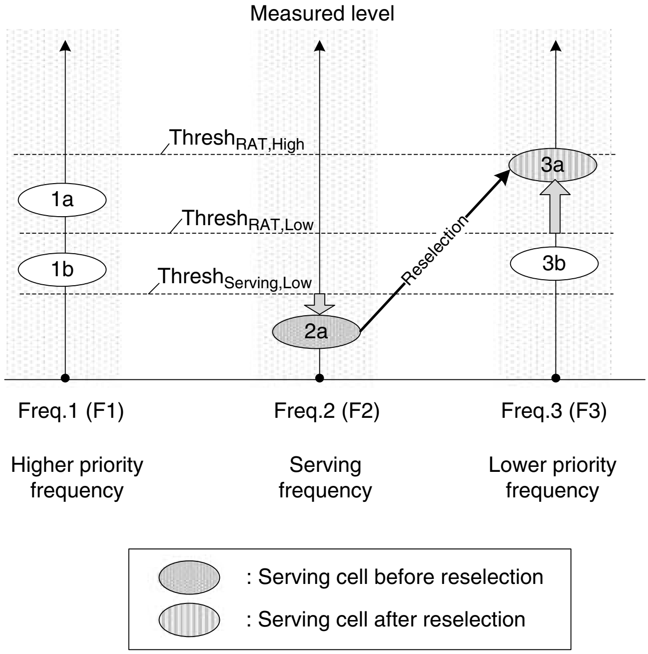

Figure 2.6 Reselection to a lower priority frequency cell

2.7.3.1 Mobility State Detection

In RRC_IDLE, the mobility state can be detected at the UE by relying on reselection counting, generally assuming that a higher number of reselections indicates higher mobility and a lower number of reselections indicates lower mobility. The UE can detect one of three mobility states: Normal-mobility, Medium-mobility, or High-mobility.

Mobility state detection is enabled if the serving cell provides the relevant parameters (T_{CRmax}, $T_{CRmaxHyst}$, N_{CR_H}, and N_{CR_M}) in the *speedStateReselectionPars* included in the *SystemInformationBlockType3*. If the number of reselections during T_{CRmax} exceeds N_{CR_H}, the UE determines that it is in the High-mobility state. If the number of reselections during T_{CRmax} exceeds N_{CR_M} but is equal to or less than N_{CR_H}, the UE considers itself to be in the Medium-mobility state. If the criterion for High-mobility or for Medium-mobility is not met for $T_{CRmaxHyst}$, the UE considers itself to be in the Normal-mobility state. Note that during reselection counting, consecutive reselections between the two cells (i.e., from cell A to cell B, and then from cell B to cell A) are counted only once.

2.7.3.2 Scaling of Reselection Parameters

If the UE detects a High- or Medium-mobility state, it applies speed-dependent scaling to the reselection parameters. The reselection parameters that are subject to the speed-dependent scaling are Q_{hyst} and $T_{reselection}$, which are provided in *SystemInformationBlockType3*. If the UE is in High-/Medium-mobility, these reselection parameters are scaled down by applying High-/Medium-mobility scaling parameters, respectively, such that the reselection can be performed more quickly to adapt to a fast-varying channel; that is, Q_{hyst} becomes smaller and $T_{reselection}$ is shortened. The UE then applies the scaled parameters for further reselection evaluations. For a UE in a Normal-mobility state, no scaling of the reselection parameters is applied; that is, the broadcast values of Q_{hyst} and $T_{reselection}$ are used without scaling.

2.8 Access Verification

Upon selecting/reselecting a new cell, the UE performs access verification to check if the selected cell is suitable to camp on. During access verification, the UE may observe that the cell is barred, or the UE may realize that the cell is not suitable due to forbidden TAs for roaming/non-allowed PLMN. How the UE behaves in such cases differs depending on the result of the access verification. UE behaviors subsequent to access restriction imposed by the network are explained first, and the descriptions on UE actions upon selecting a cell that belongs to a forbidden TA for roaming/non-allowed PLMN follow.

2.8.1 Cell Barring Status and Cell Reservation Status

The network may impose access restriction to UEs in RRC_IDLE such that the UEs cannot even camp on a cell. This kind of access restriction can be enforced by configuring a cell barring status or a cell reservation status.

In addition, the network may control access attempts at RRC connection establishment procedures while allowing the UEs to camp on a cell. The Access Class Barring (ACB) mechanism is the mechanism that prevents some UEs from entering RRC_CONNECTED on the concerned cell, and is applied only when the UE initiates an RRC connection establishment. The details of ACB are given in Section 3.4.1.

In general, access restriction is imposed by broadcasting access-restriction-related information for the UEs in RRC_IDLE. The access-restriction-related information is provided in the *SystemInformationBlockType1* (for cell barring/reservation status) and the *SystemInformationBlockType2* (for ACB) on the E-UTRAN cell.

Every cell should indicate its cell barring status – whether it is a barred cell or not. The cell barring status is advertised by *cellBarred* in the *SystemInformationBlockType1*, which could be set to "barred" or "not barred". The cell should also indicate whether the cell is reserved for operator use, and this information is broadcast by *cellReservedForOperatorUse* in the *SystemInformationBlockType1*, which is set to either "reserved" or "not reserved". Note that a single value of cell barring status is broadcast, while cell reservation status can be broadcast per PLMN in the cell.

When a UE performs access verification, the UE behaves as follows, depending on the combination of the cell barring status and the cell reservation status:

- **Case 1:** If the cell barring status is set to "not barred" and if the cell reservation status is set to "not reserved", a UE is allowed to camp on the cell.
- **Case 2:** If the cell barring status is set to "not barred" and if the cell reservation status is set to "reserved", the cell should be considered "not barred" only for the UEs assigned to AC11 or AC15 operating in their HPLMN/EHPLMN. The UEs assigned to an AC within {0, . . . ,9} or {12, 13, 14} should consider the cell "barred" for the registered PLMN or selected PLMN.
- **Case 3:** If the cell barring status is set to "barred", all UEs should consider the cell "barred".

If a cell is considered "barred", the UE is not allowed to camp on the cell. Even an emergency call is not allowed in a "barred" cell. Once the cell turns out to be barred, the UE needs to move to another cell to camp on. Here, the question is whether the UE is allowed to reselect a cell on the same frequency or not. From an interference point of view, allowing the UE to select another cell on the same frequency may not be desirable. However, if there is no other frequency available, there is no choice other than to stay on the same frequency. To manage this situation, the network advertises a one-bit indicator, denoted *intraFreqReselection*, which controls whether the UE is allowed to reselect another cell on the same frequency if the selected cell is "barred". The *intraFreqReselection* indicator is provided in the *SystemInformationBlockType1*. If the indicator is set to "allowed", the UE is allowed to consider cells on the same frequency as the cell selection/reselection candidates except the barred cell. Otherwise, the UE should exclude all the cells on the same frequency as the cell selection/reselection candidates. Access restriction to the barred cell(s) is applied for 300 seconds. The *intraFreqReselection* indicator is not applied if the barred cell is a CSG cell. The details for cell selection/reselection to/from a CSG cell are explained in Section 9.5.

2.8.1.1 Unsuitable Cell due to Forbidden TAs for Roaming/Non-allowed PLMN

Even if an access attempt is allowed for a cell, based on the cell barring and reservation status, access verification may reveal that the reselected cell is not suitable, for the reason that the reselected cell belongs to the "list of forbidden Tracking Areas for roaming" or the PLMN is not equivalent to the RPLMN.

If a reselected cell that is unsuitable due to the reason(s) above is an E-UTRAN cell, the UE excludes all the cells on that frequency from being reselection candidates for a maximum of 300 seconds. If a reselected cell that is unsuitable for the same reason(s) is an inter-RAT cell, the UE excludes only that cell from being a reselection candidate, and the other cells on the same frequency, if available, can still be considered for reselection.

If a certain condition is met, the restrictions on cell reselection, if applied, are released. For example, if the UE enters the *AnyCellSelection* state, or if the network redirects the UE to a frequency where the (maximum) 300 seconds timer is already running, any restrictions are released such that the UE can find an acceptable cell quickly. If the highest-ranked cell on the frequency changes, the UE does not apply any restriction on that frequency.

2.9 Paging Reception

From the network side, if there is a terminating call towards a UE in RRC_IDLE, the MME sends paging messages over the S1 interface – that is, an S1 Application Protocol (S1AP) paging message [7] – to relevant eNBs belonging to the tracking area(s) at which the UE is expected to be placed. The paging message includes a paging identity of the paged UE and tracking area(s) for which the paged UE is expected to be located. The paging identity indicates the UE identity represented in the form of either an SAE-Temporary Mobile Subscriber Identity (S-TMSI) or an International Mobile Subscriber Identity (IMSI). Upon receiving the paging message from the MME, the eNB performs paging of the UE over the radio interface by sending *Paging* messages on the cells belonging to the tracking areas indicated in the S1AP paging message. The *Paging* message includes a list of paging records, where each paging record contains the paging identity of the paged UE and the core network domain that originated the paging for the paged UE.

From the UE side, to receive the *Paging* message, the UE in RRC_IDLE is required to monitor the physical control channel, referred to as the Physical Downlink Control Channel (PDCCH), masked by the Paging-Radio Network Temporary Identity (P-RNTI), addressing the *Paging* message. The UE is required to monitor the PDCCH on particular occasions for *Paging* reception. The subframes where the *Paging* message can be transmitted to the UE are defined by the Paging Frame (PF) and the Paging Occasion (PO) within the PF. The PF contains one or more POs. The allocation of PO(s) is made by utilizing the International Mobile Subscriber Identity (IMSI) of the paged UEs such that the POs allocated to different UEs (so different IMSIs) are spread over subframes. The calculation of the PF and the PO can be found in [1]. When a UE in RRC_IDLE uses DRX operation, it needs to monitor at least one PO during a paging DRX cycle.

When decoding the *Paging* message, the UE considers that it is being paged only if its UE identity exists in the paging record(s). The UE ignores a *Paging* message that does not include its UE identity unless the *Paging* message is used for other purposes, for example, system information change notification.

References

1. 3GPP Technical Specification 36.304, "User Equipment (UE) procedures in idle mode (Release 10)", www.3gpp.org, accessed on 14 March 2012.
2. 3GPP Technical Specification 23.122, "Non-Access-Stratum (NAS) functions related to Mobile Station (MS) in idle mode (Release 10)", www.3gpp.org, accessed on 14 March 2012.
3. 3GPP Technical Specification 36.331, "Radio Resource Control (RRC); Protocol Specification (Release 10)", www.3gpp.org, accessed on 14 March 2012.
4. 3GPP Technical Specification 31.102, "Characteristics of the Universal Subscriber Identity Module (USIM) application (Release 10)", www.3gpp.org, accessed on 14 March 2012.
5. 3GPP Technical Specification 36.101, "User Equipment (UE) radio transmission and reception (Release 10)", www.3gpp.org, accessed on 14 March 2012.
6. 3GPP Technical Specification 36.133, "Requirements for support of radio resource management (Release 10)", www.3gpp.org, accessed on 14 March 2012.
7. 3GPP Technical Specification 36.413, "S1 Application Protocol (S1AP) (Release 10)", www.3gpp.org, accessed on 14 March 2012.

3

Radio Resource Control (RRC)

The Radio Resource Control (RRC) layer controls communications between a UE and an eNB at the radio interface and the mobility of a UE crossing cells. The RRC is the highest layer in the control plane of the Access Stratum (AS). The RRC also transfers messages of the Non-Access Stratum (NAS), which is located above the RRC layer. NAS messages are used to control communications between a UE and the Evolved Packet Core (EPC).

Signaling used to control transmissions of user data at the radio interface is provided by the RRC layer, the MAC layer, or the Physical layer in the AS. Compared to signaling provided by the MAC layer (i.e., MAC Control Elements, see Section 6.11.1) and signaling provided by the Physical layer (i.e., control information on the PDCCH), signaling provided the RRC layer (i.e., RRC messages) has the following characteristics:

- **Higher extensibility:** RRC messages can be extended easily to accommodate additional control information, for example, for enhancements in future releases of LTE.
- **Secured transmission:** Integrity protection and ciphering of an RRC message carried on the DCCH can be provided by the PDCP layer. Neither integrity protection nor ciphering is applicable for signaling provided by the MAC layer or the Physical layer.
- **Reliable transmission:** Missing parts of an RRC message carried on the DCCH can be retransmitted by ARQ operation provided in the RLC layer and HARQ operation provided in the MAC layer. The RRC layer may also repeatedly transmit the same RRC message, such as system information, to increase the reliability.
- **Longer processing delay:** All RRC messages should be processed by lower layers (i.e., the PDCP layer, the RLC layer, the MAC layer, and the Physical layer) to be transmitted on the radio interface. Due to the processing in lower layers, control by the RRC layer is slower relative to control by other layers.

In a UE, there are two RRC states: RRC_IDLE and RRC_CONNECTED. If an RRC connection exists between the UE and the eNB, the UE is in RRC_CONNECTED. If no RRC connection exists between the UE and the eNB, the UE is in RRC_IDLE.

If the UE is in RRC_IDLE, the UE performs cell selection and reselection based on parameters provided by the RRC in the network, as explained in Section 2. If the UE is in

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

RRC_CONNECTED, the UE maintains an RRC connection with the eNB and communicates with the network via dedicated resources at the radio interface. The UE in RRC_CONNECTED also performs mobility under the control of the network.

In the UTRAN, the connected mode consists of four substates (i.e., CELL_DCH, CELL_FACH, CELL_PCH, and URA_PCH), which makes the RRC layer of UTRAN complicated, due to, for example, defining different UE behaviors in different substates and handling state transitions among four substates. Compared to the RRC layer of UTRAN, the RRC layer of E-UTRAN is much simpler, in that it has only one state in RRC_CONNECTED. Hence, the RRC layer of E-UTRAN provides much simpler operation than the RRC layer of the UTRAN.

The RRC states in the AS are linked to the EPS Connection Management (ECM) states of the NAS because the NAS signaling connection between the UE and the MME consists of the concatenation of the RRC connection (via the Uu interface between the UE and the eNB) and the S1 connection (via the S1 interface between the eNB and the MME). When the RRC connection is established, the UE enters ECM-CONNECTED as well as RRC_CONNECTED and considers that the NAS signaling connection is established. Conversely, when the RRC connection is released, the UE enters ECM-IDLE as well as RRC_IDLE and considers that the NAS signaling connection is released.

It should be noted that the UE normally attaches to the Evolved Packet Core (EPC) for services in the Evolved Packet System (EPS). Attaching to the EPC is handled by the EPS Mobility Management (EMM) protocol of NAS, which is located above the RRC. When the UE is attached to the EPC, the UE is registered in the MME, by which the UE enters EMM-REGISTERED. In this state, the MME establishes a UE context and a default EPS bearer enabling always-on IP connectivity for the UE.

3.1 RRC Functions and Architecture

The Radio Resource Control (RRC) layer performs the following functions:

- Broadcast of system information:
  - information applicable for UEs in RRC_IDLE;
  - information applicable for UEs in RRC_CONNECTED;
  - NAS information.
- Paging.
- Establishment/release of an RRC connection.
- Transfer of NAS information.
- AS security configuration.
- Transfer of UE radio access capability.
- Radio resource configuration:
  - addition, modification, and release of radio bearers;
  - configuration of lower layers (i.e., PHY, MAC, RLC, and PDCP).
- Measurement configuration and reporting:
  - addition, modification, and removal of measurement configurations;
  - measurement reporting.
- Mobility control:
  - intra-EUTRA handover;

  - inter-RAT handover;
  - re-direction/cell change order.
- Recovery from failures of AS layers by re-establishment of an RRC connection.

The RRC messages are carried on the Broadcast Control Channel (BCCH), the Paging Control Channel (PCCH), the Common Control Channel (CCCH), or the Dedicated Control Channel (DCCH). Apart from system information carried on the BCCH and paging messages carried on the PCCH, RRC messages are exchanged between the UE and the eNB over the Signaling Radio Bearer (SRB) by using either the CCCH or the DCCH. There are three types of SRB in E-UTRAN: SRB0, SRB1, and SRB2.

SRB0 transports RRC messages on the CCCH using Transparent Mode (TM) RLC. The RRC messages carried on SRB0 are used only in limited cases, such as when a DCCH logical channel is not configured (i.e., for the RRC Connection Establishment procedure) or when all the other SRBs are suspended (i.e., for the RRC Connection Re-establishment procedure). The logical channel identity of SRB0 is fixed at "0".

SRB1 and SRB2 transport RRC messages on the DCCH using Acknowledged Mode (AM) RLC. Since AM RLC provides reliable transmission using the Automatic Repeat reQuest (ARQ), RRC messages transmitted on SRB1 and SRB2 are not likely to be lost. The logical channel identities of SRB1 and SRB2 are fixed at "1" and "2", respectively.

Between SRB1 and SRB2, the latter is used to transport lower priority RRC messages. Thus, most RRC messages are carried on SRB1, and only the RRC messages transferring NAS information (i.e., the Downlink/Uplink Information Transfer messages) or MDT logged measurement results (see Section 14.2) are carried on SRB2. Exceptionally, when SRB2 is not established, or when NAS information is piggybacked in an RRC message during the RRC Connection Establishment procedure, the RRC messages transferring NAS information can be transmitted on SRB1.

The Radio Bearer (RB) configuration of the UE is normally provided by the eNB using RRC procedures. However, RB configurations for the BCCH, PCCH, and SRB0 are fixed, as specified in [1]. The specified configurations enable UEs to read system information, read *Paging* messages, and perform the RRC Connection Establishment procedure, even without the RRC connection.

Table 3.1 explains how reliability and security are guaranteed for RRC message transmission. For most RRC messages, reliability is guaranteed by the ARQ and HARQ operations provided by the RLC and MAC layers, respectively, and security is guaranteed by the ciphering and integrity protection functions provided by the PDCP layer. However, for RRC messages carried on the BCCH, PCCH, and SRB0, reliability is guaranteed by repeated transmission in the RRC layer, and security is not guaranteed in the AS layer. This is because these RRC messages use TM RLC and bypass the PDCP layer.

The RRC layer performs various RRC procedures. The RRC procedures and corresponding RRC messages in LTE Release 8 are summarized in Table 3.2. Details of these are given throughout this section. Note that additional RRC procedures and messages have been introduced in Release 9 and Release 10 in order to support new features.

RRC messages requiring a response message and RRC messages that are responses to request messages contain an RRC transaction identifier. The same value of transaction identifier is used to identify the RRC procedure to which the RRC messages belong.

Table 3.1 Layer 2/3 functions for RRC message transmission

	BCCH on BCH/ PCCH on PCH	BCCH on DL-SCH	SRB0 (CCCH)	SRB1/SRB2 (DCCH)
RRC layer	Repetition of message transmission	Repetition of message transmission	N/A	N/A
PDCP layer	No integrity protection and ciphering	No integrity protection and ciphering	No integrity protection and ciphering	Integrity protection and ciphering are applicable
RLC layer	Transparent mode	Transparent mode	Transparent mode	Acknowledged mode (ARQ retransmissions)
MAC layer	N/A	HARQ retransmissions without feedback	HARQ retransmissions with feedback	HARQ retransmissions with feedback

3.2 System Information

System information provides UEs in both RRC_IDLE and RRC_CONNECTED with essential information, such as cell access information and common radio resource configuration. The system information also includes the information applicable only for UEs in RRC_IDLE, for example, cell selection and reselection information. The system information is broadcast via RRC messages.

A UE is required to obtain system information for a cell before making an attempt to access the cell. The UE is also required to obtain system information upon selecting a cell, upon reselecting a cell, after completing handover, upon returning from out of coverage, after receiving a notification that the system information is to be changed, and so on.

System information consists of a number of information blocks in the RRC layer. There are several types of information block, as described in Table 3.3. Among the information blocks, the block that contains the most essential and most frequently transmitted system information is called the *MasterInformationBlock* (MIB), and other information blocks except the MIB are called *SystemInformationBlocks* (SIBs).

Note that SIB9 through SIB13 are not listed in Table 3.3. SIB9 is described in Section 9.3 for HeNB; SIB10 through SIB12 are described in Section 10.4.1 for PWS; and SIB13 is described in Section 11.5.2 for MBMS.

3.2.1 Scheduling of System Information

For transmissions of system information blocks, each information block is mapped to one of three types of RRC message: the *MasterInformationBlock* (MIB), the *SystemInformationBlockType1* (SIB1), and the *SystemInformation*, as shown in Figure 3.1.

Table 3.2 Summary of RRC procedures and messages in LTE Release 8Ⓢ: a message used upon success of the corresponding procedure; Ⓕ: a message used upon failure of the corresponding procedure

Procedure	Messages	Logical Channel	Purpose
System Information Acquisition	MasterInformationBlock(DL) SystemInformationBlockType1(DL) SystemInformation(DL)	BCCH	Acquiring system information (see Section 3.2)
Paging	Paging(DL)	PCCH	Triggering mobile terminating calls, indicating system information updates, and PWS notifications (see Section 3.3)
RRC Connection Establishment	RRCConnectionRequest(UL)	CCCH (SRB0)	Entering RRC_CONNECTED (see Section 3.4)
	ⓈRRCConnectionSetup(DL) ⒻRRCConnectionReject(DL)		
	ⓈRRCConnectionSetup Complete(UL)	DCCH (SRB1)	
Initial Security Activation	SecurityModeCommand(DL)	DCCH (SRB1)	Activating integrity protection and ciphering in AS layer (see Section 3.5)
	ⒻSecurityModeFailure(UL) ⓈSecurityModeComplete(UL)	DCCH (SRB1)	
RRC Connection Reconfiguration	RRCConnectionReconfiguration(DL)	DCCH (SRB1)	Addition, modification, and removal of radio bearers, measurement configurations, and handover command (see Sections 3.6, 3.8, 3.9 and 3.11)
	ⓈRRCConnection ReconfigurationComplete(UL)		
Counter Check	CounterCheck(DL) CounterCheckResponse(UL)	DCCH (SRB1)	AS security verification
DL Information Transfer	DLInformationTransfer(DL)	DCCH (SRB1/SRB2)	Transferring NAS or non-3GPP dedicated information

(*continued*)

Table 3.2 (*Continued*)

Procedure	Messages	Logical Channel	Purpose
UL Information Transfer	ULInformationTransfer(UL)	DCCH (SRB1/SRB2)	Transferring NAS or non-3GPP dedicated information
UE Capability Transfer	UECapabilityEnquiry(DL) UECapabilityInformation(UL)	DCCH (SRB1)	Transferring UE radio access capability information (see Section 3.7)
Measurement Reporting	MeasurementReport(UL)	DCCH (SRB1)	Reporting measurement results (see Section 3.9)
Mobility from E-UTRA	MobilityFromEUTRA Command(DL)	DCCH (SRB1)	Mobility to a cell using another RAT, such as GERAN, UTRAN, and CDMA2000 (see Section 3.11)
Handover from E-UTRA Preparation Request (CDMA2000)	HandoverFromEUTRAPreparationRequest (DL)	DCCH (SRB1)	Preparation for mobility to CDMA2000 (see Section 3.11)
UL Handover Preparation Transfer (CDMA2000)	ULHandoverPreparation Transfer(UL)	DCCH (SRB1)	Transferring the CDMA2000 dedicated information upon reception of the *HandoverFromEUTRAPreparationRequest* (see Section 3.11)
RRC Connection Re-establishment	RRCConnection ReestablishmentRequest(UL)	CCCH (SRB0)	Recovering from failures on the radio interface by reestablishing an RRC connection (see Section 3.10)
	ⓈRRCConnection Reestablishment(DL) ⒻRRCConnection ReestablishmentReject(DL)		
	ⓈRRCConnection ReestablishmentComplete(UL)	DCCH (SRB1)	
RRC Connection Release	RRCConnectionRelease(DL)	DCCH (SRB1)	Leaving RRC_CONNECTED (see Section 3.12)

Table 3.3 Summary of system information blocks in LTE (except SIB9 through SIB13)

Type	Main Content	Transport Channel	Applicable RRC state
MIB	DL bandwidth, System frame number	BCH	RRC_IDLE and RRC_CONNECTED
SIB1	Cell access related information, System information value tag, Scheduling of other system information	DL-SCH	RRC_IDLE and RRC_CONNECTED
SIB2	Common radio resource configuration, AC barring information, UL carrier frequency and bandwidth, MBSFN subframe configuration	DL-SCH	RRC_IDLE and RRC_CONNECTED
SIB3	Cell reselection information common for intra-frequency, inter-frequency, and inter-RAT cell reselection, Intra-frequency cell reselection information other than neighboring cell information	DL-SCH	RRC_IDLE only
SIB4	Neighboring cell information for intra-frequency cell reselection	DL-SCH	RRC_IDLE only
SIB5	Other E-UTRA frequencies, Neighboring cell information for inter-frequency cell reselection	DL-SCH	RRC_IDLE only
SIB6	UTRA frequencies for FDD and TDD, Inter-RAT cell reselection information for UTRA	DL-SCH	RRC_IDLE for UEs supporting UTRAN
SIB7	GERAN frequencies, Inter-RAT cell reselection information for GERAN	DL-SCH	RRC_IDLE for UEs supporting GERAN
SIB8	CDMA2000 frequencies, Inter-RAT cell reselection information for CDMA2000	DL-SCH	RRC_IDLE and RRC_CONNECTED for UEs supporting CDMA2000

The MIB is mapped onto the MIB message. The MIB message is the only RRC message carried on the Broadcast Channel (BCH), for which only specified configuration and fixed scheduling are provided.

All SIBs except SIB1 are mapped onto System Information (SI) messages. The SI messages are carried on the DL-SCH, for which channel reconfiguration and flexible scheduling can be provided.

Mapping of SIBs other than SIB1 to SI messages is flexible and indicated by SIB1. A UE should acquire SIB1 to know how the other SIBs are scheduled on the concerned cell. Each SIB is contained in a single SI message, and a single SI message can contain multiple SIBs with the same periodicity. There may be multiple SI messages transmitted with the same periodicity.

Time domain scheduling of the MIB and SIB1 is fixed. The MIB is scheduled with a periodicity of 40 ms and is repeated every 10 ms within 40 ms. SIB1 is scheduled with a periodicity of 80 ms and is repeated every 20 ms within 80 ms.

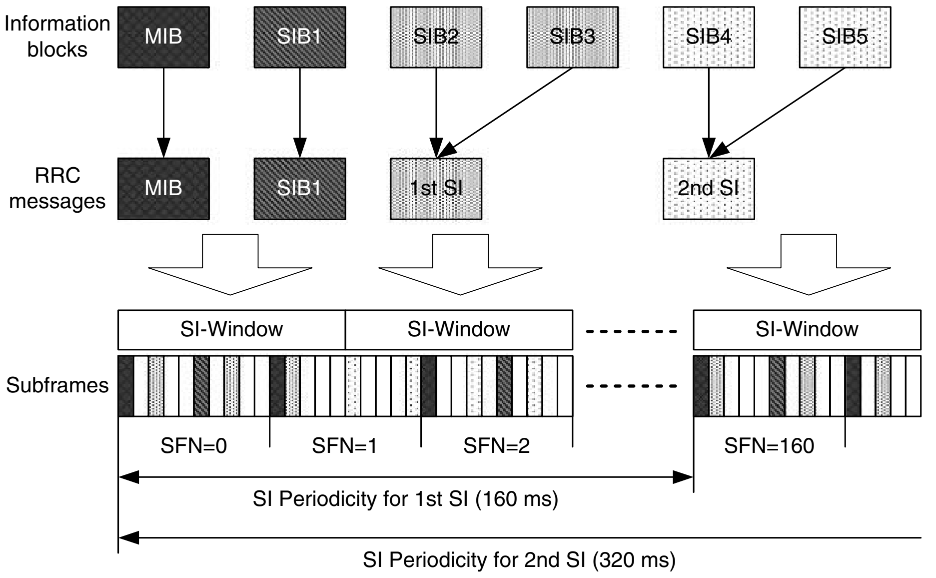

Figure 3.1 An example of system information scheduling

Scheduling of the SI messages is dynamic. The SI messages are transmitted within periodical time windows called "SI-windows". Each SI message is associated with an SI-window, and different SI messages are broadcast on different SI-windows that do not overlap and occur consecutively. Within an SI-window, only the associated SI message can be transmitted multiple times.

When SIB1 and the SI message are broadcast on the DL-SCH, the PDCCH associated with the DL-SCH uses a single System Information-Radio Network Temporary Identity (SI-RNTI) to address SIB1 and all SI messages at a cell. The UE decodes the SI-RNTI on the PDCCH within an SI-window to acquire detailed scheduling of system information corresponding to the SI-window.

3.2.2 System Information Update

The system information can be updated only at periodically occurring radio frames. The periodicity of system information update is called the "modification period", which is configured by SIB2.

If an update of the system information is required, an eNB needs to notify UEs about the update before sending the new system information. Figure 3.2 illustrates the system information update mechanism.

When the new system information is transmitted at the $(n+1)$th modification period, the notification is given to UEs at the nth modification period via *Paging* messages including *systemInfoModification*; that is, the indication of system information modification.

If the UE receives a *Paging* message including *systemInfoModification*, it knows that the system information will change at the next modification period boundary. UEs in both RRC_IDLE and RRC_CONNECTED are required to monitor *Paging* messages to obtain notification of system information modifications.

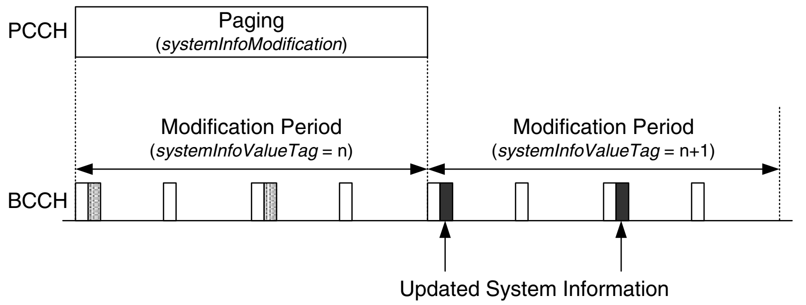

Figure 3.2 Modification of system information. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

The *Paging* message indicates only a change in system information; that is, detailed information about which system information block will be changed is not indicated. Thus, upon receiving the *Paging* message including *systemInfoModification*, the UE should read all the required system information transmitted in the next modification period. Note that the required system information is different depending on whether the UE is in RRC_IDLE or RRC_CONNECTED. After reading the required system information, the UE applies the new system information as a whole. Until the UE acquires the new system information, the UE should apply the previously acquired system information.

In addition to the system information update mechanism described previously, the value tag – that is, *systemInfoValueTag* broadcast in SIB1 – can be used to notify UEs about a change in system information. The eNB informs UEs of system information modification by incrementing the value tag if it transmits updated system information. If a UE returns from an outage of service due to, for example, coverage loss, it can verify whether the previously stored SI messages are still valid by comparing the newly obtained value tag with the one previously stored in the UE. Note that the value tag is applicable to changes of SIB provided in SI messages, meaning that a change of MIB and/or SIB1 is not relevant to the value tag.

3.3 Paging

The Paging procedure is used to inform a UE in RRC_IDLE about an incoming call. When the MME requests the eNB to page a UE, the RRC of the eNB initiates the Paging procedure. As described in Section 3.2, the RRC of the eNB also uses the Paging procedure to inform UEs in RRC_IDLE and RRC_CONNECTED about modifications to system information.

When the UE becomes attached to the EPC for services in the EPS, the UE enters EMM-REGISTERED. In this state, the UE is registered in the MME and the location of the UE is known to the MME such that the MME can page the UE.

While the UE is in EMM-REGISTERED, if the UE is in ECM-IDLE, the MME knows the location of the UE with the accuracy of a list of tracking areas, as described in Section 2.5. Hence, the MME sends a *Paging* message to all eNBs belonging to the tracking area in which the UE is registered, by using an S1AP message, as shown in Figure 3.3. On the other hand, if the UE is in ECM-CONNECTED, the MME does not utilize the Paging procedure

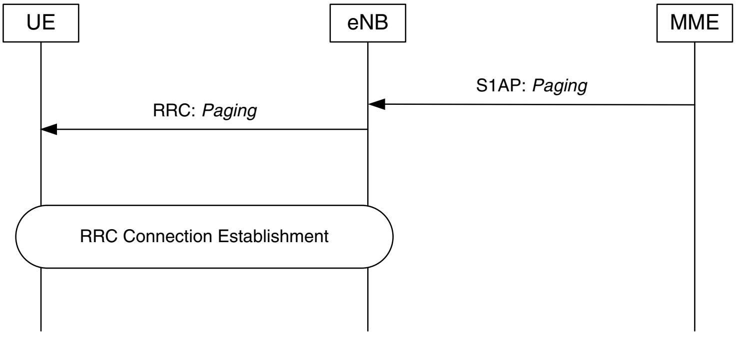

Figure 3.3 Paging for RRC Connection Establishment. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

specified in the S1AP protocol and the RRC layer, because the MME knows which eNB is serving the UE.

Figure 3.3 shows the Paging procedure triggering establishment of the RRC connection. When the Serving GW (S-GW) receives downlink data packets for the UE, if the UE is EMM-REGISTERED and ECM-IDLE, the MME sends the *Paging* message over the S1 interface to all eNBs belonging to the tracking area in which the UE is registered. Upon receipt of the *Paging* message from the MME, the RRC in the eNBs transmits the *Paging* message over the Uu interface.

The UE in RRC_IDLE is required to monitor a *Paging* message on the paging occasions assigned to the UE. The *Paging* message contains one or more UE identities to be paged, which can be either SAE-Temporary Mobile Subscriber Identity (S-TMSI) or International Mobile Subscriber Identity (IMSI). Upon receipt of the *Paging* message addressed to the UE, the NAS layer of the UE attempts to send a Service Request message to the MME. The service request from the NAS layer triggers the AS layer to establish an RRC connection, and then the RRC of the UE initiates the RRC Connection Establishment procedure.

The eNBs do not use the Paging procedure to page a UE in RRC_CONNECTED. However, the eNB uses the Paging procedure to provide notification of system information modification. Thus, UEs in RRC_CONNECTED as well as UEs in RRC_IDLE are required to monitor *Paging* messages, in order to be aware of changes in system information. Note that when a UE in RRC_CONNECTED monitors a *Paging* message, the UE is allowed to monitor any paging occasions.

Additionally, the Paging procedure can be used to indicate warning notifications broadcast on the BCCH. The indication of a warning notification via the Paging procedure is applicable to UEs in RRC_IDLE and RRC_CONNECTED. How to monitor indication of warning notifications via a *Paging* message is explained in Section 10.4.

3.4 Connection Establishment

When the NAS in the UE requests establishment of an RRC connection, the RRC in the UE initiates the RRC Connection Establishment procedure. The NAS in the UE can request establishment of an RRC connection in order to establish the NAS signaling connection between the UE and the MME by sending an initial NAS message such as the Attach Request, Tracking Area Update Request, Detach Request, Service Request, or Extended Service Request.

Upon request from the NAS, the RRC in the UE initiates the RRC Connection Establishment procedure. The RRC Connection Establishment procedure consists of the following steps, as illustrated in Figure 3.4: (1) applying Access Class Barring, (2) transmission of the *RRCConnectionRequest* message, (3) receipt of the *RRCConnectionSetup* message, (4) transmission of the *RRCConnectionSetupComplete* message.

3.4.1 Step 1: Access Class Barring

In abnormal circumstances such as network overload, it may be desirable for the network to restrict access attempts from some UEs. The Access Class Barring (ACB) mechanism allows the network to control access attempts from UEs over the radio interface. ACB is performed on a cell basis with the ACB parameters broadcast in SIB2.

During the RRC Connection Establishment procedure, the UE applies ACB by considering the following:

- one or more Access Classes (ACs) to which the UE belongs;
- the AC barring information provided by system information;
- the establishment cause for which the UE is establishing the RRC connection.

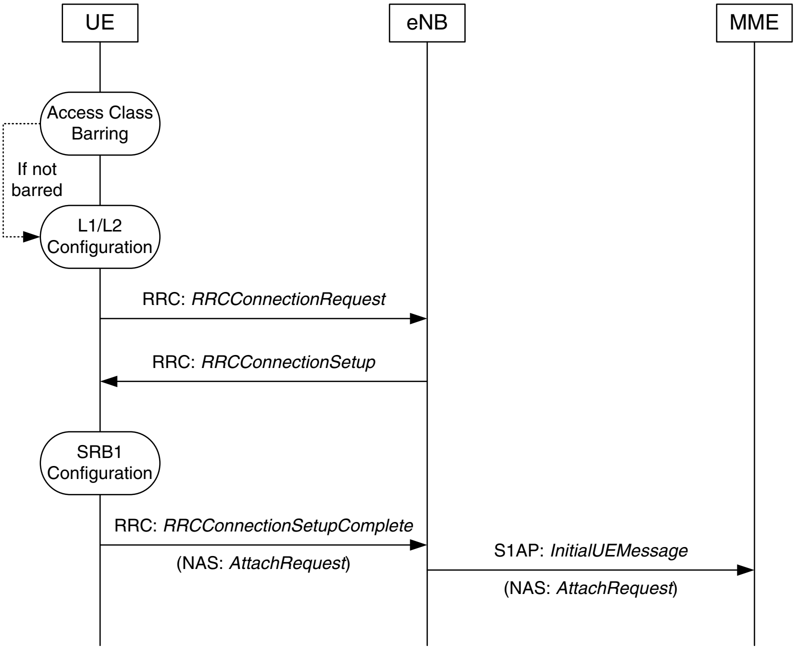

Figure 3.4 The RRC Connection Establishment procedure with the Attach Request message (successful case). Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

For any UE, one AC out of ten ACs (AC0, . . . , AC9) is assigned. Additionally, high-priority ACs (AC11, . . . , AC15) can be assigned to high-priority users such as PLMN staff and public utilities. AC10 is assigned to emergency access attempts.

A cell can control access attempts by broadcasting AC barring information; that is, *ac-BarringInfo*, via *SystemInformationBlockType2*. The AC barring information can contain different AC barring configurations for mobile originating signaling and mobile originating data – *ac-BarringForMO-Signaling* and *ac-BarringForMO-Data*. Each AC barring configuration consists of a barring factor (i.e., *ac-BarringFactor*), a barring time (i.e., *ac-BarringTime*), and a bit string for high-priority ACs (i.e., *ac-BarringForSpecialAC*).

For UEs without high-priority ACs, when the UE is establishing an RRC connection with a cell, it checks whether it is allowed for the access to the cell by referring to the AC barring information in *SystemInformationBlockType2*. If AC barring information is present, the UE draws a random number and checks whether the cell is barred or not by comparing the random number with the barring factor. If the random number is lower than the barring factor, the UE considers the cell not barred, otherwise the UE considers the cell barred. If the cell is not barred, the UE attempts to make an RRC connection with the cell. If the cell is barred, the UE cannot access the cell for the duration calculated with the barring time and the random number.

For UEs with high-priority ACs, the network can selectively allow special access for those UEs by using a bit string in the AC barring configuration. There is one-to-one mapping between the bits and the high-priority ACs, where each bit of the bit string indicates whether the corresponding AC of AC11, . . . , AC15 is barred. While the UE is establishing an RRC connection to the cell, if at least one bit of the bit string to which the UE belongs indicates "not barred", the UE considers the cell not barred, otherwise the UE considers the cell barred.

Access control for emergency calls is independent of that for other types of calls. The network indicates, in *SystemInformationBlockType2*, whether or not ACB is applied to an emergency call. For a UE making an emergency call, the UE is allowed to make an RRC connection unless it is indicated that ACB is applied to emergency calls.

Note that ACB is not applied to mobile terminating calls, as the network can control access attempts for mobile terminating calls by coordinating *Paging* messages. However, a mobile terminating call may be barred from a cell if the previous RRC connection request from a UE was rejected by the *RRCConnectionReject* message, including the wait time considered by the UE as the barring time for the cell.

3.4.2 Step 2: Transmission of the RRCConnectionRequest Message

If the cell is not barred according to step 1, the UE can transmit the *RRCConnectionRequest* message. Since the *RRCConnectionRequest* message is sent over SRB0, the RRC configures the lower layers with the specified CCCH configuration. After the configurations, the UE transmits the *RRCConnectionRequest* message including the UE identity and the establishment cause of the RRC connection.

The UE uses S-TMSI as the UE identity in the *RRCConnectionRequest* message if the UE NAS has given S-TMSI to the UE AS for this RRC connection establishment. If the S-TMSI has not been provided, the UE draws a random number in the range from 0 to $2^{40} - 1$ and then uses the random number as the UE identity.

When the NAS in the UE requests establishment of an RRC connection, the NAS indicates the establishment cause for the RRC connection. The establishment cause indicates the reason why the UE is establishing the RRC connection, for example, mobile originating data, mobile originating signaling, emergency call, or high-priority access.

In order to transmit the *RRCConnectionRequest* message, the MAC in the UE performs the Random Access procedure. The UE transmits the *RRCConnectionRequest* message during the Random Access procedure after successful receipt of the Random Access Response. How to perform the Random Access procedure is explained in Section 6.9.

It should be noted that the measurements and evaluation for cell reselection are continued until the UE receives the *RRCConnectionSetup* message. When the UE reselects a cell during the Random Access procedure, the RRC in the UE informs the NAS about the failure of the RRC Connection Establishment procedure.

3.4.3 Step 3: Receipt of the RRCConnectionSetup Message

When the eNB receives the *RRCConnectionRequest* message from the UE, the eNB decides whether or not to accept the request for establishment of an RRC connection. If the eNB decides to accept the request, the eNB sends the *RRCConnectionSetup* message to the UE. The *RRCConnectionSetup* message includes the dedicated radio resource configuration applied to the UE for SRB1 configuration.

When the UE receives the *RRCConnectionSetup* message, the UE applies the configuration included in the message and enters RRC_CONNECTED. Since the UE is in RRC_CONNECTED from this point in time, the UE stops the cell reselection process.

The eNB may reject the request by sending an *RRCConnectionReject* message to the UE, for example owing to network overload. When the UE receives an *RRCConnectionReject* message, the RRC in the UE informs the NAS about the failure of the RRC Connection Establishment procedure.

The *RRCConnectionReject* message includes the wait time. When the UE receives an *RRCConnectionReject* message, the UE starts a timer that is set to the wait time. While the timer is running, the UE cannot access the cell.

3.4.4 Step 4: Transmission of the RRCConnectionSetupComplete Message

In response to the *RRCConnectionSetup* message, the UE transmits the *RRCConnectionSetupComplete* message to the eNB. Before transmitting the *RRCConnectionSetupComplete* message, the UE needs to configure SRB1 according to the radio resource configuration in the *RRCConnectionSetup* message. The *RRCConnectionSetupComplete* message is sent over SRB1.

A cell may broadcast multiple PLMNs to support RAN sharing. When making an RRC connection with a cell broadcasting multiple PLMNs, the UE indicates, in the *RRCConnectionSetupComplete* message, the PLMN identity that was selected by the upper layers as a result of the PLMN selection process (see Section 2.4).

Furthermore, the *RRCConnectionSetupComplete* message includes an initial NAS message, which is delivered to an MME, and the routing information of the NAS message. If the UE has previously registered in the MME, the UE includes the Global Unique MME Identity (GUMMEI) corresponding to the registered MME. The eNB may use the GUMMEI to route the NAS message to the MME indicated by the GUMMEI.

The initial NAS message triggers the MME to initiate the Initial Context Setup procedure between the MME and the eNB to establish an S1 connection and a UE context in the eNB. Upon receiving the Initial NAS message from the UE via the eNB, the MME sends the Initial Context Setup Request message including UE-associated information to the eNB. The Initial Context Setup procedure over the S1 interface is linked to both initial activation of AS security and the establishment of DRBs and SRB2. The Initial Context Setup Request message can also contain the UE Radio Access Capability Information.

3.5 Security

The security function provides integrity protection and ciphering. The integrity protection prevents user data and signaling from being altered in an unauthorized manner, and the ciphering provides confidentiality of user data and signaling.

There are two levels of security between the UE and the network:

- **AS security:** This protects the RRC signaling and the user data between the UE and the E-UTRAN. It provides integrity protection and ciphering of the RRC signaling on the control plane of the radio protocols. It also provides ciphering of the user data on the user plane of the radio protocols. The Security Mode Command procedure in the RRC is used to activate AS security between the UE and the E-UTRAN.
- **NAS security:** This protects the NAS signaling between the UE and the MME. It provides integrity protection and ciphering of the NAS signaling. The Security Mode Command procedure in the NAS is used to activate NAS security between the UE and the MME.

The security function between the UE and the network is based on a secret key called K_{ASME}. The K_{ASME} is derived from the permanent key that is stored in both the USIM and the Home Subscriber Server (HSS). The MME receives the K_{ASME} from the HSS. The UE derives the K_{ASME} as a result of the Authentication and Key Agreement (AKA) procedure in the NAS protocols. During the AKA procedure, the UE and the network perform mutual authentication and agree on the K_{ASME}.

How to derive keys from the K_{ASME} for protection of signaling and user data is shown in Figure 3.5. The UE and the MME derive the K_{NASint} for integrity protection of the NAS messages and the K_{NASenc} for ciphering of the NAS messages using the K_{ASME}. The UE and the eNB derive the K_{eNB} from the K_{ASME} for protection of signaling and user data on the Uu interface between the UE and the eNB. From the K_{eNB}, the UE and the eNB derive the K_{RRCint} for integrity protection of the RRC messages, the K_{RRCenc} for ciphering of the RRC messages, and the K_{UPenc} for ciphering of user data.

The four AS keys (K_{eNB}, K_{RRCint}, K_{RRCenc}, and K_{UPenc}) change upon every handover and connection re-establishment. For handover from a source eNB to a target eNB, the UE and the source eNB derive the $K_{eNB}*$, which is a new K_{eNB} used at the target cell. The other AS keys – K_{RRCint}, K_{RRCenc}, and K_{UPenc} – are derived from the $K_{eNB}*$. The $K_{eNB}*$ is derived based either on the current K_{eNB} or a fresh Next Hop (NH) in the UE and the eNB. The physical cell ID and downlink carrier frequency of the target cell are also used for derivation of the $K_{eNB}*$. The NH is derived in the UE and the MME from the K_{ASME}. The eNB receives the NH from the MME to derive the $K_{eNB}*$ for handovers. An intra-cell handover procedure may be used to change the AS keys in RRC_CONNECTED (see Section 3.8).

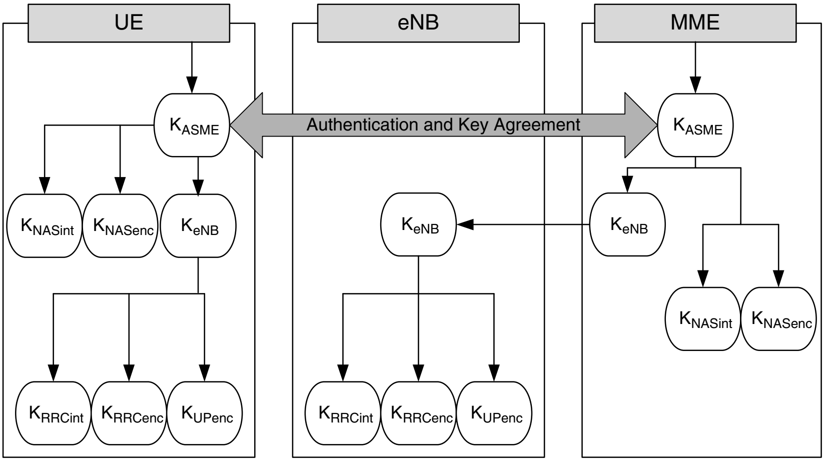

Figure 3.5 Key derivation for AS security and NAS security except handover. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

Activation of AS security is initiated by the eNB after the Initial Context Setup procedure over the S1 interface is initiated by the MME. During the Initial Context Setup procedure, the MME informs the eNB about the K_{eNB} derived directly from K_{ASME} for AS key derivation and the algorithms of integrity protection and ciphering supported in the UE. Thus, after receiving the K_{eNB} and the algorithms from the MME, the eNB can initiate activation of AS security with the UE.

For initial activation of AS security, the eNB sends the *SecurityModeCommand* message, indicating the integrity protection algorithm and the ciphering algorithm to the UE after the RRC Connection Establishment procedure is completed. The overall procedure of AS security activation is shown in Figure 3.6.

The RRC layer of the UE receives the *SecurityModeCommand* message before configuring the PDCP layer to apply integrity protection. Thus, upon receiving the *SecurityModeCommand* message, the RRC layer derives the K_{RRCint} for integrity protection of the RRC messages with the algorithm indicated by the received *SecurityModeCommand* message, and then requests the PDCP layer to verify the integrity of the received *SecurityModeCommand* message using the algorithm and the K_{RRCint}.

Note that the *SecurityModeCommand* message is transmitted with integrity protection even though AS security is not activated in the UE. This is because the UE can verify the integrity of the message using the algorithm included in the message. However, ciphering is not applied to the *SecurityModeCommand* message because the UE cannot decode the message until the ciphering algorithm included in the message is applied.

If the UE fails to verify the integrity of the *SecurityModeCommand* message, the UE sends a *SecurityModeFailure* message to the eNB in response to the *SecurityModeCommand* message. In this case, neither integrity protection nor ciphering is applied to the *SecurityModeFailure* message because valid algorithms are not available in the UE.

If the received *SecurityModeCommand* message passes the integrity verification in the PDCP layer, the RRC layer of the UE further derives the K_{RRCenc} and the K_{UPenc}. The

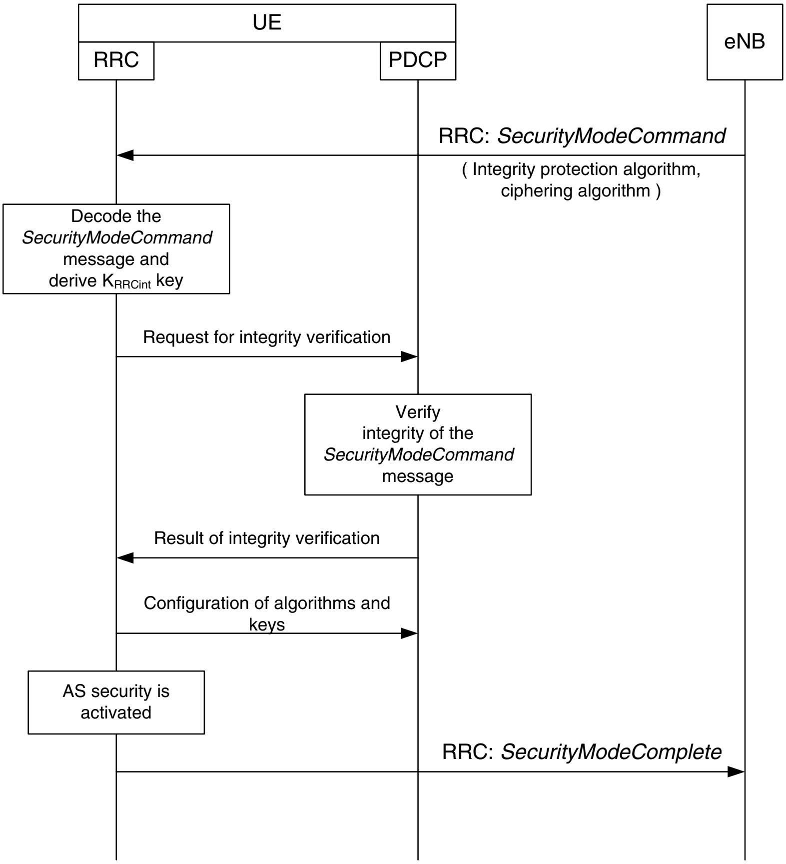

Figure 3.6 Initial activation of AS security. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

RRC layer configures the PDCP layer to apply integrity protection using the integrity protection algorithm and the K_{RRCint}, and to apply ciphering using the ciphering algorithm, the K_{RRCenc}, and the K_{UPenc}. Then, the UE considers AS security to be activated and applies both integrity protection and ciphering to all subsequent RRC messages received and sent by the UE.

Upon the activation of AS security, the UE sends a *SecurityModeComplete* message to the eNB in response to the *SecurityModeCommand* message. The *SecurityModeComplete* message is integrity protected but not ciphered. The reason for not applying ciphering to the *SecurityModeComplete* message is that by transmitting both response messages – that is, *SecurityModeFailure* and *SecurityModeComplete* messages – unciphered, the eNB can easily decode the response message without deciphering, regardless of whether security activation is successful or not in the UE.

After AS security has been activated, the eNB establishes SRB2 and DRBs. The eNB does not establish SRB2 and DRBs prior to activating AS security. Once AS security is activated, all RRC messages over SRB1 and SRB2 are integrity protected and ciphered, and all user data over DRBs are ciphered by the PDCP layer. However, neither integrity protection nor ciphering applies for SRB0.

The integrity protection algorithm is common for signaling radio bearers SRB1 and SRB2, and the ciphering algorithm is common for all radio bearers (i.e., SRB1, SRB2, and DRBs). The integrity protection and ciphering algorithms can be changed only upon handover.

3.6 RRC Connection Reconfiguration

The eNB may trigger the RRC Connection Reconfiguration procedure for the following purposes:

- establishment and reconfiguration of SRB2 and one or more DRBs;
- handover;
- configuration of measurements.

The RRC Connection Reconfiguration procedure is used for the establishment of SRB2 and DRB(s). The procedure is also used to modify or release the established radio bearers. Note that the establishment of SRB2 and DRB(s) is possible only after AS security has been activated. After the UE enters RRC_CONNECTED, at least one DRB should be configured in the UE for exchange of user traffic.

The RRC Connection Reconfiguration procedure is also used to perform handover. In this case, the Mobility Control Information (MCI) – that is, *mobilityControlInfo* – is included in the *RRCConnectionReconfiguration* message. In addition, this procedure can be used to configure measurements of intra-LTE/inter-RAT.

It should be noted that in cases of reconfiguration failure (e.g., when the UE cannot comply with any of the configuration in the *RRCConnectionReconfiguration* message), if security has been activated, the UE performs the RRC Connection Re-establishment procedure. If security has not been activated, the UE goes to RRC_IDLE.

3.6.1 SRB2 Establishment

For SRB2 establishment, the eNB includes the necessary parameters to configure SRB2 in the *RRCConnectionReconfiguration* message. After SRB2 has been established, NAS information is always carried over SRB2. In cases where SRB2 has not yet been established, for example, during the first RRC Connection Reconfiguration procedure after the UE enters RRC_CONNECTED, NAS information is transferred over SRB1. Note that if the RRC layer of the UE receives the NAS information, it does not touch the NAS information, but just forwards it to the NAS layer.

3.6.2 DRB Establishment

To provide a user plane service, the eNB needs to establish at least one or more DRBs for the UE. For the establishment of DRBs, the eNB provides dedicated radio resource configuration in the *RRCConnectionReconfiguration* message. The scope of radio resource configuration includes the configuration for PDCP, RLC, MAC, and PHY layers. The maximum number of DRBs that can be set up for one UE is eight.

The parameters included in the *RRCConnectionReconfiguration* message are determined by the eNB, based on information in the Initial Context Setup Request message received from the MME. The Initial Context Setup Request message includes the list of E-RABs for which the eNB is requested to configure DRBs to the UE. The eNB may not configure DRBs for all E-RABs requested by the MME, due to, for example, insufficient radio resources. Thus, the eNB informs the MME of the list of E-RABs that have been established successfully and the list of E-RABs that failed to be established by the Initial Context Setup Response message sent over the S1 interface.

3.6.3 Parallel Procedure with Security Activation

The eNB is allowed to perform the Initial Security Activation procedure and the RRC Connection Reconfiguration procedure in parallel to accelerate the completion of the procedures. In other words, the eNB is allowed to send the *RRCConnectionReconfiguration* message prior to receipt of the *SecurityModeComplete* message, as shown in Figure 3.7. Since the eNB applies integrity protection and ciphering to the *RRCConnectionReconfiguration* message, the PDCP layer of the UE can decrypt the *RRCConnectionReconfiguration* message only after AS security has been activated in the UE by the concurrent *SecurityModeCommand* message.

If the parallel procedure is triggered by a single S1AP procedure, both the Security Activation and RRC Connection Reconfiguration procedures should be completed successfully. If either of the procedures fails, the eNB considers that the S1AP procedure has failed, and the eNB releases the RRC connection by sending an *RRCConnectionRelease* message to the UE.

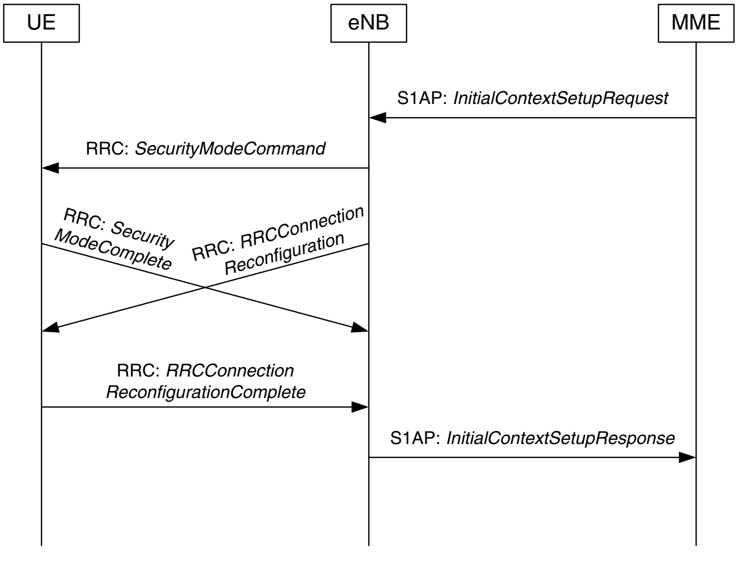

Figure 3.7 Initial AS security activation and RRC connection reconfiguration in parallel. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

3.7 UE Capability Transfer

For a UE in RRC_CONNECTED, the eNB should be precisely aware of UE capabilities in order to configure the UE properly for services. Normally, the MME stores the UE capabilities, consisting of UE Radio Access Capability information and the UE Core Network Capability information. The UE Core Network Capability is indicated by the UE via NAS signaling during, for example, the Attach procedure, and the UE Radio Access Capability is transferred from the UE to the eNB via the UE Capability Transfer procedure of the RRC, and then delivered to the MME over the S1 interface. The MME then provides the eNB with the UE Radio Access Capability, if available, whenever the UE enters RRC_CONNECTED.

The MME may trigger the eNB to request retrieval of the UE Radio Access Capability when the UE Radio Capability is not available or new retrieval is requested. Upon receiving a request for retrieval of the UE Radio Access Capability, the eNB triggers the UE Capability Transfer procedure, as shown in Figure 3.8.

The eNB requests the UE Radio Access Capability by sending a *UECapabilityEnquiry* message to the UE. In this message, the eNB indicates *rat-Type*; that is, one or multiple RATs for which the UE Radio Access Capability is requested from the UE. The *rat-Type* could be set to E-UTRA, UTRA, GERAN CS, GERAN PS, or CDMA2000. The eNB utilizes the UE Radio Access Capability of other RATs for inter-RAT mobility. Upon receipt of the *UECapabilityEnquiry* message, the UE responds with a *UECapabilityInformation* message including its UE Radio Access Capability for the RAT indicated by the *rat-Type*.

Generally, it is assumed that the UE Radio Access Capability information does not change dynamically while the UE is attached to the EPC. If the UE Radio Access Capability for E-UTRA is changed, the UE can perform detach and re-attach to update the information for the network. To update the UE Radio Access Capability for a RAT other than E-UTRAN, the UE Capability Transfer procedure can be triggered.

3.8 Intra-EUTRA Handover

While a UE is in RRC_IDLE, UE-based mobility is used, where the UE in RRC_IDLE decides when and where the UE moves by performing cell selection and reselection processes. Since the eNB has no UE context for a UE in RRC_IDLE, the eNB cannot be

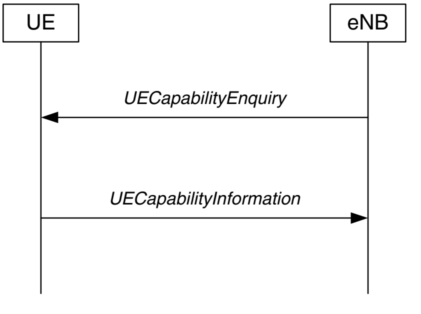

Figure 3.8 UE Capability Transfer. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

involved directly in UE mobility in RRC_IDLE. Even though the network provides information on cell selection and reselection, for example, via system information and the *RRCConnectionRelease* message, UE mobility in RRC_IDLE is processed internally on the UE side.

On the other hand, while a UE is in RRC_CONNECTED, network-controlled mobility is used, where the eNB decides when and where the UE in RRC_CONNECTED moves. The UE can assist the network-controlled mobility if the eNB configures the UE to send a measurement report when mobility is desirable. It is also possible for the eNB to initiate mobility blindly without measurement reports from the UE. Note that for the RRC Connection Re-establishment procedure, UE-based mobility is used even in RRC_CONNECTED where cell selection is performed.

Note that in E-UTRAN, the UE is always connected to a single cell only. For this reason, the E-UTRAN does not support soft handover, which is supported in UTRAN. During intra-EUTRA handover, the UE switches from its serving cell, the so-called *source cell*, to a neighboring cell, the so-called *target cell*.

Intra-EUTRA handover (i.e., both source cell and target cell are LTE cells) can be further classified into X2-based handover and S1-based handover. X2-based handover involves X2AP messages that are communicated over the X2 interface established between the source eNB and the target eNB. X2-based handover can be used when the MME serving the UE is not changed as a result of handover. It is possible for the Serving GW (S-GW) serving the UE to be changed as a result of X2-based handover. S1-based handover is used when X2-based handover cannot be used. For example, when the MME serving the UE is changed as a result of handover, S1-based handover is used. S1-based handover can also change the S-GW as well as the MME. There is no difference in the signaling over the radio interface between X2-based handover and S1-based handover; the same RRC procedure is commonly applied to both types of handover, and the UE behaviors are also the same. Only network signaling among the source eNB, the target eNB, and possibly the CN node is different for X2-based handover and S1-based handover.

In general, the handover procedure consists of the following three phases:

- handover preparation;
- handover execution;
- handover completion.

3.8.1 Handover Preparation

In E-UTRAN, the handover is successful only if the UE accesses the cell prepared for the handover. Thus, the source eNB requests the target eNB to prepare for the handover before the UE executes handover to the target cell. The source eNB is allowed to prepare handover with multiple target eNBs.

Before initiating handover preparation, the source eNB decides a target cell to which the UE will move and also decides when to move the UE to the selected target cell. As an input to the handover decision that the source eNB needs to make, the measurement results reported by the UE can be used. As another input to the handover decision, the MME may provide a handover restriction list for the UE to the eNB to limit the number of target cell candidates.

When making a handover decision, the source eNB also determines whether X2-based handover or S1-based handover will be used. When the source eNB decides to use X2-based

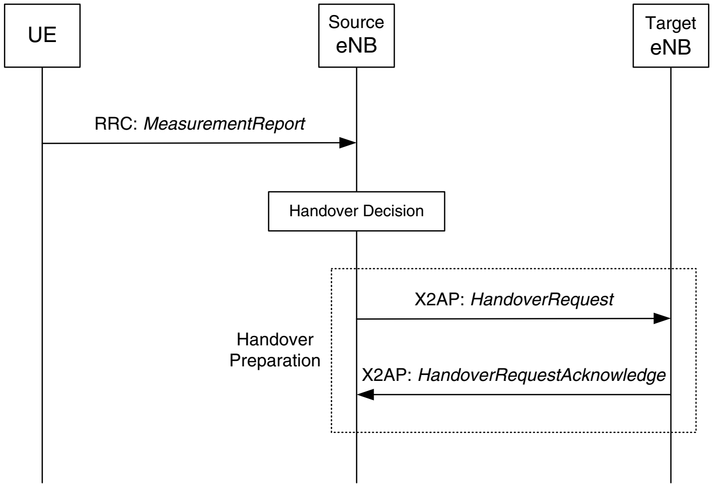

Figure 3.9 Measurement report and handover preparation for X2-based handover. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

handover, the source eNB initiates the X2-based handover by sending a Handover Request message over the X2 interface to the target eNB for handover preparation, as shown in Figure 3.9. When the source eNB decides to use S1-based handover, the source eNB initiates the S1-based handover by sending a Handover Required message over the S1 interface to the MME. The Handover Required message triggers the MME to perform handover preparation with the target eNB by sending a Handover Request message to the target eNB, as shown in Figure 3.10.

During handover preparation, the source eNB transfers inter-node RRC information for the UE to the target eNB. The inter-node RRC information includes the UE Radio Access Capability information for the concerned UE, RRC configurations which are currently being used at the source cell, and UE-specific RRM information. The target cell utilizes the received information to configure the UE during/after handover. However, some of the inter-node RRC information may not be utilized by the target cell.

The source eNB can also include re-establishment information in the inter-node RRC information sent to the target eNB. The re-establishment information is needed for the RRC Connection Re-establishment procedure, which might happen in cases of failure of the handover procedure.

The Handover Request message includes information on QoS and priority for E-RABs to be set up at the target cell. Upon receiving the Handover Request message, the target eNB may perform admission control, based on the information in the Handover Request message. Some E-RABs may fail to be admitted at the target cell due to, for example, lack of radio resource or unsupported configurations. If at least one of the requested E-RABs is admitted at the target cell, the target eNB reserves the necessary resources and sends a Handover Request Acknowledge message to the source eNB. It should be noted that for each admitted E-RAB, data forwarding can be supported during the handover execution phase. For data

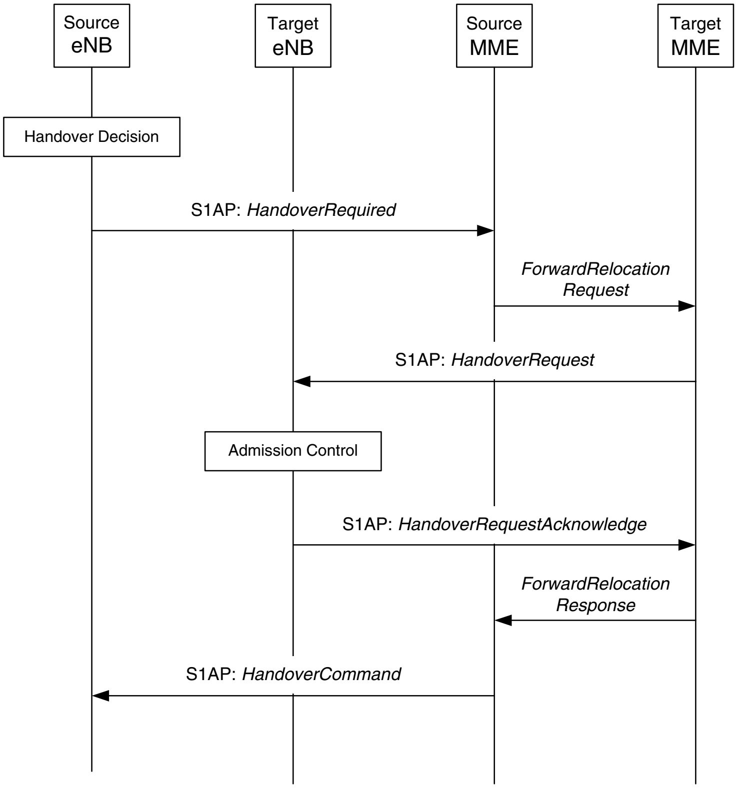

Figure 3.10 Handover preparation for S1-based handover with MME relocation. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

forwarding in handover execution, the source eNB and the target eNB should exchange information about GTP tunnels where data forwarding is applied.

The Handover Request Acknowledge message sent by the target eNB includes a transparent container carrying the *RRCConnectionReconfiguration* message. The *RRCConnectionReconfiguration* message, which is generated by the RRC layer of the target cell, contains configurations that the UE should use at the target cell. The target eNB can use “delta signaling” for the *RRCConnectionReconfiguration* message, which means that only changes to configurations, compared to configurations used at the source eNB, are signaled in the *RRCConnectionReconfiguration* message.

For X2-based handover, the target eNB sends the *RRCConnectionReconfiguration* message directly to the source eNB via the Handover Request Acknowledge message over the X2 interface, as shown in Figure 3.9. For S1-based handover, the target eNB sends the Handover Request Acknowledge message to the MME. Thus, the source eNB receives the *RRCConnectionReconfiguration* message from a transparent container in the Handover Command message that the MME sends to the source eNB over the S1 interface, as shown in Figure 3.10.

For AS security between the UE and the target eNB, the source eNB derives the $K_{eNB}*$ and then sends the $K_{eNB}*$ to the target eNB during handover preparation. The target eNB uses the $K_{eNB}*$ to secure communication on the radio interface with the UE. The target eNB also selects integrity protection and ciphering algorithms, and then indicates the selected algorithms in the *RRCConnectionReconfiguration* message sent to the UE.

3.8.2 Handover Execution

The sequential step of handover execution is shown in Figure 3.11. Upon receiving the transparent container carrying the *RRCConnectionReconfiguration* message from the target eNB, the source eNB executes the handover by forwarding the *RRCConnectionReconfiguration* message to the UE without modifying the contents of the message. The *RRCConnectionReconfiguration* message used for handover includes the *mobilityControlInfo*; that is, the "*RRCConnectionReconfiguration* message including *mobilityControlInfo*" corresponds to the "handover command".

Upon receiving the handover command, the UE resets the MAC and re-establishes the PDCP and the RLC for all RBs that are established. Then, the UE starts synchronizing with the downlink of the target cell.

The *mobilityControlInfo* in the *RRCConnectionReconfiguration* message informs the UE about the target cell by including the physical cell ID, the carrier frequency, and the UL/DL bandwidths of the target cell. In addition, the *mobilityControlInfo* provides a new C-RNTI

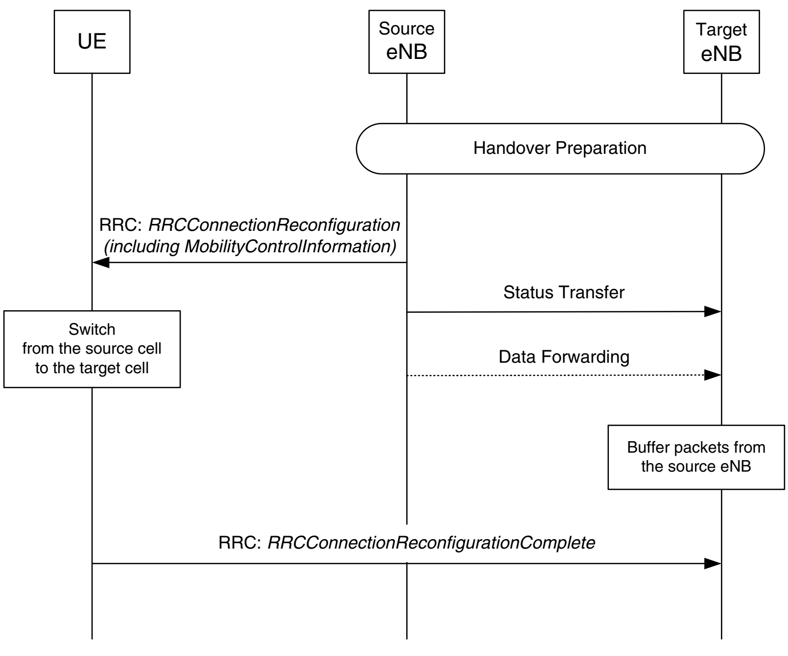

Figure 3.11 Handover execution. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

that is used to identify the UE at the target cell and also gives configurations of common channels, such as RACH, that are used at the target cell. The UE applies the new C-RNTI for the target cell and then uses the RACH configuration to perform the random access procedure at the target cell for transmission of the *RRCConnectionReconfigurationComplete* message. The RACH configuration can include a dedicated preamble, by which the UE performs non-contention-based random access at the target cell (see Section 6.9).

For AS security, the handover command also includes the security configuration used at the target cell. Upon receiving the security configuration, the UE derives the new K_{eNB} key (i.e., $K_{eNB}*$) to be used at the target cell. For intra-EUTRA handovers, except for intra-cell handovers where the source cell is equal to the target cell, the UE derives the $K_{eNB}*$ based on the current K_{eNB} or the NH, according to the security configuration. An intra-cell handover may be used to change AS keys in RRC_CONNECTED.

If the target eNB selects different algorithms compared to the source eNB, it indicates the selected algorithms in the security configuration of the handover command. The UE then uses the selected algorithms for derivation of AS keys. If the UE does not receive any selected integrity and ciphering algorithms, it continues to use the algorithms used in the source eNB.

In the meantime, the source eNB can send a Status Transfer message to the target eNB. The Status Transfer message includes the uplink PDCP receiver status and the downlink PDCP transmitter status to support SDU reordering in the PDCP layer (see Section 4.6).

To avoid data loss, the source eNB may also perform data forwarding of user packets from the source eNB to the target eNB, via GTP tunnels established during handover preparation. For X2-based handover, data forwarding is done over a direct forwarding path between the source eNB and the target eNB. For S1-based handover, an indirect forwarding path passing through the S-GW is used. The source eNB should continue forwarding of packets to the target eNB as long as the source eNB receives packets from the EPC or as long as the buffer of the source eNB has not been emptied.

When the UE can comply with the configuration in the handover command, the UE synchronizes with the downlink of the target cell and then attempts to access the target cell with the Random Access procedure. Upon receiving the handover command, the UE starts a timer that governs the maximum duration of handover execution. This timer is stopped when random access is completed successfully in the MAC layer. If random access is not completed until the timer has expired, the UE declares a handover failure. Then, the UE performs the RRC Connection Re-establishment procedure to recover from the handover failure. During the re-establishment procedure, the UE may access the source cell or another cell. In any case, re-establishment can be successful only with the cell of the prepared eNB.

While the timer is running, the UE performs the Random Access procedure to the target cell. The UE transmits a random access preamble at the first available RACH occasion following receipt of the handover command. Since the UE's preamble transmission timing is unknown to the target cell, the target cell should prepare to receive a random access preamble from the UE for some time after providing the handover command to the source cell.

To reduce handover latency, the UE does not read the system information from the target cell before performing the Random Access procedure, but acquires the system information after the completion of the handover. Until the acquisition of required system information, the UE may be temporarily unable to apply some parts of configurations right after the handover.

3.8.3 Handover Completion

From the UE side, the handover procedure is completed if the UE transmits the *RRCConnectionReconfigurationComplete* message successfully to the target cell. From the network side, further processes are performed, which include switching of the packet data path towards the target eNB and releasing UE-associated resources in the source eNB. Depending on the type of handover – X2-based or S1-based – different procedures are applied, as shown in Figures 3.12 and 3.13, respectively.

When the UE has been identified at the target cell, the target eNB can indicate to the MME that handover is complete on the radio interface, by sending a Path Switch Request message to the MME for X2-based handover or by sending a Handover Notify message to the MME for S1-based handover. When the MME receives either a Path Switch Request message or a Handover Notify message from the target eNB, depending on the type of handover, the MME communicates with the S-GW to switch the data path from the source eNB to the target eNB.

Once the data path has been switched to the target eNB, the source eNB may receive a UE Context Release message from the target eNB for X2-based handover, or a UE Context Release Command message from the MME for S1-based handover. Upon receiving a UE Context Release message or a UE Context Release Command message, the source eNB releases resources associated with the UE, and then the handover procedure is complete on the network side.

When the handover preparation is finished, the source eNB starts a timer governing the maximum delay allowed for handover execution and handover completion. If the source eNB does not receive the UE Context Release message or the UE Context Release Command message until the timer has expired, the source eNB sends the MME a UE Context Release

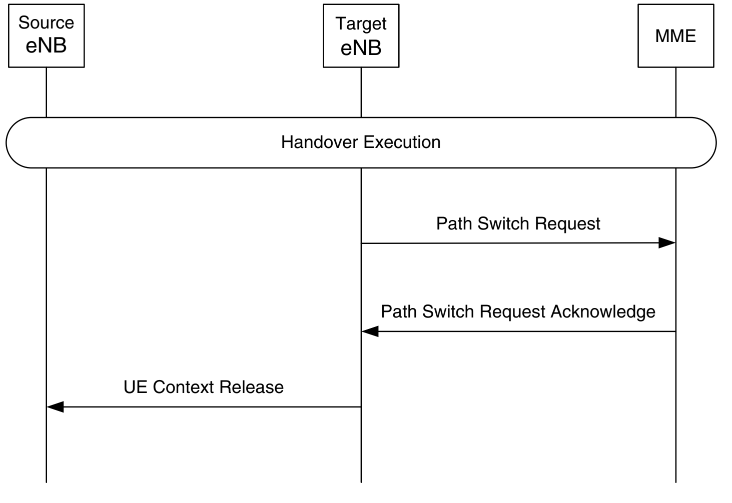

Figure 3.12 Handover completion for X2-based handover. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

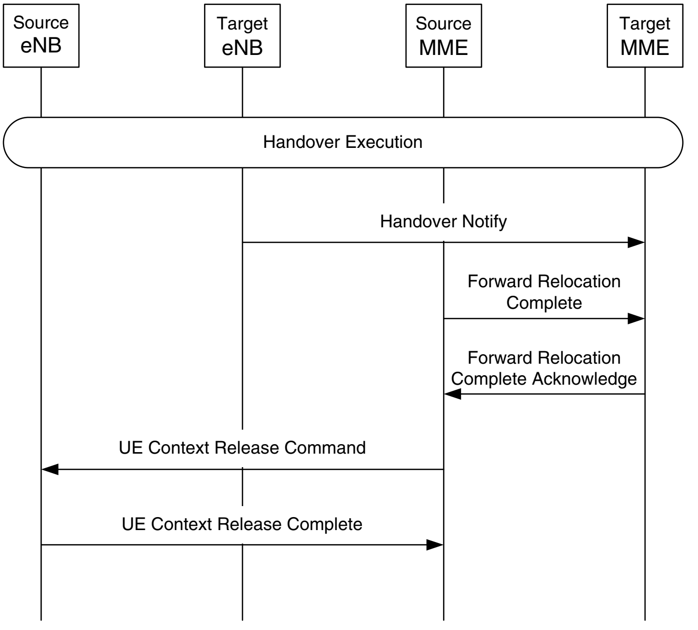

Figure 3.13 Handover completion for S1-based handover with MME relocation. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

Request message, which causes the MME to send the UE Context Release Command message to the source eNB.

3.9 Measurement Control

LTE supports the following types of measurement:

- intra-frequency measurements for the E-UTRA;
- inter-frequency measurements for the E-UTRA;
- inter-RAT measurements:
  - inter-RAT measurements of UTRA frequencies;
  - inter-RAT measurements of GERAN frequencies;
  - inter-RAT measurements of CDMA2000 1xRTT or HRPD frequencies.

Intra-frequency measurement is defined as measurement at the downlink carrier frequency of the serving cell. Inter-frequency measurement is defined as measurement at frequencies different from the downlink carrier frequency of the serving cell. Inter-RAT measurement is defined as measurement for RATs different from the E-UTRA.

While a UE is in RRC_CONNECTED, it can be configured with measurements applicable to the UE. The eNB provides the measurement configuration to a UE in RRC_CONNECTED by using the *RRCConnectionReconfiguration* message.

In the UE, the Physical layer performs measurements and reports measured data to the RRC layer. The Physical layer may perform L1 filtering according to performance requirements in the 3GPP specifications. When measured data are reported from the Physical layer, the RRC layer performs L3 filtering, which is configured by the measurement configuration.

Such filtered data, which are processed after L3 filtering, are used for evaluation of the reporting criteria. The reporting criteria that are configured by the measurement configuration determine whether or not reporting should be triggered. Upon triggering of reporting, the UE reports measurement results according to the measurement configuration.

3.9.1 Measurement Configuration

The measurement configuration in the *RRCConnectionReconfiguration* message includes the following parameters:

- measurement objects;
- reporting configurations;
- measurement identities;
- quantity configurations;
- measurement gaps.

Measurement objects are defined as the objects on which the UE performs measurements, such a set of frequencies or a set of cells. The E-UTRAN configures a single measurement object for a given frequency. The E-UTRAN can configure blacklisted cells that are not considered for event evaluation or measurement reporting.

With regard to inter-RAT measurements, the measurement object for the UTRA is a set of cells on a single carrier frequency, the measurement object for the GERAN is a set of carrier frequencies, and the measurement object for the CDMA2000 is a set of cells on a single carrier frequency.

The reporting configuration defines the reporting criterion and the reporting format. The reporting criterion corresponds to the criterion that triggers the UE to send a measurement report. The reporting format corresponds to the quantities to be included in a measurement report.

The quantities are the Reference Signal Received Power (RSRP) and the Reference Signal Received Quality (RSRQ) for the E-UTRA, and the maximum number of cells. The RSRP corresponds to the signal strength of an E-UTRA cell. The RSRQ comes from the ratio of the RSRP to the Received Signal Strength Indicator (RSSI). The RSSI is the received wideband power including interference, and therefore the RSRQ is used as an additional metric to take into account interference in LTE. For inter-RAT measurements, the quantities are the Received Signal Code Power (RSCP), which is comparable to the RSRP, and the Ec/No for the UTRA, and the Received Signal Strength Indicator (RSSI) for the GERAN.

A measurement identity is used to identify a linkage between one measurement object and one reporting configuration. It is possible for a single measurement object to be linked to more than one reporting configuration by configuring multiple measurement identities. It is also possible for a single reporting configuration to be linked to more than one measurement object by configuring multiple measurement identities. When the *MeasurementReport* message is sent to the eNB, the measurement identity is included in the message as a reference to a combination of the concerned measurement object and the concerned reporting configuration.

The quantity configuration configures the measurement quantity and filtering associated with the measurement quantity. This configuration is used for evaluation of the reporting criteria and measurement reporting.

Measurement gaps are defined as the periods in which no transmissions are scheduled in uplink or downlink towards the UE. The measurement gap can be configured to assist the inter-frequency or inter-RAT measurements. During the measurement gap, the UE is allowed to leave its serving cell to perform measurements of other frequencies. The UE indicates to the eNB whether or not measurement gaps are required to perform inter-frequency measurements and inter-RAT measurements separately, as part of the UE Radio Access Capability information. The eNB provides one single configuration of the measurement gaps in the UE, which is common to all types of measurements including inter-frequency E-UTRA measurements and inter-RAT measurements on the GERAN, the UTRA, and the CDMA2000.

3.9.2 Measurement Report Triggering

The UE may measure and report the following types of cells:

- the serving cell;
- listed cells;
- detected cells.

Listed cells are defined as cells listed in the measurement objects. Detected cells are defined as cells that are not listed in the measurement objects but are detected by the UE on the carrier frequencies indicated by the measurement objects.

The UE performs measurements and reporting for the serving cell, listed cells, and detected cells in the E-UTRA, for listed cells in the UTRA and the CDMA2000, and for detected cells in the GERAN.

The reporting can be triggered either periodically (i.e., periodical reporting) or based on a configured event (i.e., event-triggered reporting), according to the reporting configuration of the measurement configuration. In addition, the UE can provide a number of periodic reports to the eNB after reporting has been triggered, based on a configured event. Such "event-triggered periodic reporting" is also configured by the reporting configuration with the number and the interval of measurement reports.

For event-triggered reporting, several events are defined in the 3GPP specification. The following events are defined as criteria for measurements within the E-UTRA:

- Event A1: Serving becomes better than threshold;
- Event A2: Serving becomes worse than threshold;
- Event A3: Neighbor becomes offset better than serving;
- Event A4: Neighbor becomes better than threshold;
- Event A5: Serving becomes worse than one threshold and neighbor becomes better than another threshold.

In addition, the following events are defined as criteria for inter-RAT measurements on the GERAN, the UTRA, and the CDMA2000:

- Event B1: Inter-RAT neighbor becomes better than threshold;
- Event B2: Serving becomes worse than one threshold and inter-RAT neighbor becomes better than another threshold.

When one or more cells satisfy the condition triggering the event, the UE triggers the event to report one or more cells with measurement results in the *MeasurementReport* message. The condition is configured via the measurement configuration with several parameters such as one or more thresholds, a hysteresis, and an offset.

The eNB also configures the time-to-trigger parameter in the reporting configuration for event-triggered reporting. The UE triggers the concerned event only when the condition is satisfied during a time interval corresponding to the time-to-trigger parameter. According to the measurement configuration, the UE may scale the time to trigger an event, depending on how fast the UE is moving. For instance, when the UE is moving quickly, the UE may shorten the time to trigger. When the UE is moving slowly, the UE may lengthen the time to trigger.

3.9.3 Measurement Reporting

The UE uses the measurement reporting procedure to report measurement results to the eNB. The UE includes the measurement results in the *MeasurementReport* message. The eNB can use the measurement results in handover decisions.

3.10 RRC Connection Re-establishment

The RRC Connection Re-establishment procedure is used to recover a connection between the UE and the eNB from various failures on the radio interface. The overall procedure is shown in Figure 3.14. The UE initiates the RRC Connection Re-establishment procedure when one of the following failure conditions is met:

- when a Radio Link Failure (RLF) is detected;
- when handover failure occurs;
- when the Mobility From E-UTRA procedure fails;
- when an integrity check failure is indicated from the PDCP layer (see Section 4.3);
- when an RRC connection reconfiguration fails.

The UE considers an RLF to have been detected when one of the following cases occurs:

- when out-of-synchronization occurs in the Physical layer;
- when the random access problem occurs in the MAC layer (see Section 6.9);
- when the maximum number of retransmissions is reached in the RLC layer (see Section 5.4.3).

The UE can initiate the RRC Connection Re-establishment procedure while AS security is activated. AS security is re-activated during this procedure. If one of the failure conditions occurs while AS security is not activated, the UE goes directly into RRC_IDLE without initiation of the RRC Connection Re-establishment procedure.

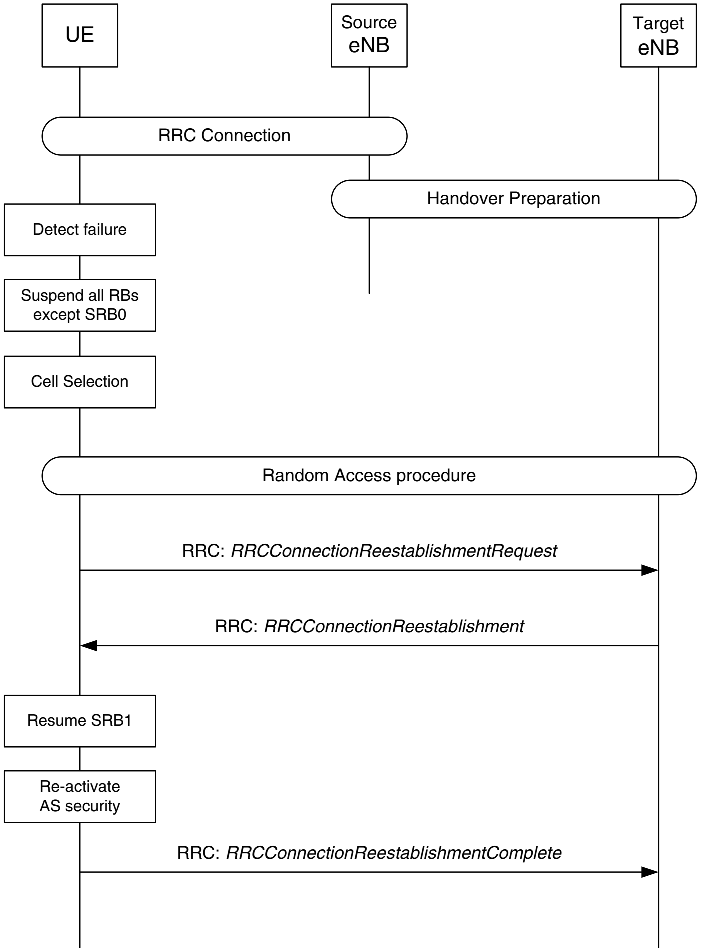

Figure 3.14 RRC Connection Re-establishment. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

As a consequence of RRC Connection Re-establishment, the UE may move from one eNB – the source eNB – to another eNB – the target eNB. A UE in RRC_CONNECTED uses UE-based mobility for the RRC Connection Re-establishment. The UE can re-establish the RRC connection successfully only with the prepared eNB that admitted the UE and established a UE context via handover preparation before this re-establishment, because the target eNB does not fetch a UE context from the source eNB after the re-establishment. The source eNB is allowed to prepare more than one neighboring eNB for the UE.

When one of the failure conditions is met, the UE performs the cell selection process with UE-based mobility to find a suitable cell (see Section 2.2), and then tries to recover a connection with the selected suitable cell as a target cell by requesting an RRC Connection Re-establishment from the target eNB controlling the target cell. It is possible for the source cell controlled by the source eNB to be selected as the target cell as a consequence of this cell selection process.

The UE needs to send the *RRCConnectionReestablishmentRequest* message to the target eNB controlling the selected suitable cell by initiating the Random Access procedure in the MAC layer. System information used in the target cell should be available in the UE before transmission of the *RRCConnectionReestablishmentRequest* message. This message is transmitted on the UL-SCH during the Random Access procedure. The following information elements with a re-establishment cause are included in the *RRCConnectionReestablishmentRequest* message:

- C-RNTI;
- Physical Cell ID (PCI);
- Short MAC-I.

If the re-establishment is initiated due to failure of the handover procedure or the Mobility From E-UTRA procedure, the *RRCConnectionReestablishmentRequest* message includes the C-RNTI used in the source cell, the PCI of the source cell, and the Short MAC-I based on the keys of the source cell. Otherwise, the *RRCConnectionReestablishmentRequest* message includes the C-RNTI used for the cell where the RLF occurs, the PCI of the cell where the RLF occurs, and the Short MAC-I based on the cell where the RLF occurs.

The Short MAC-I corresponds to the 16 least significant bits of the MAC-I (see Section 4.3). Since the PDCP layer does not apply integrity protection and ciphering to the *RRCConnectionReestablishmentRequest* message sent over SRB0, the target eNB uses the Short MAC-I to verify the *RRCConnectionReestablishmentRequest* message. RRC Connection Re-establishment is completed successfully only when the target eNB has been prepared with a UE context. When the target eNB receives the message for re-establishment, the target eNB retrieves a stored UE context that matches the C-RNTI used in the cell identified by the PCI, and then verifies this message with the Short MAC-I. The target eNB can accept the request for RRC Connection Re-establishment only if the target eNB has been prepared with a UE context. Thus, when the target eNB fails to retrieve a matched UE context, the target eNB sends an *RRCConnectionReestablishmentReject* message to the UE to reject the re-establishment.

When the target eNB accepts re-establishment, an *RRCConnectionReestablishment* message is sent to the UE over SRB0. The PDCP layer does not apply integrity protection and ciphering to this message. Upon receiving this message, the UE resumes SRB1, which had been suspended. Then, without algorithm change, the UE updates the K_{eNB} key and derives three AS keys (K_{RRCint}, K_{RRCenc}, and K_{UPenc}) to re-activate AS security. After re-activation of AS security, integrity protection and ciphering are applied to all subsequent messages including the *RRCConnectionReestablishmentComplete* message, which is sent over the resumed SRB1.

The RRC Connection Re-establishment procedure is completed successfully when the UE sends the *RRCConnectionReestablishmentComplete* message to the target cell. Following this re-establishment, the target eNB immediately performs the RRC Connection Reconfiguration procedure to re-establish the RLC and the PDCP for SRB2 and all DRBs, and resumes SRB2 and all DRBs that were suspended. The RRC Connection Reconfiguration procedure will also configure measurements for the UE.

The RRC Connection Re-establishment procedure may fail for several reasons, for example, when the target eNB has not been prepared (i.e., no UE context), when the UE fails to find a suitable cell for the cell selection process, or when the Random Access procedure does not succeed within a certain interval. When the RRC Connection Re-establishment procedure fails, the RRC layer indicates “RRC connection failure” to the NAS layer in the UE, and then the UE goes to RRC_IDLE. The NAS layer can then initiate recovery of this connection; that is, the NAS signaling connection recovery.

3.11 Inter-RAT Mobility

The E-UTRAN supports inter-RAT mobility from/to the other 3GPP access systems and inter-RAT mobility from/to CDMA2000 access systems. The UE can move from/to the other access systems in RRC_IDLE and RRC_CONNECTED. The UEs in the E-UTRAN can perform inter-RAT mobility from/to the following Radio Access Technologies (RATs):

- 3GPP access systems:
  - GSM/EDGE Radio Access Network (GERAN);
  - Universal Terrestrial Radio Access Network (UTRAN).
- CDMA2000 access systems:
  - 1x Radio Transmission Technology (1xRTT);
  - High Rate Packet Data (HRPD).

3.11.1 Inter-RAT Mobility from/to 3GPP Access Systems

Figure 3.15 illustrates inter-RAT mobility between RRC states in the E-UTRAN and states in the UTRAN and the GERAN. The following options can be used for inter-RAT mobility between 3GPP access systems:

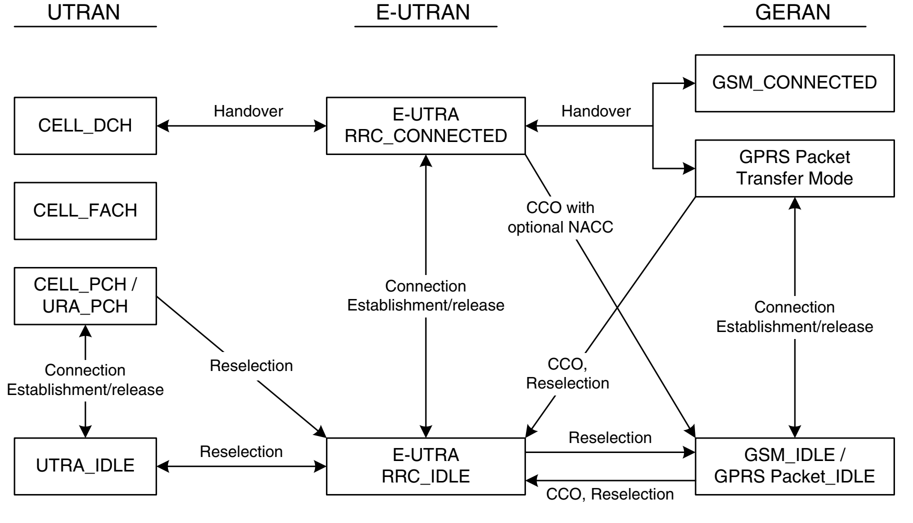

Figure 3.15 Inter-RAT mobility between 3GPP access systems. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

- inter-RAT handover;
- Cell Change Order (CCO) with optional Network Assisted Cell Change (NACC);
- inter-RAT cell reselection.

Inter-RAT handover procedures are used for mobility between the E-UTRAN and the other 3GPP access systems, when the UE is in RRC_CONNECTED on the E-UTRAN, in CELL_DCH on the UTRAN, or in GSM_CONNECTED/GPRS Packet Transfer Mode on the GERAN.

In principle, inter-RAT handover between 3GPP access systems is network-controlled mobility; that is, the source access system decides when to start handover preparation and handover execution. The source access system may apply the inter-RAT measurements procedure for inter-RAT handovers.

For inter-RAT handover, the target access system should prepare radio resources for the UE before the UE executes the handover. The source access system provides the target access system with the necessary information required to prepare for the handover command. The handover command generated by the target access system is transferred to the UE via the source access system. The core network nodes of the source and target access systems, such as the SGSN for the GERAN and the UTRAN, and the MME and the S-GW for the E-UTRAN, are involved in the inter-RAT handover procedure.

For mobility between E-UTRAN and GERAN, the source access system may trigger CCO with the optional NACC. If a UE receives the handover command for CCO with optional NACC, it goes into the idle state of the target access system and then establishes a connection to the target access system. The source access system may provide the NACC in the CCO command to reduce interruption time during cell change, where the UE is provided with the system information of the target cell.

The idle UEs in 3GPP access systems use inter-RAT cell reselection for mobility between the E-UTRAN and the other 3GPP access systems. Inter-RAT cell reselection is also used for mobility of the UE from CELL_PCH or URA_PCH in the UTRAN to RRC_IDLE in the E-UTRAN, and from GPRS Packet Transfer Mode of the GERAN to RRC_IDLE in the E-UTRAN. Details of the cell reselection process for inter-RAT mobility are given in Section 2.7.

Note that no inter-RAT mobility between the E-UTRAN and CELL_FACH in the UTRAN is supported because CELL_FACH is considered a transient state used, for example, for intermittent transmission of packets.

The RRC layer in the E-UTRAN provides the following RRC procedures for inter-RAT mobility from/to the GERAN or the UTRAN:

- Mobility from the E-UTRAN to another 3GPP access system:
  - The Mobility from E-UTRA procedure.
- Mobility from another 3GPP access system to the E-UTRAN:
  - The Handover to E-UTRA procedure.
  - The Inter-RAT Cell Change Order to E-UTRAN procedure.

3.11.1.1 The Mobility from E-UTRA Procedure

The Mobility from E-UTRA procedure supports inter-RAT handovers to the GERAN or the UTRAN, and the CCOs to the GERAN. The RRC layer specifies the *MobilityFromEUTRA-Command* message for the Mobility from E-UTRA procedure. This message indicates the

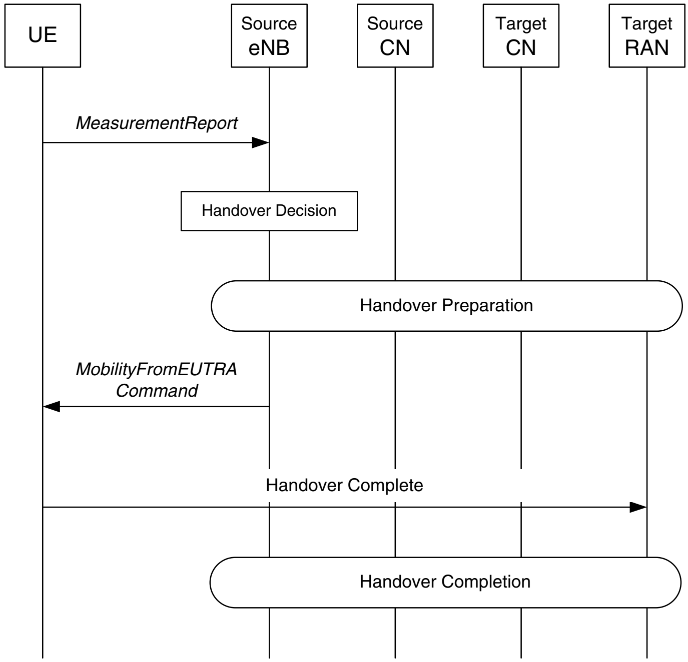

Figure 3.16 Inter-RAT handover from the E-UTRAN to the other 3GPP access systems. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

purpose of the procedure – either inter-RAT handover or cell change order. This procedure also supports inter-RAT mobility from the E-UTRAN to CDMA2000 access systems.

Figure 3.16 explains inter-RAT handover from the E-UTRAN to the other 3GPP access systems with the Mobility from E-UTRA procedure. The source eNB may use inter-RAT measurements to make decisions on inter-RAT handovers. When the source eNB decides to initiate inter-RAT handover to a target 3GPP access system such as the GERAN or the UTRAN, the access networks and the core networks prepare handover before handover execution.

During handover preparation, the target RAN sends a handover command to the UE via the source eNB. The source eNB uses the *MobilityFromEUTRACommand* message to forward the handover command to the UE. This message indicates the type of target access system; that is, either the GERAN or the UTRAN.

The *MobilityFromEUTRACommand* message indicating *handover* triggers the UE to initiate handover execution for inter-RAT handover. When this message is received, the UE switches from the E-UTRAN to the target RAN of the target cell indicated in the received message. When the UE synchronizes successfully with the target cell, the UE sends a message completing this inter-RAT handover to the target cell. For handover completion, the network switches the data path of packets from the source eNB to the target RAN and also releases UE-associated resources still left in the source eNB.

When the Mobility from E-UTRA procedure is used for a CCO from the E-UTRAN to the GERAN, the *MobilityFromEUTRACommand* message indicates *cellChangeOrder*. Upon receiving this message from the E-UTRAN, the UE initiates the CCO and establishes a connection to the GERAN cell. For support of the optional NACC, this message can contain system information of the target GERAN cell. If no system information is contained in this message, the UE should acquire system information from the GERAN cell when accessing the GERAN cell.

The Mobility from E-UTRA procedure may fail, for instance, when the UE fails to establish a connection to the target cell, or when the UE cannot comply with the configuration included in the *MobilityFromEUTRACommand* message. When the Mobility from E-UTRA procedure fails, the source access system – the E-UTRAN – handles this failure by performing the RRC Connection Re-establishment procedure.

3.11.1.2 The Handover to E-UTRA Procedure

For inter-RAT handovers from the GERAN or the UTRAN to the E-UTRAN, the RRC layer provides the Handover to E-UTRA procedure. As a result of this procedure, the UE activates AS security for the E-UTRAN if security has not yet been activated in the source RAT, and also establishes SRB1, SRB2, and one or more DRBs.

Figure 3.17 explains inter-RAT handover from the other 3GPP access systems to the E-UTRAN with the RRC Connection Reconfiguration procedure. This Handover to E-UTRA procedure utilizes the RRC Connection Reconfiguration procedure.

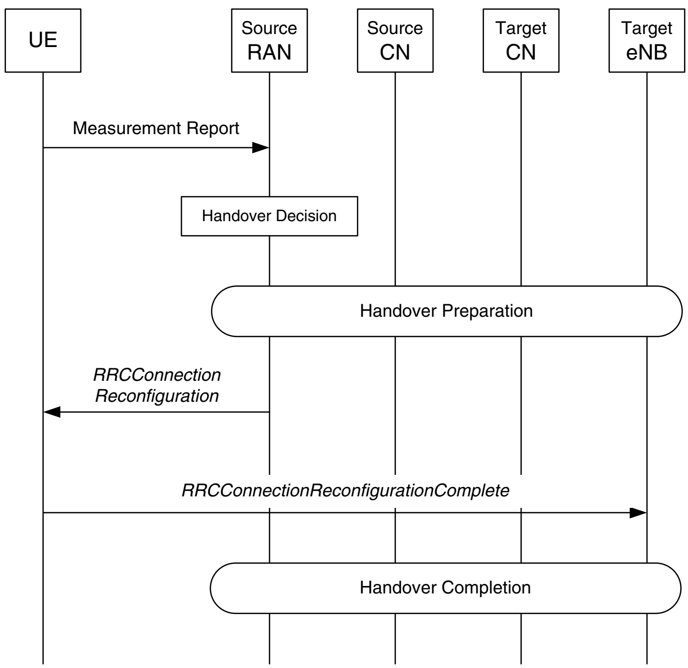

Figure 3.17 Inter-RAT handover from the other 3GPP access systems to the E-UTRAN. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

When the source RAN makes a decision on inter-RAT handover to the E-UTRAN, the target eNB sends the *RRCConnectionReconfiguration* message to the source RAN, and then the source RAN forwards this message to the UE. Upon receiving this message, the UE accesses the target eNB and then sends the *RRCConnectionReconfigurationComplete* message.

In general, handling of inter-RAT mobility failure is specified in the specifications of the source RAN rather than the target RAN. Thus, when the Handover to E-UTRA procedure fails, the UE handles the failure according to the specifications of either the GERAN or the UTRAN that corresponds to the source RAN.

3.11.1.3 The Inter-RAT Cell Change Order to E-UTRAN Procedure

For a CCO from the GERAN to the E-UTRAN, the RRC layer provides the Inter-RAT Cell Change Order to E-UTRAN procedure. Upon receiving a message corresponding to cell change order from the GERAN to E-UTRAN, the UE initiates the Inter-RAT Cell Change Order to E-UTRAN procedure. Then, the UE performs the RRC Connection Establishment procedure for the target E-UTRAN cell.

If the Inter-RAT Cell Change Order to E-UTRAN procedure fails, the UE returns to the GERAN. The UE handles the failure according to the specifications of the GERAN.

3.11.2 Inter-RAT Mobility from/to CDMA2000 Systems

Figure 3.18 illustrates inter-RAT mobility between RRC states in the E-UTRAN and states in the CDMA access systems that include 1xRTT and the HRPD. Inter-RAT handover as well as inter-RAT cell reselection can be used for mobility between E-UTRAN and CDMA2000 systems.

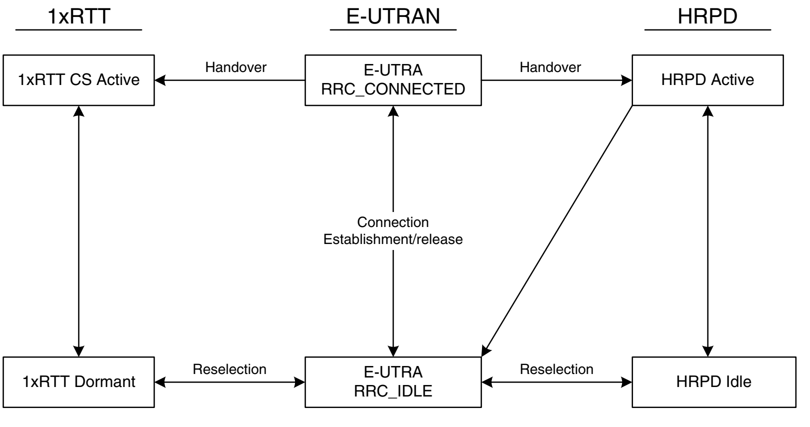

Figure 3.18 Inter-RAT mobility between the E-UTRAN and CDMA2000 access systems. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

3.11.2.1 Inter-RAT Handover from the E-UTRAN

Figure 3.19 explains inter-RAT handover to the CDMA2000 access system. In general, inter-RAT handover from the E-UTRAN to the CDMA2000 access system consists of two phases:

- a pre-registration phase;
- a handover phase.

To enable inter-RAT handover from the E-UTRAN to the CDMA2000 access system, the UE connected to E-UTRAN may perform pre-registration, where the UE registers with the CDMA2000 system while the UE is still connected to the E-UTRAN. The CDMA2000 messages used for pre-registration to CDMA2000 are exchanged via tunneled signaling over the EPS between the UE and the CDMA2000 system. The *DLInformationTransfer* message and the *ULInformationTransfer* message are used to exchange the CDMA2000 information for

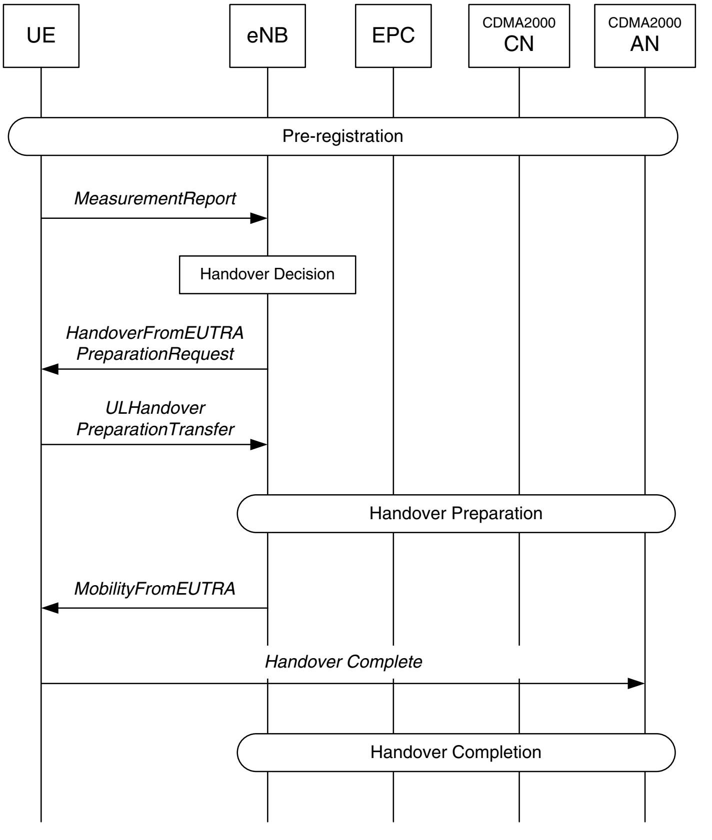

Figure 3.19 Inter-RAT mobility from the E-UTRAN to the CDMA2000 access system. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

session establishment and authentication with the CDMA2000 system in the pre-registration phase.

The eNB can indicate whether or not the UE is allowed to perform pre-registration in the E-UTRAN via system information. This is indicated in *SystemInformationBlockType8* that contains information on inter-RAT cell reselection to the CDMA2000 system. The eNB can also indicate this via the *RRCConnectionReconfiguration* message.

For a UE that has registered with the CDMA2000 access system, the source eNB can trigger handover of the UE to the CDMA2000 access system. The *MeasurementReport* message for inter-RAT measurements on CDMA2000 frequencies may trigger the source eNB to initiate handover.

The RRC layer in the E-UTRAN additionally provides the following RRC procedures for handover preparation toward the CDMA2000 access system:

- the Handover from E-UTRA Preparation Request procedure;
- the UL Handover Preparation Transfer procedure.

Upon a handover decision, the source eNB requests the UE to prepare for handover by initiating the Handover from E-UTRA Preparation Request procedure. For this procedure, the source eNB sends the *HandoverFromEUTRAPreparationRequest* message to the UE. Upon receiving the message, the RRC layer of the UE forwards information contained in this message to the CDMA2000 upper layers. In response to this message, the CDMA2000 upper layers request the RRC layer to initiate the UL Handover Preparation Transfer procedure.

When the UL Handover Preparation Transfer procedure is initiated, the RRC layer of the UE sends handover-related information to the eNB via the *ULHandoverPreparationTransfer* message. This message is transported transparently to the CDMA2000 system via tunneling during handover preparation. When the CDMA2000 system is prepared for handover, information necessary to access the target cell is transported to the source eNB during handover preparation. The source eNB sends this information to the UE via the *MobilityFromEUTRACommand* message.

For inter-RAT mobility from the E-UTRAN to the CDMA2000 access system, the source eNB may re-direct the UE to the CDMA2000 access system by initiating the RRC Connection Release procedure after the pre-registration phase, instead of inter-RAT handover.

3.12 RRC Connection Release

The eNB sends the *RRCConnectionRelease* message to the UE in order to release an RRC connection. Upon receipt of the *RRCConnectionRelease* message, the UE leaves RRC_CONNECTED and enters RRC_IDLE with cell selection. In addition, upon request from the NAS, the UE may internally release an RRC connection, without receipt of the *RRCConnectionRelease* message.

The *RRCConnectionRelease* message includes *releaseCause*. In order to offload a UE in ECM-CONNECTED, the eNB can transmit the *RRCConnectionRelease* message including *releaseCause* set to *loadBalancingTAURequired* to the UE. The *loadBalancingTAURequired* in the *RRCConnectionRelease* message triggers the UE to establish an RRC connection without indicating S-TMSI and the GUMMEI, which may cause the eNB to select another MME for the UE for offloading.

In addition, the eNB can optionally include "redirection information" in the *RRCConnectionRelease* message – that is, *redirectedCarrierInfo* – in order to redirect the UE to another carrier frequency or another RAT including the GERAN, UTRAN, and CDMA2000 systems. When the UE receives an *RRCConnectionRelease* message with redirection information, the UE attempts to access a cell on the carrier frequency/RAT indicated by the redirection information, after release of the RRC connection. The redirection information in the *RRCConnectionRelease* message can be used by the UE to apply CS fallback (see Section 8.3).

Furthermore, the eNB can optionally include *idleModeMobilityControlInfo* in the *RRCConnectionRelease* message in order to provide information about cell reselection priority dedicated to the UE. How the UE handles the cell reselection priority after leaving RRC_CONNECTED is explained in Section 2.7.

Reference

1. 3GPP Technical Specification 36.331, "Radio Resource Control (RRC); Protocol Specification (Release 10)", www.3gpp.org, accessed on 14 March 2012.

4

Packet Data Convergence Protocol (PDCP)

In the LTE radio protocol stack, the PDCP layer is located above the RLC layer and below the IP layer (in the user plane) or the RRC layer (in the control plane). 3GPP provides a PDCP specification in [1]. From UMTS, the main function of the PDCP layer has been the header compression of IP packets – this is the reason why it is called the Packet Data Convergence Protocol.

In LTE, the PDCP layer has evolved to support a security function; that is, integrity protection and ciphering. To provide time-varying characteristics to the security function, a PDCP sequence number has been introduced in the PDCP layer. The PDCP sequence number is attached to each PDCP PDU, and it is used to generate different security output per PDCP PDU. Thanks to the PDCP sequence number, the PDCP layer is able to perform ARQ-related functions, which can improve the radio efficiency at handover.

The PDCP layer is running on top of the RLC layer. A PDCP entity which is used to perform the PDCP functions can be configured either with both transmitting and receiving sides (for a bidirectional radio bearer), or only one of them (for a unidirectional radio bearer). If a PDCP entity is configured with both transmitting and receiving sides, it is associated either with one AM RLC entity or two UM RLC entities (one for the uplink direction and one for the downlink direction). If a PDCP entity is configured with either a transmitting or a receiving side, it is associated with one UM RLC entity of the same direction. One PDCP entity is used for only one radio bearer.

4.1 PDCP Functions and Architecture

Radio bearers utilizing PDCP entities can be categorized into three types:

- **SRB:** Signaling radio bearer using AM RLC;
- **AM DRB:** Data radio bearer using AM RLC;
- **UM DRB:** Data radio bearer using UM RLC.

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

Depending on the radio bearer characteristics and the mode of the associated RLC entity, the following functions are selectively performed by the PDCP entity:

- header compression using Robust Header Compression (ROHC) for DRB;
- security functions:
  - integrity protection for SRB;
  - ciphering for SRB and DRB;
- maintenance of PDCP Sequence Numbers for SRB and DRB;
- handover support functions:
  - status reporting for AM DRB;
  - duplicate elimination of lower layer SDUs for AM DRB;
  - in-order delivery of upper layer PDUs for AM DRB;
- timer-based SDU discard for SRB and DRB.

The functional view of the PDCP entity is shown in Figure 4.1. Note that the PDCP entity in Figure 4.1 consists of both transmitting and receiving sides, and hence it is used for a

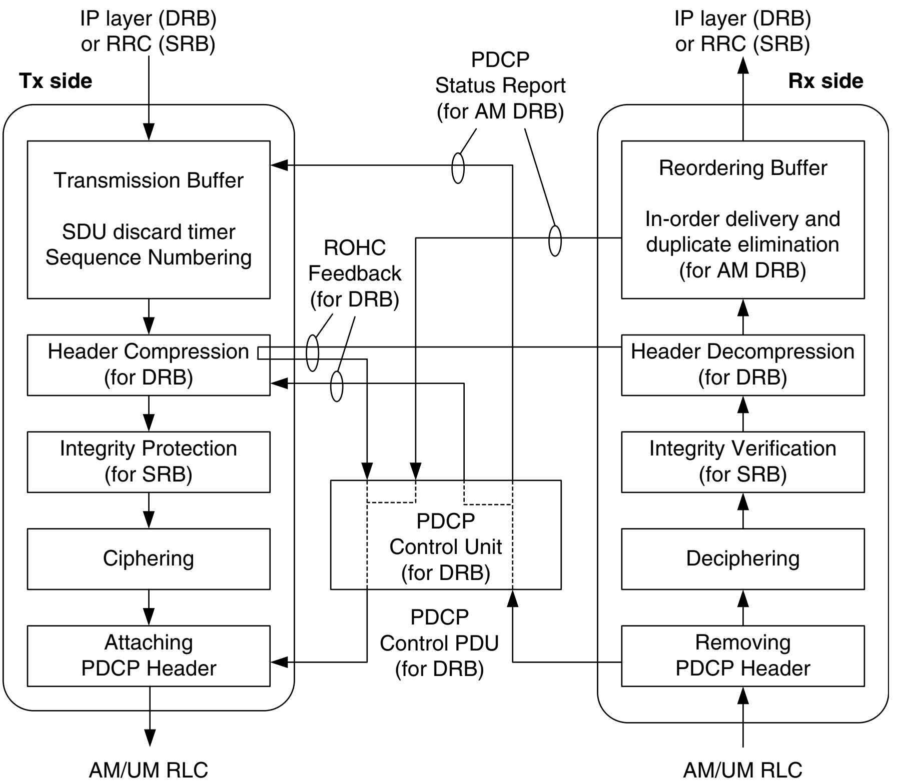

Figure 4.1 Functional view of PDCP entity. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

bidirectional radio bearer. For a unidirectional radio bearer, the PDCP entity consists of either a transmitting side or a receiving side without a PDCP control unit.

The PDCP control unit manages control information generated by the PDCP entity. Currently, two kinds of control information are defined: PDCP status report and ROHC feedback. Since they are not related to PDCP SDUs, no PDCP functions (e.g., sequence numbering, header compression, and security) are applied.

4.2 Header Compression

Every IP packet has a few tens of bytes of header, which may be a considerable amount compared with the payload size. For example, in a Voice over IP (VoIP) service where RTP/UDP/IPv6 protocols are used, every packet contains 60 bytes of header while the payload size is typically 20 $\sim$ 30 bytes. Therefore, header compression, a mechanism to reduce the size of header overhead, is essential to improve transmission efficiency, especially for VoIP services.

The basic idea of header compression is to transmit only the difference from the reference header, as shown in Figure 4.2. For a single IP flow, many header fields are identical between packets. Thus, transmission of the same information in every packet is redundant and the root cause of inefficiency.

To avoid such redundancy in every packet, the compressor first transmits the full header packet to establish the header context for an IP flow in the decompressor. The header context contains the reference header together with the identifier for the IP flow, called Context ID (CID). Once the header context is established in the decompressor, for the subsequent packets of the IP flow, the compressor transmits the compressed packet including the compressed header carrying the difference between the original packet header and the established header context. The difference is usually the difference of dynamic parts of the packet header. Then,

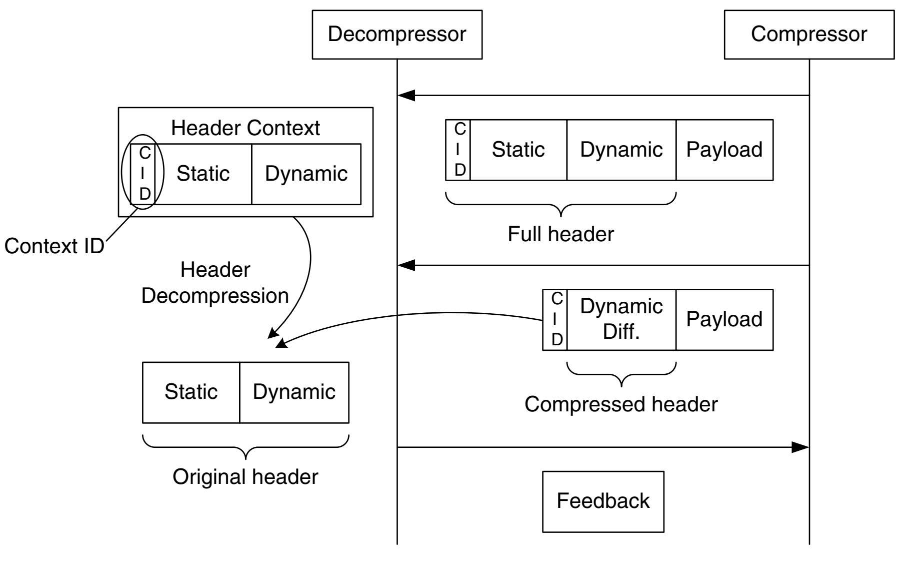

Figure 4.2 Conceptual view of header compression algorithm

Table 4.1 Supported ROHC profiles. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

Profiles	Type of IP Flows	Header Compression Algorithm
ROHC uncompressed	No compression	RFC4995
ROHC RTP	RTP/UDP/IP	RFC3095, RFC4815
ROHC UDP	UDP/IP	RFC3095, RFC4815
ROHC ESP	ESP/IP	RFC3095, RFC4815
ROHC IP	IP	RFC3843, RFC4815
ROHC TCP	TCP/IP	RFC4996
ROHCv2 RTP	RTP/UDP/IP	RFC5225
ROHCv2 UDP	UDP/IP	RFC5225
ROHCv2 ESP	ESP/IP	RFC5225
ROHCv2 IP	IP	RFC5225

when the decompressor receives the compressed packet, it can reconstruct the original packet by recovering the packet header based on the established header context and the difference in the compressed packet.

If the header decompression is not successful, which can be detected by a CRC error in header decompression, the decompressor transmits feedback information to the compressor to indicate that the header context has been corrupted. This feedback information can be transmitted either as a standalone packet or piggybacked to a data packet in the opposite direction. When the compressor receives the feedback information, it will transmit the full header packet again to re-establish the correct header context.

In LTE, header compression is based on the Robust Header Compression (ROHC) framework defined by the Internet Engineering Task Force (IETF). ROHC is an enhanced version of the header compression algorithm that is used in wired links. The term "robust" means that this header compression algorithm is performed well in lossy links, such as wireless links.

In the ROHC framework, multiple header compression algorithms are defined, each one being called an ROHC profile. The profile is specific to the particular combination of IP flows, such as RTP/UDP/IP or TCP/IP. The supported ROHC profiles are listed in Table 4.1.

Support for ROHC is not mandatory for the UE, but is dependent on the UE's capability. Support for each ROHC profile also depends on the UE's capability. Thus, the UE is required to inform the eNB of its supported ROHC profiles. From this information, the eNB selects the ROHC profiles that will be used for each DRB. The UE is then informed of the selected ROHC profiles by RRC signaling. In this way, the ROHC profiles to be used are synchronized between the compressor and the decompressor peers for each DRB.

Among the ROHC profiles, the profile "ROHC uncompressed" is mandatorily supported if the UE wants to support at least one ROHC profile. This is because the algorithm "RFC4995" is the basic framework of ROHC profiles, including configuration parameters for compressor and decompressor peers.

The UE wanting to support a VoIP service is required to support the profiles "ROHC RTP" and "ROHC UDP". Since a VoIP packet is transported over RTP/UDP/IP layers, support for those profiles is essential for a VoIP service. Using ROHC, the header size of an RTP/UDP/IP packet can be reduced from 40 (for IPv4) or 60 (for IPv6) bytes to $1 \sim 3$ bytes.

4.3 Security

The security function includes both integrity protection and ciphering. The configuration of the security function is the responsibility of the RRC, but the responsibility for actually performing the security function falls to the PDCP. When security activation is indicated by the RRC layer, the PDCP layer applies the security function to all PDCP SDUs for the downlink and the uplink.

The PDCP layer manages one of the security input parameters, called COUNT. The COUNT has a length of 32 bits, consisting of the Hyper Frame Number (HFN) and the PDCP Sequence Number (SN). The length of the PDCP SN can be 5 bits, 7 bits, or 12 bits depending on the radio bearer's characteristics. The PDCP SN constitutes the Lease Significant Bit (LSB) part of the COUNT, with the HFN comprising the remainder of COUNT.

The PDCP SN is attached explicitly to a PDCP PDU, while the HFN is maintained internally in both transmitter and receiver. The PDCP SN increases by one for each PDCP SDU while the HFN increases by one for each PDCP SN wrap around. A PDCP SN wrap around means that the value of PDCP SN becomes zero after reaching the maximum value due to the limited number of bits for the PDCP SN. Therefore, the resulting COUNT value can provide a time-varying input to the security algorithm. The time-varying characteristic of COUNT provides protection against the well-known replay attack.

4.3.1 Integrity Protection

Integrity protection is used to detect whether the packet has been replaced or inserted by unauthorized parties during delivery. It is realized by a special code, called the Message Authentication Code-Integrity (MAC-I). The MAC-I has a length of 4 bytes, and is generated from the integrity protection algorithm with the following input parameters:

- **KEY:** Access Stratum (AS) derived key for integrity protection;
- **COUNT:** HFN + PDCP SN;
- **BEARER:** The radio bearer ID;
- **DIRECTION:** The direction of the radio bearer – that is, uplink or downlink;
- **MESSAGE:** The message itself to be integrity protected.

For each PDCP SDU, the PDCP transmitter generates a MAC-I and attaches it to the PDCP SDU to generate a PDCP PDU. When the PDCP receiver receives a PDCP PDU, it also generates its own MAC-I, called XMAC-I, based on the received PDCP PDU, and compares the generated XMAC-I with the received MAC-I. If they are identical, the integrity is verified. But if they are different, the PDCP receiver declares integrity verification failure, and indicates the failure to the RRC to initiate the RRC Connection Re-establishment procedure. The overall operation of integrity protection is shown in Figure 4.3.

Since the 4-byte MAC-I is attached to each PDCP PDU, the integrity protection increases overhead in the PDCP PDU. Therefore, only important packets such as RRC messages are integrity protected. In other words, integrity protection is applied only for SRBs.

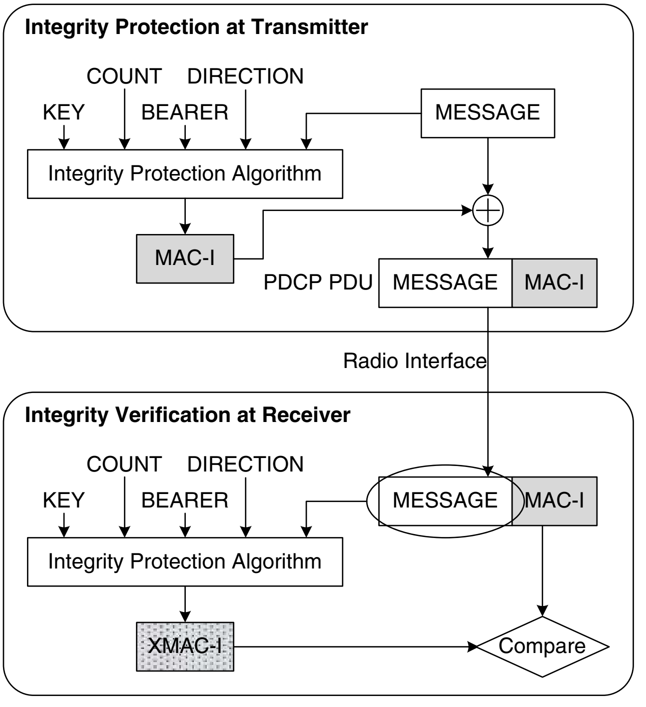

Figure 4.3 Integrity protection operation

4.3.2 Ciphering

The purpose of ciphering is to maintain the confidentiality of messages between a sender and a receiver. The original message is masked (i.e., XOR operation) by a ciphering keystream, and a third party cannot recover the original message unless it has the same ciphering keystream. The ciphering keystream is the output of the ciphering algorithm with the following input parameters:

- **KEY:** AS-derived key for ciphering;
- **COUNT:** HFN + PDCP SN;
- **BEARER:** The radio bearer ID;
- **DIRECTION:** The direction of the radio bearer – that is, uplink or downlink;
- **LENGTH:** The length of the ciphering keystream to be generated.

Ciphering is performed for each PDCP SDU. The PDCP transmitter generates a ciphering keystream and performs a XOR operation on the PDCP SDU and the ciphering keystream. On the receiver side, another XOR operation on the received PDCP PDU and the same ciphering keystream recovers the original PDCP SDU. The overall operation of ciphering is shown in Figure 4.4.

It is important to notice that there are no means for the PDCP receiver to detect ciphering failure. In other words, there are no means for the PDCP receiver to be sure that it is applying

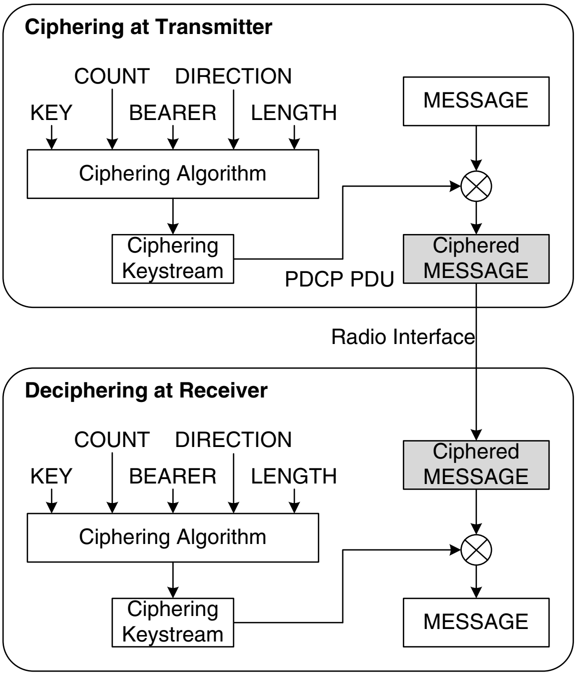

Figure 4.4 Ciphering operation

the same ciphering keystream as that of the PDCP transmitter. However, for SRBs, because an incorrect ciphering keystream would also cause corruption of the integrity verification key, ciphering failure can be handled by integrity verification failure. On the other hand, for DRBs, ciphering failure cannot be handled by integrity verification failure (because integrity protection is not applied to DRBs), but it does cause failure in header decompression.

When ciphering failure occurs due to the de-synchronization of security parameters such as COUNT, an incorrect ciphering keystream would corrupt the compressed header of a PDCP PDU. In this case, header decompression would keep failing without performing any recovery mechanism. As a solution to this situation, the PDCP receiver may be implemented to detect ciphering failure by monitoring whether there is continuous header decompression failure. An alternative solution would be to make the upper layer responsible for detection. When service quality deteriorates due to ciphering failure, the user of the UE may be able to refresh the RRC connection.

The ciphering operation does not increase any overhead, and thus it is applied for all packets passing through the PDCP entity.

4.4 Data Transfer

The data transfer procedure in PDCP is relatively simple, in that all applicable functions in Figure 4.1 are applied sequentially. However, the applicable functions are different depending on the type of radio bearer.

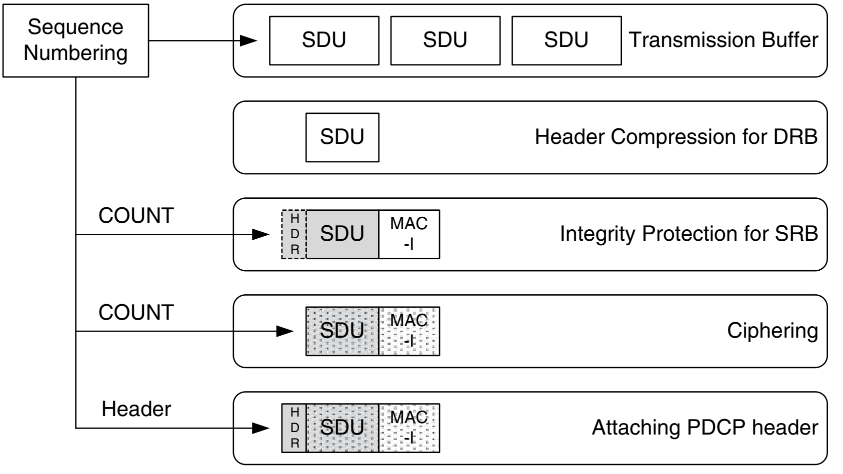

Figure 4.5 Sequence numbering in the PDCP transmitter

In the PDCP transmitter, the PDCP SDUs received from the upper layer are stored in the transmission buffer. The PDCP transmitter may, at any time, allocate a PDCP SN to a stored PDCP SDU. Note that not all the PDCP SDUs stored in the transmission buffer need to be allocated with PDCP SNs. When the PDCP transmitter decides to transmit the PDCP SDUs, it starts to process them in ascending order of the associated PDCP SN. The process takes the sequential steps: header compression (for DRBs), integrity protection (for SRBs), and ciphering (for DRBs and SRBs). The allocated PDCP SN is used to generate a COUNT value that will be used for integrity protection and ciphering. Integrity protection is applied to the virtual PDU header that will be attached after ciphering and the data part of the PDCP PDU. Ciphering is applied to the data part of the PDCP PDU and the MAC-I field if integrity protection is applied. After ciphering, the PDCP SN is attached to a PDCP header to construct the PDCP PDU. The constructed PDCP PDU is then delivered to the lower RLC layer at a suitable point in time. The transmitter behavior with sequence numbering is shown in Figure 4.5.

The PDCP receiver takes the reverse steps to the PDCP transmitter. The PDCP PDU received from the RLC layer is first subjected to header removal and then passed through the functions deciphering (for DRBs and SRBs), integrity verification (for SRBs), and header decompression (for DRBs), sequentially. With these steps, the original PDCP SDU is restored, and it is then delivered to the upper layer.

One aspect that needs to be considered in the PDCP receiver is PDCP SDU reordering. In normal situations, reordering is not performed. SDU reordering is required for AM DRBs at handover to handle the PDCP SDUs retransmitted after handover. This is explained further in Section 4.6.

4.5 SDU Discard

The SDU discard function is used to prevent buffer overflow in the transmitter. Each SDU can stay in the transmission buffer only for the configured time period in the PDCP entity.

The value of the configured time period depends on the delay requirement of the radio bearer. Thus, the PDCP SDU discard function is also used to ensure that the requirement of the maximum transmission delay of the PDCP SDU is met.

Upon receipt of an SDU from the upper layer, the PDCP transmitter starts a discard timer for the SDU. Once started, the discard timer never stops until expiry. When the discard timer expires, the PDCP transmitter discards the PDCP SDU and indicates to the RLC to discard the corresponding PDCP PDU if the PDU has already been delivered to the RLC layer. The RLC layer then discards the PDCP PDU if no segments of the PDCP PDU have been transmitted to the receiver (see Section 5.6).

The PDCP transmitter may also discard the PDCP SDU when successful transmission of the PDCP PDU is confirmed by the PDCP status report. Note, however, that the use of a PDCP status report is optional for AM DRBs at handover.

4.6 Handover

The UE performs handover when it changes its serving cell. During handover, the PDCP layer as well as the RLC layer is re-established upon request from the RRC layer. During re-establishment, the PDCP layer performs several actions such as resetting the header compression algorithm and context, changing the security algorithm and key, and so on.

The important thing to note in PDCP re-establishment is that some behaviors are specific to the type of radio bearer:

- maintenance of COUNT (i.e., HFN and PDCP SN);
- processing of PDCP SDUs stored in the transmission buffer;
- processing of PDCP PDUs received due to RLC re-establishment (see Section 5.7).

The following subsections describe the PDCP behavior at handover for each type of radio bearer.

4.6.1 SRB Behavior at Handover

For SRBs, all parameters and buffers are reset in both the transmitter and receiver at handover. Since the RRC messages generated in the source cell may not be relevant any more in the target cell, it is of no use to transmit or receive them after handover. Though the RRC messages are used also for the transport of NAS layer messages, the NAS layer has its own mechanism to handle the unsuccessful delivery of NAS layer messages.

At PDCP re-establishment, the COUNT value is initialized to zero and the PDCP SDUs stored in the transmission buffer are discarded. If some PDCP PDUs are received from the RLC due to RLC re-establishment, they are also discarded. To summarize, PDCP re-establishment for an SRB is equivalent to re-initialization of the PDCP entity.

4.6.2 UM DRB Behavior at Handover

For UM DRBs, the COUNT value is initialized but the PDCP SDUs and PDUs are not discarded at handover. Services using UM DRBs are usually not sensitive to packet loss. However, discarding the PDCP SDUs and PDUs degrades the QoS, so it is better to minimize packet discarding as much as possible.

During handover, the RLC layer as well as the PDCP layer is re-established. As a result, in the PDCP receiver, some PDCP PDUs can be received from the RLC layer due to RLC re-establishment. To minimize the loss of these PDCP PDUs, the PDCP receiver processes the received PDCP PDUs to reconstruct the PDCP SDUs before performing PDCP re-establishment. The reconstructed PDCP SDUs are delivered to the upper layer in ascending order of the associated PDCP SN. Only after delivering them to the upper layer can the PDCP receiver perform the PDCP re-establishment.

In the PDCP transmitter, the PDCP SDUs stored in the transmission buffer are processed again after handover as if they are newly received from the upper layer. Regardless of whether PDCP SNs were allocated or not to the PDCP SDUs before handover, new PDCP SNs are allocated to the PDCP SDUs. It should be noted, however, that the discard timers are not restarted for those PDCP SDUs, as explained in Section 4.5. On the network side, the source eNB forwards to the target eNB downlink PDCP SDUs that have not been transmitted before handover. Then, the target eNB treats the forwarded PDCP SDUs as new PDCP SDUs received from the upper layer. The target eNB prioritizes the forwarded PDCP SDUs over the PDCP SDUs received from the Serving Gateway (S-GW).

Figure 4.6 shows an example of UM DRB behavior at handover for the downlink case. In this example, the source eNB first transmits PDCP SDUs 11 ~ 13 before handover. Upon handover, the source eNB forwards not transmitted PDCP SDUs 14 and 15 to the target eNB. The target eNB then reallocates new PDCP SNs 0 and 1 to the forwarded PDCP SDUs 14 and 15, respectively, and transmits them to the UE. The new PDCP SDUs received from the S-GW are numbered after the forwarded PDCP SDUs as 2, 3, and 4, and transmitted sequentially.

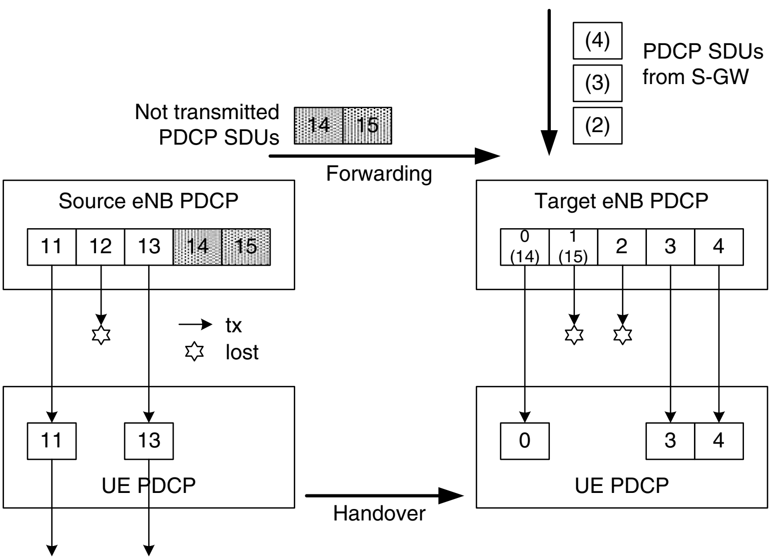

Figure 4.6 UM DRB behavior at handover – downlink case

4.6.3 AM DRB Behavior at Handover

An AM DRB is used to support lossless transmission. In normal situations, lossless transmission is guaranteed by the AM RLC entity. The AM RLC receiver negatively acknowledges RLC PDUs that have not been received correctly using the RLC status report, and the AM RLC transmitter retransmits them when the RLC status report is received. Retransmission is performed repeatedly until all RLC PDUs have been received correctly by the AM RLC receiver or the maximum number of retransmissions is reached (see Section 5.4.3).

At handover, however, lossless transmission cannot be guaranteed by the AM RLC entity because the RLC status report is usually not up to date. Even if the RLC status report is transmitted just before handover, it cannot be guaranteed that the RLC status report will be received successfully by the AM RLC receiver. Thus, at the point of time of handover, there may be outstanding RLC PDUs that are not yet acknowledged by the RLC status report. The issue, then, is how to guarantee lossless transmission of the PDCP SDUs included in the outstanding RLC PDUs.

A simple solution is to perform retransmission in the PDCP layer. If some PDCP SDUs were not acknowledged (by the RLC layer) before handover, the PDCP transmitter retransmits them after handover. Since only the unacknowledged PDCP SDUs are retransmitted, this scheme is called "selective retransmission".

To support selective PDCP SDU retransmission, the PDCP receiver is required to process the PDCP PDUs received as a result of RLC re-establishment before performing PDCP re-establishment. The reconstructed PDCP SDUs are stored in the reordering buffer if they are out of sequence, or delivered to the upper layer otherwise. After handover, the PDCP SDUs not acknowledged before handover are retransmitted by the PDCP transmitter, and the PDCP receiver reorders them together with the PDCP SDUs stored in the reordering buffer to provide in-order delivery to the upper layer.

For the reordering to work, all PDCP state variables including COUNT need to be maintained at handover. Since the reordering is performed based on the PDCP SN, the PDCP transmitter needs to retransmit the unacknowledged PDCP SDUs with the same PDCP SNs.

The network-side behavior to support retransmission and reordering is more complex because the location of the PDCP entity is changed from the source eNB to the target eNB. For continuous PDCP operation before and after handover, the PDCP state variables are forwarded from the source eNB to the target eNB. In addition, the PDCP transmitter forwards unacknowledged and not transmitted PDCP SDUs, and the PDCP receiver forwards out-of-sequence PDCP SDUs stored in the reordering buffer.

PDCP retransmission can be further optimized using the PDCP status report. As explained before, the RLC status report is not up to date. Thus, even though receipt of a PDCP SDU has not been acknowledged by the RLC status report, it is possible that the PDCP SDU was transmitted successfully before handover. In this case, the retransmission of this PDCP SDU after handover is useless.

To avoid unnecessary PDCP retransmission, the PDCP receiver sends a PDCP status report to the PDCP transmitter after handover. The PDCP status report describes the latest reordering buffer status of the PDCP receiver, so it can avoid unnecessary retransmission of PDCP SDUs that have already been received correctly. Note that the use of a PDCP status report is limited to the handover case, and is optionally configured for each AM DRB.

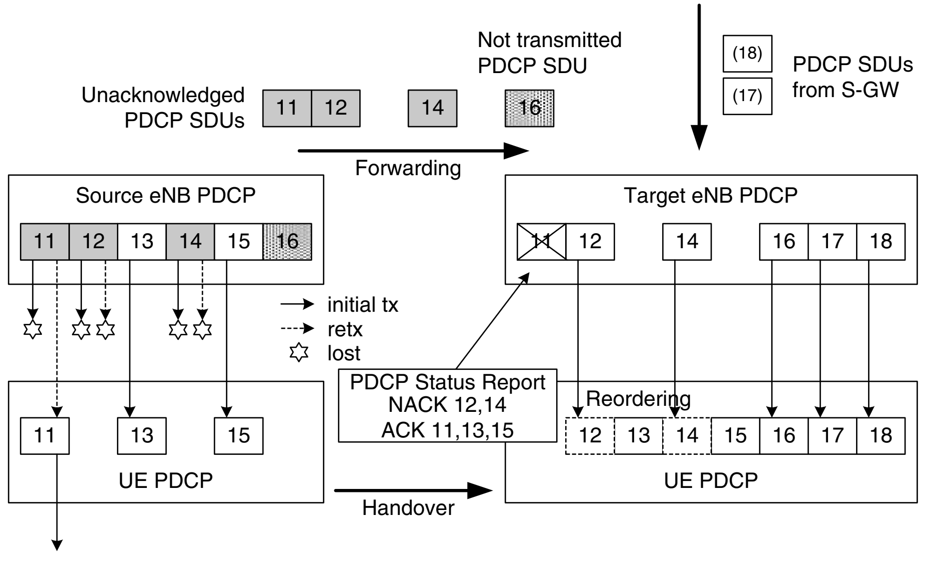

Figure 4.7 AM DRB behavior at handover – downlink case

An example of AM DRB behavior at handover is illustrated in Figure 4.7. The figure shows only the downlink case, but similar behavior applies to the uplink case. In this example, the source eNB first transmits PDCP SDUs 11 $\sim$ 15, but only PDCP SDUs 13 and 15 are received correctly by the UE. Based on the RLC status report, the source eNB retransmits PDCP SDUs 11, 12, and 14 in the RLC layer, and this time only PDCP SDU 11 is received correctly by the UE. The PDCP SDU 11 is in sequence, and thus is delivered to the upper layer. Handover occurs at this point in time.

During handover, the source eNB forwards unacknowledged PDCP SDUs 11, 12, and 14, and not transmitted PDCP SDU 16 to the target eNB. The target eNB stores them in the transmission buffer to prepare for transmission after handover. In the meantime, the UE PDCP processes the out-of-sequence PDCP SDUs 13 and 15 received due to RLC re-establishment and stores them in the reordering buffer.

After handover, if the use of a PDCP status report is configured, the UE sends the PDCP status report to the target eNB, indicating that PDCP SDUs 12 and 14 have not been received, and that PDCP SDUs 11, 13, and 15 have been received correctly. Based on the PDCP status report, the target eNB recognizes that PDCP SDU 11 was transmitted successfully before handover, and discards it from the transmission buffer. The target eNB retransmits only PDCP SDUs 12 and 14 to the UE. Upon receipt of PDCP SDUs 12 and 14, the UE reorders them with the already stored PDCP SDUs 13 and 15, and delivers PDCP SDUs 12 $\sim$ 15 to the upper layer in sequential order.

In the target eNB, the not-transmitted PDCP SDU 16 forwarded by the source eNB and the new PDCP SDUs 17 and 18 received from the S-GW are numbered after the last PDCP SDU transmitted by the source eNB. For this purpose, the source eNB provides the PDCP SN of the last transmitted PDCP SDU to the target eNB. When allocating PDCP SNs, the target PDCP prioritizes PDCP SDUs forwarded by the source eNB over new PDCP SDUs received from the S-GW.

If the use of a PDCP status report is not configured, the UE does not send a PDCP status report and the target eNB will retransmit PDCP SDU 11 in addition to SDUs 12 and 14. In this case, PDCP SDU 11 is discarded in the UE PDCP based on a sequence number check of the PDCP SDUs. This functionality is called *duplicate elimination* and ensures that the upper layer receives a packet only once.

4.7 PDCP PDU Formats

In PDCP, two types of PDU are defined: PDCP Data PDUs and PDCP Control PDUs. A PDCP Data PDU is used to transport user plane data such as an IP packet or control plane data such as an RRC message. A PDCP Control PDU is used to transport PDCP control information such as ROHC feedback or a PDCP status report. The PDCP Control PDU is not used for SRBs.

In what follows, the PDU format describes only the essential fields of the PDCP PDU. The actual PDU format may contain some reserved fields that are used for byte alignment of the PDCP PDU.

4.7.1 PDCP Data PDU Formats

The PDCP Data PDU consists of a PDCP Sequence Number (SN) field and a Data field. The PDCP SN field is used to identify each PDCP SDU, and the Data field is used to carry one PDCP SDU.

For SRBs, where integrity protection is applied, a 4-byte MAC-I field is attached at the end of the PDCP Data PDU. As a result, the format in Figure 4.8 is used for SRBs. The length of the PDCP SN field is 5 bits, which is deemed sufficient for RRC message transmission.

For DRBs, the MAC-I field is not attached because integrity protection is not applied. Instead, a 1-bit identifier is attached in order to distinguish between a PDCP Data PDU and a PDCP Control PDU. The 1-bit identifier is called “D/C”, and is attached at the front of the PDCP PDU.

Figure 4.9 shows the PDCP Data PDU format used for DRBs. In this figure, the length of the PDCP SN field is either 12 bits or 7 bits. The 12-bit SN is used for normal IP packet

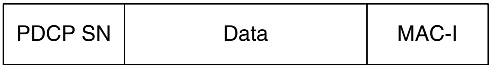

Figure 4.8 PDCP Data PDU format for SRBs. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

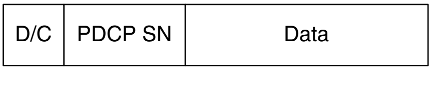

Figure 4.9 PDCP Data PDU format for DRBs. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

D/C	PDU Type	ROHC feedback

Figure 4.10 PDCP Control PDU format for ROHC feedback. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

D/C	PDU Type	FMS	BITMAP

Figure 4.11 PDCP Control PDU format for a PDCP status report. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

transmission, and thus can be used for both AM DRBs and UM DRBs. The 7-bit SN is an optimized length for VoIP transmission, and thus can only be used for UM DRBs.

4.7.2 PDCP Control PDU Formats

The PDCP Control PDU consists of a 1-bit "D/C" field, a 3-bit "PDU Type" field, and a variable length of control information. The "D/C" field is used to distinguish between PDCP Data PDUs and PDCP Control PDUs, and the "PDU Type" field is used to indicate what kind of control information is carried by the PDCP Control PDU. Currently, two kinds of control information are defined: ROHC feedback and PDCP status report.

The PDCP Control PDU format for ROHC feedback is shown in Figure 4.10. The ROHC feedback itself is generated by the ROHC protocol, but is carried by the PDCP Control PDU. Note that only one standalone ROHC feedback is carried by the PDCP Control PDU. As the ROHC protocol is used for both AM DRBs and UM DRBs, this PDCP Control PDU is also used for both DRBs.

The PDCP status report consists of a 12-bit "FMS" field, and a variable-length "BITMAP" field. "FMS" stands for "First Missing SN", indicating the PDCP SN of the first missing PDCP SDU. The "BITMAP" field indicates whether or not the PDCP SDU with PDCP N = (FMS + bit position) is missing in the PDCP receiver. If there are no missing PDCP SDUs, the "BITMAP" field is not included. Figure 4.11 shows the PDCP Control PDU format for a PDCP status report.

Reference

1. 3GPP Technical Specification 36.323, "Packet Data Convergence Protocol (PDCP) Specification (Release 10)", www.3gpp.org, accessed on 14 March 2012.

5

Radio Link Control (RLC)

3GPP provides the RLC specification in [1]. The RLC layer performs framing of RLC SDUs to put them into the size indicated by the lower MAC layer. The RLC transmitter segments and/or concatenates RLC SDUs to construct RLC PDUs, and the RLC receiver reassembles RLC PDUs to reconstruct RLC SDUs.

The RLC layer is connected to the lower MAC layer via logical channels. Each logical channel transports different types of traffic, and the name of the logical channel reflects the characteristic of the traffic it transports. In Release 8, a total of five types of logical channel are defined, as explained in Section 6.3.2.

The layer above the RLC layer is typically the PDCP layer, but in some cases it is the RRC layer. RRC messages transmitted on the logical channels PCCH, BCCH, and CCCH do not require security protection, and thus go directly to the RLC layer, bypassing the PDCP layer. Other than those RRC messages, all control plane and user plane traffic goes through the PDCP layer and RLC layer.

RLC functions are performed by the RLC entity. An RLC entity is established when a radio bearer is set up, and removed when the associated radio bearer is released. When established, the RLC entity is configured in one of three operating modes: Transparent Mode (TM), Unacknowledged Mode (UM), and Acknowledged Mode (AM). Depending on the operating mode, different RLC functions are performed by the RLC entity.

An RLC entity is used for only one radio bearer. However, a bidirectional radio bearer can have two RLC entities, in which case two UM RLC entities with different directions are used for the radio bearer.

5.1 RLC Functions and Architecture

This section gives an overview of RLC functions depending on the RLC operating mode. A detailed description of each RLC function is given in later sections.

5.1.1 Transparent Mode (TM) RLC

The TM RLC is an operating mode utilized only for the transport of special RRC messages that do not require security protection. RRC messages such as *Paging* messages transmitted

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

on the PCCH, system information broadcast messages transmitted on the BCCH, and SRB0 messages transmitted on the CCCH do not require security protection, so they bypass the PDCP layer and go directly to the TM RLC. Therefore, the upper layer of the TM RLC is the RRC layer. The radio bearers used for these RRC messages have fixed RB configurations, as explained in Section 3.1.

The TM RLC entity is virtually a null entity that does not perform any RLC function. RLC SDUs or RLC PDUs are transparently passing through the TM RLC entity, and thus an RLC SDU is mapped directly to an RLC PDU, and vice versa. This is the reason why the operating mode is called "Transparent Mode".

Since an RLC SDU is mapped directly to an RLC PDU, the RLC PDU header is not included in the RLC PDU. Due to the unavailability of the RLC PDU header, there are some requirements associated with use of the TM RLC. Firstly, an RLC PDU must contain one complete RRC message. As a result, different RLC PDUs are uncorrelated with each other, and this can avoid the need for reordering even if the RLC PDUs are received out of sequence. Secondly, duplicate reception of an RLC PDU must be handled properly without using RLC sequence numbers. For RRC messages on the PCCH and BCCH, because neither RLC retransmission nor MAC HARQ retransmission is performed, there is no duplicate reception. For RRC messages on the CCCH, because MAC HARQ retransmission can be performed, duplicate reception can occur. In this case, duplicate reception is detected by the HARQ process and the duplicate RLC PDU is discarded.

The TM RLC supports only unidirectional radio bearers. Therefore, a TM RLC entity is configured either as a transmitting TM RLC entity or a receiving TM RLC entity. In a transmitting TM RLC entity, a transmission buffer is used to store RLC SDUs. The stored RLC SDUs are delivered to the MAC layer only when a transmission opportunity is indicated by the MAC layer. The functional view of the TM RLC entity is shown in Figure 5.1.

5.1.2 Unacknowledged Mode (UM) RLC

The UM RLC is an operating mode optimized for delay-sensitive user traffic such as VoIP or audio/video streaming. The UM RLC is used only for user plane traffic, so the associated logical channel is only the DTCH. The upper layer of the UM RLC is the PDCP layer.

To support the delay-sensitive characteristics, retransmission is not performed in the UM RLC. The side effect of "no retransmission" is that RLC PDUs are subject to loss, but some loss of packets is usually acceptable to delay-sensitive traffic. Since retransmission is not

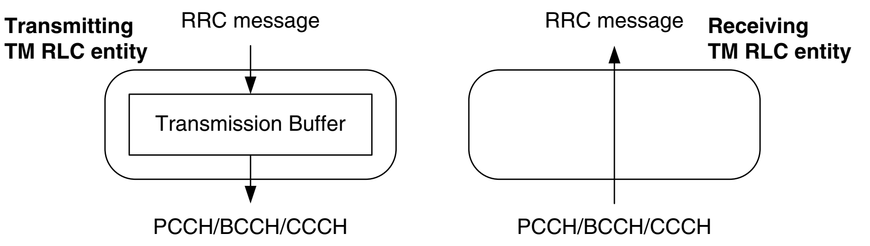

Figure 5.1 Functional view of the TM RLC entity. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

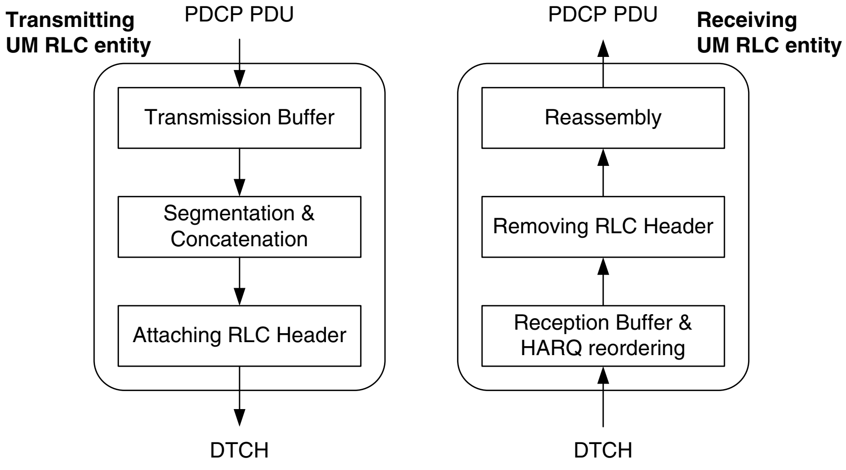

Figure 5.2 Functional view of the UM RLC entity. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

supported, feedback from the receiver is not needed either. The absence of feedback means the transmission result of the RLC PDU is not acknowledged, hence the name "Unacknowledged Mode".

The UM RLC entity performs a framing function to make RLC SDUs fit into RLC PDUs whose size is indicated by the lower MAC layer. The framing function comprises segmentation and concatenation in the RLC transmitter, and reassembly in the RLC receiver. To support the framing function, some information related to framing is included in the RLC PDU header.

The PDU header also contains an RLC sequence number that can be used for RLC PDU identification. Based on the RLC sequence number, the receiving UM RLC entity can perform reordering of the received RLC PDUs. The reordering function is required because RLC PDUs can be received out of sequence due to the use of multiple HARQ processes in the MAC layer. The HARQ reordering is performed in the RLC reception buffer, and it is realized by reordering timer and state variables.

Similar to the TM RLC, a UM RLC entity supports unidirectional radio bearers. Thus, a UM RLC entity can be configured either as a transmitting entity or as a receiving entity. However, the UM RLC entity can also support bidirectional radio bearers, in which case, one transmitting entity and one receiving entity are used. The functional view of the UM RLC entity is shown in Figure 5.2.

5.1.3 Acknowledged Mode (AM) RLC

The key feature of the AM RLC is "retransmission". The AM RLC guarantees lossless transmission with the aid of retransmission. Since the transmission is lossless, the AM RLC is utilized by most of the traffic in both control plane and user plane. In the control plane, all RRC messages other than those using TM RLC utilize the AM RLC. They are transmitted on the logical channel DCCH. In the user plane, the interactive/background type user traffic such as Web browsing and file transfer utilize the AM RLC. Audio/video streaming may also utilize the AM RLC if the delay requirement is not stringent. User plane traffic is transmitted on the logical channel DTCH. For both control plane and user plane, the upper layer of the AM RLC is the PDCP layer.

The AM RLC entity performs framing and HARQ reordering similar to the UM RLC entity. In addition, the AM RLC entity performs Automatic Repeat reQuest (ARQ) related functions such as polling, status reporting, retransmission, and re-segmentation. These functions are realized by transmission/reception windows and state variables.

The transmitter performs retransmission based on feedback from the receiver. The feedback is called an "RLC status report", which is transported by an RLC Control PDU. Therefore, in an AM RLC entity, two types of RLC PDU – RLC Data PDUs and RLC Control PDUs – are used. The two are differentiated by a 1-bit flag (the "D/C" field) in the RLC PDU header. For RLC Data PDUs, the RLC PDU header further includes ARQ-related information, for example, a polling bit and re-segmentation information, in addition to an RLC sequence number and framing-related information.

For transmission in AM RLC, RLC Control PDUs are prioritized over RLC Data PDUs, and, among the RLC Data PDUs, RLC Data PDUs for retransmission are prioritized over RLC Data PDUs for new transmission.

An AM RLC entity supports bidirectional radio bearers, and is configured with both a transmitting side and a receiving side. The functional view of the AM RLC entity is shown in Figure 5.3.

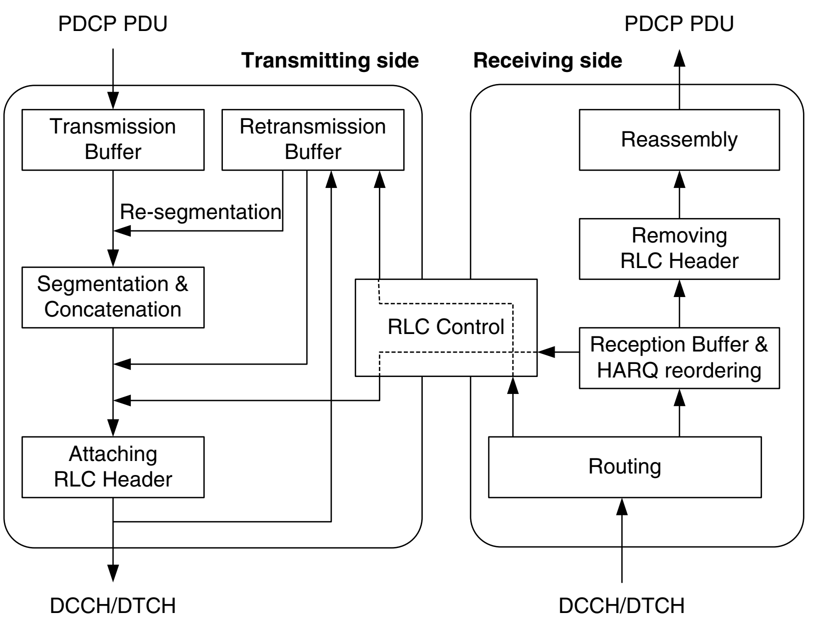

Figure 5.3 Functional view of the AM RLC entity. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

5.2 Framing

The RLC transmitter constructs RLC Data PDUs from RLC SDUs when a transmission opportunity is indicated by the MAC layer. For UM and AM RLC, the MAC layer also indicates the size of RLC resource allocated to the RLC transmitter. The RLC resource size is decided by the MAC layer on consideration of various aspects such as transmission resource, transmission power, radio condition, QoS of RBs, and so on. The RLC transmitter fills up the indicated size by constructing new RLC Data PDUs, possibly with an RLC Control PDU and RLC Data PDUs for retransmission in the AM RLC case. Typically, the RLC transmitter constructs only one new RLC Data PDU for each transmission opportunity.

When constructing an RLC Data PDU, the RLC transmitter performs segmentation and/or concatenation of SDUs in the order of arrival at the RLC transmitter. The payload part of an RLC Data PDU is composed of one or more SDU elements (i.e., RLC SDU or RLC SDU segment), each being byte-aligned. Segmentation is allowed only for the first and last SDU elements. In the RLC Data PDU, at least one SDU element is always included because padding is not allowed in the payload part. Figure 5.4 shows the format of the payload part of an RLC Data PDU.

Segmentation may be performed at the beginning and at the end of the payload part. To indicate whether segmentation was actually performed, a 2-bit indicator called "Framing Info (FI)" is included in the PDU header, as shown in Figure 5.5. Specifically, the first bit indicates whether the first byte of the payload part is the first byte of an SDU, and the second bit indicates whether the last byte of the payload part is the last byte of an SDU.

If more than one SDU element is concatenated in an RLC Data PDU, the RLC receiver should know the length of each element to parse the PDU. Therefore, for each SDU element, an 11-bit indicator called the "Length Indicator (LI)" is included in the PDU header to indicate the length in bytes of the corresponding SDU element, as shown in Figure 5.6. Since the length of the "Length Indicator" is 11 bits, the maximum size of an SDU element is 2048 bytes. However, this does not mean that the maximum size of an SDU is 2048 bytes, because an SDU larger than 2048 bytes can be segmented into multiple RLC Data PDUs. Note that

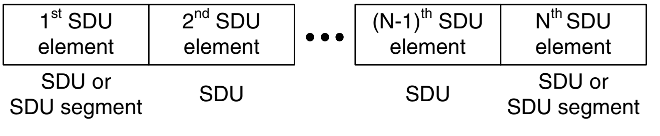

Figure 5.4 Payload part of an RLC Data PDU

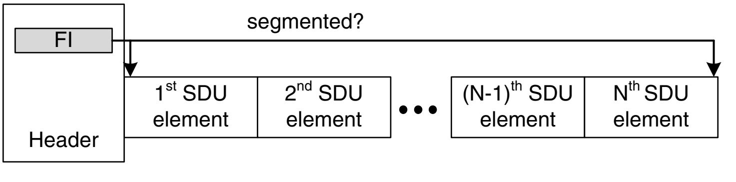

Figure 5.5 Framing Info

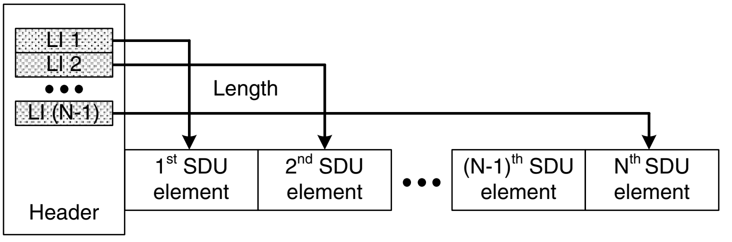

Figure 5.6 Length Indicator

the "Length Indicator" is not included for the last SDU element because the length of the last SDU element can be deduced by the RLC PDU size and other "Length Indicators". Consequently, there is no "Length Indicator" in the PDU header if only one SDU element is included in the PDU.

In the RLC receiver, the reassembly function is performed to reconstruct SDUs using the "Framing Info" and "Length Indicator". Only SDUs for which all segments are available are reassembled from RLC Data PDUs. If an SDU has any missing segments, it is either discarded (in UM RLC) or a request is sent for retransmission (in AM RLC). The reassembled SDUs, with the exception of the discarded SDUs in UM RLC, are delivered to the upper layer in ascending order of RLC sequence number.

5.3 Reordering

Reordering is performed in UM and AM RLC to rearrange the processing order of RLC Data PDUs in sequential order if they are received out of sequence. The out-of-sequence reception of RLC Data PDUs is likely to occur due to the multiple HARQ processes running in parallel with "Stop-And-Wait" operation in the MAC layer (see Section 6.8). Since out-of-sequence reception is caused by HARQ operation, this function is also called "HARQ reordering".

The detection of out-of-sequence RLC Data PDUs is based on the RLC Sequence Number (SN) included in the PDU header. The RLC SN uniquely identifies the corresponding RLC Data PDU, with a length of 5 or 10 bits for UM RLC and 10 bits for AM RLC. The RLC receiver detects out-of-sequence PDUs by comparing the RLC SN of the last in-sequence PDU and that of the received PDU. If there is a gap between the two, the RLC receiver decides that the received RLC Data PDU is out of sequence, and stores it in the reception buffer at the position indicated by the RLC SN. The stored PDU is processed only after all the previous missing PDUs are received and processed.

Sometimes HARQ transmission may not succeed for RLC Data PDUs. For example, there may be a HARQ failure – that is, transmission has not succeeded for the maximum number of HARQ retransmissions – or a NACK-to-ACK error – that is, the feedback of NACK by the HARQ process receiver is misinterpreted as ACK by the HARQ process transmitter. In this case, the RLC receiver will not be able to receive the RLC Data PDUs from the HARQ processes, and thus it does not need to wait for them any longer.

To detect transmission loss in a HARQ process and to limit the maximum wait time for missing PDUs, a reordering timer is used in the RLC receiver. Note that only one reordering timer is used in an RLC entity. A reordering timer is started when an out-of-sequence PDU is

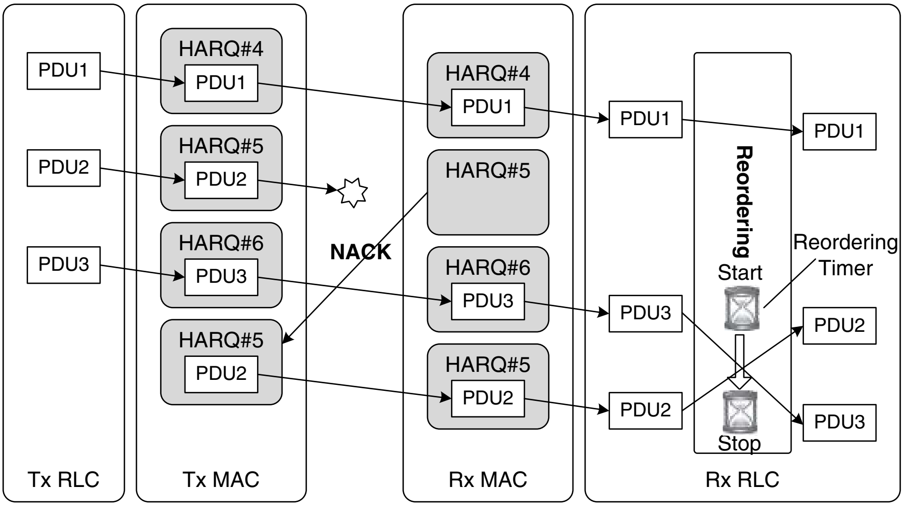

Figure 5.7 Out-of-sequence reception and HARQ reordering

received, and stopped when the out-of-sequence PDU becomes in sequence (i.e., all missing PDUs with lower RLC SNs than the out-of-sequence PDU are received). While the timer is running, the RLC receiver waits for the missing PDUs. If the missing PDUs are not received until the timer expires, the RLC receiver decides that transmission loss has occurred for the missing PDUs in the HARQ process and stops waiting for these PDUs. HARQ reordering for the out-of-sequence PDU is considered complete when the reordering timer for the out-of-sequence PDU stops or expires.

An example of out-of-sequence reception and HARQ reordering is shown in Figure 5.7. In this figure, the PDU3 (the PDU with RLC SN = 3) is received out of sequence because the HARQ transmission of PDU2 (the PDU with RLC SN = 2) failed at the initial attempt while that of PDU3 succeeded. For the out-of-sequence PDU3, the RLC receiver starts a reordering timer and stores PDU3 in the RLC reception buffer. Then, when the RLC receiver receives PDU2, it stops the reordering timer and processes PDU2 while keeping PDU3 in the reception buffer. Only after processing PDU2 is the RLC receiver allowed to process PDU3.

Upon the expiry of the reordering timer, the RLC receiver stops waiting for missing PDUs and starts performing subsequent behaviors; either reassembling RLC SDUs from the stored RLC Data PDUs (for UM RLC) or transmitting a status report to request retransmission of the missing PDUs (for AM RLC). The reassembly or retransmission request is performed for the PDUs with RLC SNs lower than that of the next missing PDU following the out-of-sequence PDU that triggered the reordering timer. Among these PDUs, the UM RLC receiver reassembles SDUs from the stored PDUs, and the AM RLC receiver requests retransmission of the missing PDUs. After that, if any PDUs remain in the reception buffer following the next missing PDU, the RLC receiver restarts the reordering timer for the PDU with the highest RLC SN. This is to ensure that similar reordering delay is applied to different PDUs and also that reordering delay is not accumulated.

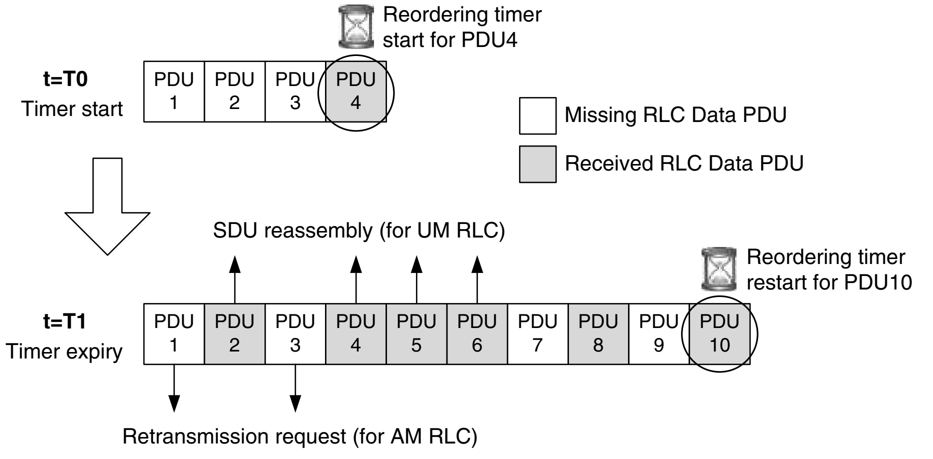

Figure 5.8 Example of the HARQ reordering operation

Figure 5.8 shows an example of the HARQ reordering operation. At $t = \mathrm{T0}$, the RLC receiver receives an out-of-sequence PDU4 and starts a reordering timer. While the timer is running, the RLC receiver receives more PDUs such as PDU2, PDU5, PDU6, PDU8, and PDU10, but PDU4 is still stored in the reception buffer because some of the previous PDUs – PDU1 and PDU3 – are still missing. At $t = \mathrm{T1}$, the reordering timer expires and the RLC receiver starts to perform reassembly or request retransmission. These actions are performed for the PDUs up to PDU7, because PDU7 is the next missing PDU following PDU4 that triggered the timer. Specifically, the UM RLC receiver reassembles SDUs from the stored PDUs of PDU2, PDU4, PDU5, and PDU6, and the AM RLC receiver requests retransmission for the missing PDUs of PDU1 and PDU3. Then, the reordering timer is restarted to perform HARQ reordering for the later PDUs. The reordering timer is restarted for PDU10, which is the PDU with the highest RLC SN among the PDUs stored in the reception buffer.

5.4 ARQ Operation

The Automatic Repeat reQuest (ARQ) operation is performed in AM RLC to guarantee lossless transmission. In basic ARQ operation, the transmitter retransmits packets if they were negatively acknowledged by the receiver. In addition to this, the AM RLC transmitter has the ability to request a status report from the receiver (called “Polling”), and to perform re-segmentation of the PDUs for retransmission. Each ARQ-related function is described below.

5.4.1 Polling

The Polling function is used by the RLC transmitter to obtain a status report from the RLC receiver. In the transmitter, transmission is allowed only for the PDUs within the transmission window, and the transmission window is updated only by the status report. Therefore, if the status report is delayed, the transmission window is stalled and the transmission may be stuck. Moreover, since unacknowledged PDUs are kept in the retransmission buffer until they are acknowledged, the buffer in the transmitter may overflow. To avoid such problems, the RLC transmitter needs to receive the status report in a timely manner. Thus, the following four triggers are defined in the RLC transmitter to trigger a poll:

- **Predefined number of PDUs has been transmitted:** A poll is triggered if the predefined number of RLC Data PDUs has been transmitted without a poll. The RLC transmitter maintains a PDU counter to count the number of newly transmitted PDUs. The PDU counter is reset to zero when a poll is transmitted by any of the polling triggers.
- **Predefined number of bytes has been transmitted:** A poll is triggered if RLC Data PDUs containing the predefined number of bytes has been transmitted without a poll. The RLC transmitter maintains a byte counter to count the number of newly transmitted bytes. The byte counter is reset to zero when a poll is transmitted by any of the polling triggers.
- **Last PDU in buffer is transmitted:** A poll is triggered if the RLC Data PDU to be transmitted is the last PDU in both the transmission and retransmission buffers (excluding transmitted RLC Data PDUs waiting for acknowledgment).
- **Poll retransmit timer expires:** A poll is triggered when the poll retransmit timer expires. The timer starts or restarts when a poll is transmitted by any of the polling triggers, and the timer stops when the RLC Data PDU for which the timer started is positively or negatively acknowledged by a status report.

When a poll is triggered, it is included in one of the RLC Data PDUs to be transmitted. One bit in the PDU header is allocated for polling purposes. This bit is called the “Polling” field, and the RLC transmitter sets the “Polling” field to “1” to transmit a poll. Since the “Polling” field is always included in the PDU header, there is no additional overhead to transmit a poll.

5.4.2 Status Reporting

The status report is used by the RLC receiver to inform the RLC transmitter of the reception buffer status. The status report provides positive acknowledgment (ACK) or negative acknowledgment (NACK) information on RLC Data PDUs or portions of them, up to the last RLC Data PDU whose HARQ reordering is complete. The RLC receiver triggers a status report when one of the following conditions is met after performing HARQ reordering of the relevant RLC Data PDUs:

- an RLC Data PDU is received with “Polling” field set to “1”;
- an RLC Data PDU is detected as missing.

The triggered status report needs to be transmitted to the RLC transmitter. However, if multiple status reports were triggered within a short time interval, it would not be radio efficient to transmit them all because they contain almost the same information. Moreover, a duplicated NACK of one RLC Data PDU would cause unnecessary retransmission and duplicate reception. Therefore, it is prohibited to transmit multiple status reports within a short time interval.

For this purpose, a “Status Prohibit” function is used in the RLC receiver. The prohibition interval is controlled by a timer called the “status prohibit timer”. The timer starts when a status report is transmitted and never stops until expiry. While the timer is running, the RLC receiver is not allowed to transmit further status reports. In addition, even if multiple status reports are triggered while the timer is running, only one status report is transmitted by expiry of the timer. In this way, the "Status Prohibit" function prevents frequent transmission of status reports. However, it should be noted that a long prohibition interval may cause window stalling in the RLC transmitter, so the length of the prohibition interval should be controlled carefully considering a trade-off between radio efficiency and transmission delay.

The RLC receiver generates a status report if transmission is not prohibited. The status report is transported by the STATUS PDU, as explained in Section 5.8. The status report includes the latest buffer status information regardless of when the status report is triggered. Since the transmission of a status report is controlled by the "Status Prohibit" function, it would be good for a status report to include as much information as possible. Normally, the status information of RLC Data PDUs up to the last RLC Data PDU whose HARQ reordering is complete is included in a status report. However, it cannot always be guaranteed that the RLC resource size indicated by MAC can cope with the required status report size. In this case, the status report can include only partial information – that is, the status information up to a certain RLC Data PDU whose RLC SN is lower than that of the last RLC Data PDU – to fit the RLC resource size.

5.4.3 Retransmission

When the RLC transmitter receives a status report from the RLC receiver, it prepares for retransmission of the RLC Data PDUs indicated as NACK in the status report. A simple rule is applied for retransmission: only RLC Data PDUs indicated as NACK can be retransmitted. In other words, the RLC transmitter is not allowed to retransmit RLC Data PDUs by its own will.

One exception to this rule is autonomous retransmission for polling purposes. If no RLC Data PDU is available for retransmission due to the lack of a status report, and, at the same time, if no new RLC Data PDU is available for transmission due to the fact that it is the end of the data stream or due to window stalling, the RLC transmitter has nothing to send, and thus there is no way to transmit a poll. In this case, a deadlock situation may occur if a triggering condition for a status report is not met in the RLC receiver. To prevent this deadlock situation, the RLC transmitter is allowed to perform autonomous retransmission if no RLC Data PDU is available for transmission. Specifically, the RLC transmitter picks one of the RLC Data PDUs that have not been acknowledged by the status report, and retransmits it with a poll in the PDU header. Autonomous retransmission also prevents serious delay at the end of a data stream or during window stalling.

When retransmitting an RLC Data PDU, the RLC transmitter may change the PDU header to set the "Polling" field. The "Polling" field is set according to the polling triggers described in Section 5.4.1. Other than this, the original PDU is retransmitted without any change unless it is re-segmented.

The maximum number of retransmissions of an RLC Data PDU is controlled by the RRC layer. There may be RLC protocol errors, for example, de-synchronized state variables, in peer AM RLC entities, in which case, in theory, the retransmission would continue to fail indefinitely. To rescue this protocol error case, a maximum number of retransmissions is defined in the AM RLC entity. Since the number of retransmissions has an impact on the QoS, the maximum number is configured per radio bearer depending on the radio bearer characteristics. When an RLC Data PDU is not transmitted successfully for the configured maximum number of retransmissions, the AM RLC entity indicates retransmission failure to the RRC, and then the RRC initiates the RRC Connection Re-establishment procedure.

5.4.4 Re-segmentation

Basically, the RLC transmitter retransmits the original PDU as it is, with the possible exception of a modification in the “Polling” field in the PDU. However, due to varying radio conditions, it is likely that the RLC resource size at retransmission will be smaller than the original PDU size. In this case, the RLC transmitter is allowed to re-segment the original PDU into smaller PDU segments in order to fit the available RLC resource size. The original PDU is called the Acknowledged Mode Data (AMD) PDU, and each segment of the original PDU is called an AMD PDU segment. Without re-segmentation, the RLC transmitter would have to skip the transmission opportunity, which would lead to radio resource wastage. Thus, re-segmentation helps utilize radio resource efficiently.

An example of re-segmentation is shown in Figure 5.9. PDU1 is constructed initially with a size of 900 bytes for transmission, but, as the radio condition becomes worse, it is re-segmented into two PDU segments of 400 and 500 bytes for retransmission. On the receiver side, all segments of a PDU are gathered together to reconstruct the original PDU.

For the RLC receiver to distinguish between AMD PDU segments, a 1-bit “Re-segmentation Flag (RF)” field is included in the PDU header. Moreover, for the RLC receiver to reconstruct the AMD PDU from the AMD PDU segments, it is necessary to indicate which portions of the original PDU are included in the PDU segment. This is achieved by a 15-bit “Segment Offset (SO)” field, which is included in the PDU header of the AMD PDU segment when the “Re-segmentation Flag” field indicates that the PDU is an AMD PDU

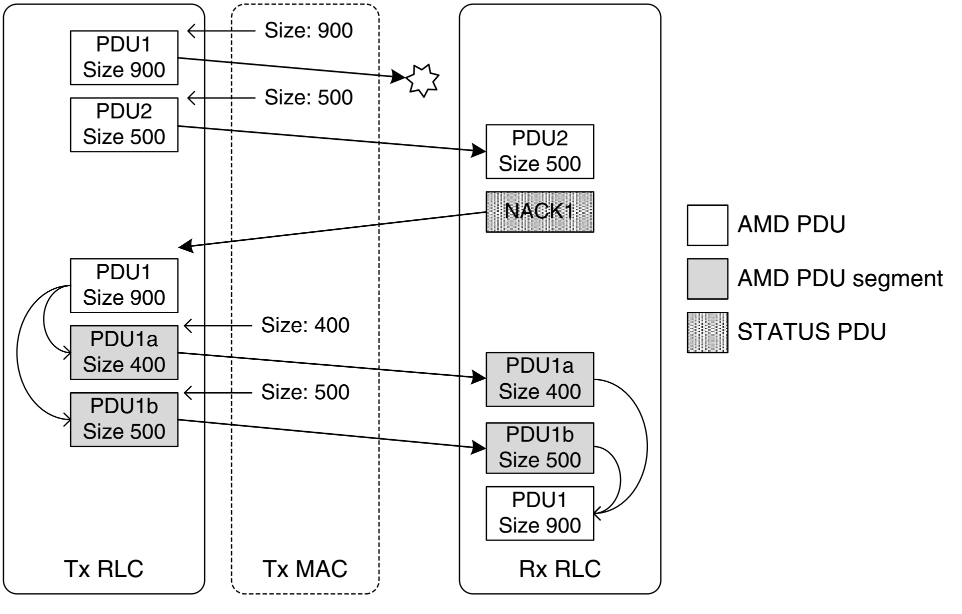

Figure 5.9 Re-segmentation at retransmission

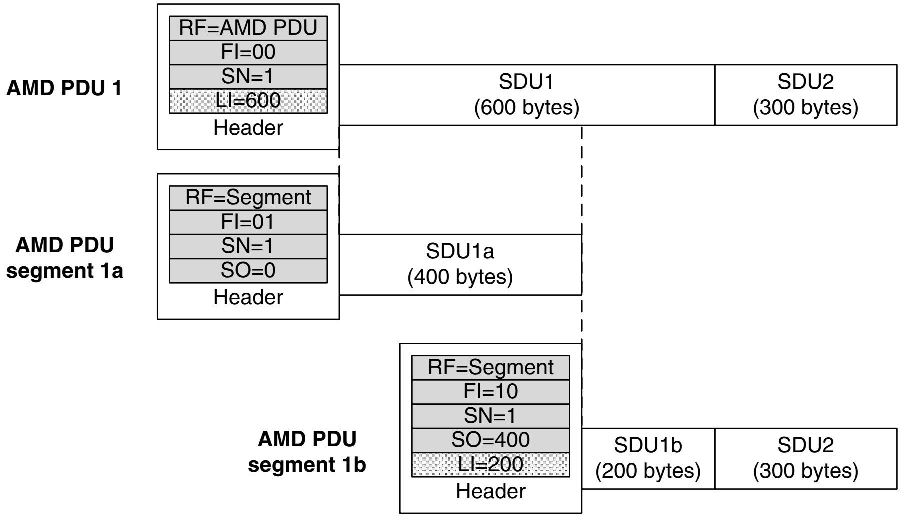

Figure 5.10 AMD PDU and AMD PDU segment

segment. The “Segment Offset” field indicates the start position of the AMD PDU segment in bytes within the original AMD PDU, considering only the payload part.

Taking PDU1 in Figure 5.9 as an example, the relationship between the AMD PDU and an AMD PDU segment is illustrated in Figure 5.10. As can be seen in the figure, the AMD PDU segment has a slightly larger overhead than the AMD PDU due to the inclusion of the “Segment Offset” field. Note that the “Framing Info” and “Length Indicator” fields in the AMD PDU segment are set according to the contents of the AMD PDU segment.

5.5 Window Operation

An RLC window is used in UM RLC and AM RLC for efficient protocol operation. The window is defined in terms of RLC SN, and the window size is configured as half of the SN space. That is, the window size is 16 (for 5-bit SN) or 512 (for 10-bit SN) for UM RLC, and 512 for AM RLC.

The purpose of the RLC window is different for UM RLC and AM RLC. Therefore, the window operation is also different depending on the RLC mode. In this section, the window operations for UM RLC and AM RLC are explained separately.

5.5.1 UM RLC Window Operation

The receiver of UM RLC maintains a reception window to identify whether the received RLC Data PDU is a new PDU or an old PDU. The received PDU is considered a new PDU if the SN is outside the reception window, and is considered an old PDU if the SN is within the reception window.

When a new PDU is received, the RLC receiver stores it in the reception buffer and updates the reception window so that the upper edge of the reception window is set equal to (1 + SN of the received new PDU). Since the reception window is updated every time a new

PDU is received, this UM window operation is called "Pull-based operation" (i.e., the PDU with the highest SN pulls the reception window).

If the received new PDU is in sequence to the last PDU delivered to the upper layer, the new PDU is delivered immediately to the upper layer. Otherwise, HARQ reordering is performed for the new PDU because it is received out of sequence. The out-of-sequence PDU is delivered to the upper layer when the HARQ reordering is complete (see Section 5.3).

When an old PDU is received, the RLC receiver does not update the reception window but performs duplicate detection. If the received old PDU is deemed to be a duplicate – that is, if a PDU with the same SN is already stored in the reception buffer or has been delivered to the upper layer – the RLC receiver discards the old PDU. Duplication is usually caused by ACK-to-NACK misinterpretation error in the HARQ process.

If the received old PDU is deemed not to be a duplicate, the RLC receiver stores the old PDU in the reception buffer. The stored old PDU is delivered to the upper layer when all previous PDUs have been delivered to the upper layer or when the old PDU falls outside the reception window due to the advancement of the window (by receipt of a new PDU). Note that the delivery of the old PDU does not update the reception window.

5.5.2 AM RLC Window Operation

The AM RLC maintains both transmission and reception windows to support ARQ operation. The RLC transmitter updates the transmission window when an in-sequence PDU – that is, the PDU with the lowest SN that equals the lower edge of the transmission window – is acknowledged by the RLC receiver. The RLC receiver updates the reception window when an in-sequence PDU – that is, the PDU with the lowest SN that equals the lower edge of the reception window – is received. The windows are updated so that the lower edge of the window is set equal to the SN of the next in-sequence PDU. Since the window is updated by an in-sequence PDU, the AM window operation is called "Push-based operation" (i.e., the PDU with the lowest SN pushes the transmission and reception windows).

The RLC transmitter is not allowed to transmit PDUs with SNs outside the transmission window. Without the acknowledgment of an in-sequence PDU from the RLC receiver, the RLC transmitter cannot update the transmission window, and thus the transmission will be stuck – that is, no new PDU can be transmitted. This case is called "window stalling" and it must be avoided for speedy data transmission.

Similar to the RLC transmitter, the RLC receiver is not allowed to receive PDUs with SNs outside the reception window. If the SN of the received PDU is outside the reception window, the received PDU is discarded by the RLC receiver. This case may happen when an ACK-to-NACK misinterpretation error occurs in the HARQ process for a PDU already received and delivered to the upper layer. If the SN of the received PDU is within the reception window, the PDU is stored in the reception buffer unless it is a duplicate, in which case the duplicated PDU is discarded. The stored PDU is delivered to the upper layer when the PDU falls in sequence – that is, all previous PDUs have been delivered to the upper layer.

5.6 SDU Discard

The RLC layer cannot trigger the discard of SDUs stored in the transmission buffer. Actually, SDU discard is triggered by the PDCP layer, as explained in Section 4.5. The RLC layer discards SDUs only when the PDCP layer provides a discard indication.

When a discard indication is provided for an RLC SDU, the RLC transmitter discards the indicated SDU if the SDU has not been transmitted. However, if any portion of the SDU has been formed into an RLC PDU, the RLC transmitter does not discard it but completes the transmission of the SDU. This means that the AM RLC transmitter keeps retransmitting the SDU until it is received correctly by the AM RLC receiver.

Note that the TM RLC transmitter does not discard SDUs in the transmission buffer because it is not connected to the PDCP layer.

5.7 RLC Re-establishment

During handover, the RRC layer requests the RLC layer and the PDCP layer to perform re-establishment. When performing re-establishment, the RLC layer resets state variables and timers to their initial values and also discards SDUs and PDUs in the transmission buffer. However, PDUs stored in the reception buffer are not discarded but reassembled into SDUs if possible. As for the normal reassembly process, only SDUs for which all segments are available are reassembled from the stored PDUs. Reassembled SDUs are then delivered to the upper layer in ascending order of RLC SN of the corresponding PDUs.

It should be noted that even the AM RLC delivers out-of-sequence SDUs at RLC re-establishment. In normal situations, the AM RLC does not deliver out-of-sequence SDUs to the upper layer. However, at RLC re-establishment, correctly received SDUs are delivered to the PDCP layer even if they are out of sequence, and selective retransmission is performed in the PDCP layer (see Section 4.6).

5.8 RLC PDU Formats

RLC PDUs are categorized into two types – RLC Data PDUs and RLC Control PDUs – depending on the contents they transport. User plane data received from the upper layer (i.e., RLC SDUs) are transported by RLC Data PDUs, and the control information generated in the RLC layer is transported by RLC Control PDUs. The RLC Data PDU is used by all RLC modes, but the RLC Control PDU is used by only AM RLC. The RLC Data PDU is further categorized into TMD PDUs, UMD PDUs, AMD PDUs, and AMD PDU segments. For the RLC Control PDU, only one type is defined – the STATUS PDU. The RLC PDU is a bit string that is byte-aligned (i.e., a multiple of 8 bits in length).

5.8.1 TMD PDU Format

A Transparent Mode Data (TMD) PDU consists of only one complete SDU. Since no framing functions are performed in the TM RLC, a PDU header is not needed for a TMD PDU. There is always a one-to-one relationship between an RLC SDU and a TMD PDU. The format of a TMD PDU is shown in Figure 5.11.

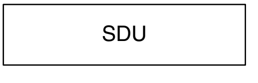

Figure 5.11 TMD PDU format. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

5.8.2 UMD PDU Format

An Unacknowledged Mode Data (UMD) PDU consists of a PDU header and one or more SDU elements; each SDU element being a complete SDU or an SDU segment. The PDU header is further categorized into a fixed header and an extension header. The fixed header consists of "Framing Info (FI)" and a "Sequence Number (SN)", and the extension header consists of one or more "Length Indicators (LIs)".

- **Framing Info (FI):** This 2-bit field indicates whether the RLC SDUs included at the beginning and at the end of the payload part are segmented or not. The first bit indicates whether the first byte of the payload part is the first byte of an SDU, and the second bit indicates whether the last byte of the payload part is the last byte of an SDU.
- **Sequence Number (SN):** This 5-bit or 10-bit field indicates the sequence number of the corresponding UMD PDU. The length of the SN field is configured by the RRC when the radio bearer is set up. The SN is incremented by one for every UMD PDU, and hence it is used to uniquely identify the UMD PDU.
- **Length Indicator (LI):** This 11-bit field indicates the length in bytes of the corresponding SDU element. An LI field is included for each SDU element except the last one. Length Indicators are included in the same order as the SDU elements present in the UMD PDU.

The extension header is present only when more than one SDU element is present in the UMD PDU. To make each of the fixed header and the extension header byte-aligned, a reserved bit or padding bit may be included in the PDU header. Note that they are not included in the payload part because each SDU element is byte-aligned. Figure 5.12 shows the format of the UMD PDU.

5.8.3 AMD PDU Format

As for the UMD PDU, the Acknowledged Mode Data (AMD) PDU consists of a PDU header and one or more SDU elements. The AMD PDU header still contains "Framing Info", a "Sequence Number", and a "Length Indicator" as in the UMD PDU header, the only difference being that only a 10-bit RLC SN is used for the AMD PDU. In addition to the above-mentioned fields, the AMD PDU header also contains a "Data/Control (D/C)" field, a "Re-segmentation Flag (RF)" field, and a "Polling (P)" field to support status reports and re-segmentation:

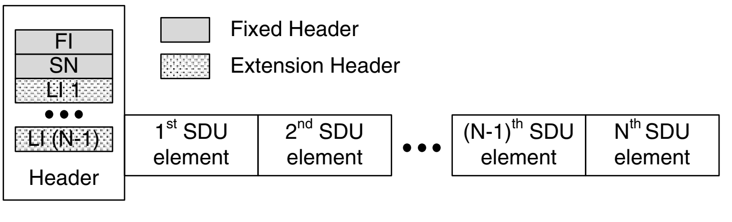

Figure 5.12 UMD PDU format. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

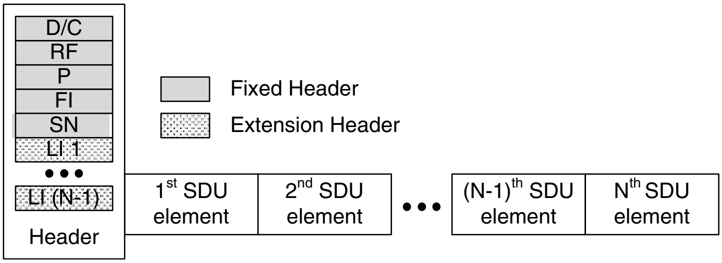

Figure 5.13 AMD PDU format. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

- **Data/Control (D/C):** This 1-bit field indicates whether the RLC PDU is an RLC Data PDU or an RLC Control PDU.
- **Re-segmentation Flag (RF):** This 1-bit field indicates whether the RLC PDU is an AMD PDU or an AMD PDU segment.
- **Polling (P):** This 1-bit field indicates whether the transmitting side of an AM RLC entity is requesting a status report from its peer receiving side of an AM RLC entity.

Figure 5.13 shows the format of the AMD PDU.

5.8.4 AMD PDU Segment Format

The AMD PDU segment is used when re-segmentation is performed at retransmission. If the "Re-segmentation Flag" field indicates that the RLC PDU is an AMD PDU segment, additional fields such as a "Last Segment Flag (LSF)" and a "Segmentation Offset (SO)" are included in the PDU header to provide re-segmentation-related information:

- **Last Segment Flag (LSF):** This 1-bit field indicates whether or not the payload part of this AMD PDU segment is the last segment of the payload part of the AMD PDU.
- **Segmentation Offset (SO):** This 15-bit field indicates the starting position of the payload part of the AMD PDU segment in bytes within the payload part of the original AMD PDU.

Figure 5.14 shows the format of the AMD PDU segment.

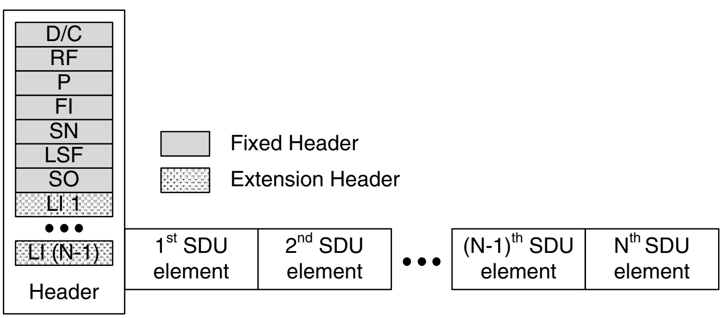

Figure 5.14 AMD PDU segment format. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

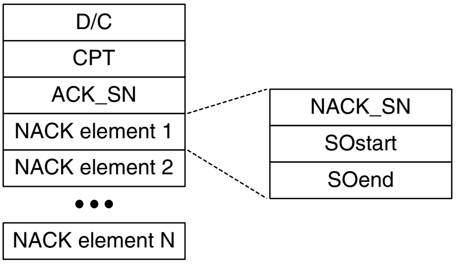

Figure 5.15 STATUS PDU format. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

5.8.5 STATUS PDU Format

The STATUS PDU is used in AM RLC to transport a status report from the RLC receiver to the RLC transmitter. The reception status – that is, ACK or NACK – of each AMD PDU is provided up to a certain AMD PDU for which HARQ reordering is complete. The STATUS PDU simply lists all the missing portions of AMD PDUs up to the certain AMD PDU with SN = ACK_SN.

- **Control PDU Type (CPT):** This 3-bit field indicates the type of RLC Control PDU. Currently, only STATUS PDU is defined for the RLC Control PDU.
- **Acknowledgment SN (ACK_SN):** This 10-bit field indicates the SN of the first AMD PDU whose reception status is not reported by the STATUS PDU. The RLC transmitter should interpret this to mean that all AMD PDUs up to but not including the AMD PDU with SN = ACK_SN have been received correctly by the RLC receiver, excluding the AMD PDUs or portions of them listed in the NACK element.
- **Negative Acknowledgment SN (NACK_SN):** This 10-bit field indicates the SN of the AMD PDU whose portions have not been received by the RLC receiver. If a portion of the AMD PDU is missing, "SOstart" and "SOend" fields follow the "NACK_SN" field in order to indicate the exact position of the missing portion in the AMD PDU. Otherwise – that is, if the whole AMD PDU is missing – only the "NACK_SN" field is included.
- **Segmentation Offset Start (SOstart):** This 15-bit field indicates the start position of the missing portion of the AMD PDU with SN = NACK_SN.
- **Segmentation Offset End (SOend):** This 15-bit field indicates the end position of the missing portion of the AMD PDU with SN = NACK_SN. A special value, all "1"s, indicates that all bytes from the "SOstart" to the last byte of the AMD PDU are missing.

Figure 5.15 shows the format of the STATUS PDU.

Reference

1. 3GPP Technical Specification 36.322, "Radio Link Control (RLC) Protocol Specification (Release 10)", www.3gpp.org, accessed on 14 March 2012.

6

Medium Access Control (MAC)

3GPP provides the MAC specification in [1]. As the name suggests, the Medium Access Control (MAC) layer controls the upper layers' access to the communication medium, which is the PHY layer. The MAC layer is connected to the PHY layer below through transport channels, and the MAC layer is connected to the RLC layer above through logical channels. Thus, the MAC layer decides which logical channels can access the transport channels at a given time, and performs multiplexing and de-multiplexing of the data between them.

The MAC layer basically provides the radio resource allocation service and the data transfer service to the upper layer. As part of the radio resource allocation service the MAC layer performs procedures such as logical channel prioritization, power headroom reporting, the handling of UL grant and DL assignment, and so on. As part of the data transfer service, the MAC layer performs procedures such as scheduling requests, buffer status reporting, random access, and HARQ.

In addition, the MAC layer handles the Semi-Persistent Scheduling (SPS) procedure and the Discontinuous Reception (DRX) procedure. The SPS procedure is used to increase the cell capacity for a voice service, and the DRX procedure is used to reduce the power consumption of the UE.

6.1 MAC Functions and Services

As the layer in charge of controlling radio resource, connecting the RLC layer and the PHY layer, and transferring data, the following functions are supported by the MAC layer:

- Connecting the upper layer with the lower layer:
  - mapping between the logical channels and the transport channels;
  - multiplexing of the MAC SDUs from one or different logical channels onto the transport blocks (TBs) to be delivered to the PHY layer on the transport channels;
  - de-multiplexing of the MAC SDUs of one or different logical channels from the transport blocks (TBs) delivered from the PHY layer on the transport channels.
- Transferring data:
  - scheduling information reporting;
  - error correction through HARQ.

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

Table 6.1 MAC functions by UE and eNB. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

MAC Functions	UE	eNB
Mapping between logical channels and transport channels	X	X
Multiplexing for uplink	X	
Multiplexing for downlink		X
De-multiplexing for downlink	X	
De-multiplexing for uplink		X
Error correction through HARQ	X	X
Transport format selection		X
Priority handling between UEs		X
Priority handling between logical channels of one UE		X
Logical channel prioritization	X	
Scheduling information reporting	X	

- Controlling radio resource:
  - priority handling between UEs by means of dynamic scheduling;
  - priority handling between the logical channels of one UE;
  - transport format selection.

Among the functions listed above, some functions are performed only in the eNB and some functions are performed only in the UE. Table 6.1 shows which functions are performed by which entities. The uplink direction is from the UE to the eNB and the downlink direction is from the eNB to the UE.

6.2 MAC Architecture

As shown in Figure 6.1, the MAC layer is composed of a Hybrid Automatic Repeat reQuest (HARQ) entity, a multiplexing/de-multiplexing entity, a logical channel prioritization entity, and a control entity.

The multiplexing and de-multiplexing entity is in charge of composing and decomposing the MAC PDUs and performs (de-)multiplexing of data from several logical channels into/from one transport channel. When the radio resources for a new transmission are allocated, the logical channel prioritization entity instructs the multiplexing and de-multiplexing entity to generate MAC PDUs from the MAC SDUs. When the MAC PDUs have been received, the multiplexing and de-multiplexing entity reassembles the MAC SDUs from the received MAC PDUs and delivers the recovered MAC SDUs to the appropriate RLC entities.

The HARQ entity performs the transmit HARQ operation and the receive HARQ operation. The transmit HARQ operation includes transmission of transport blocks and, if necessary, retransmission of the transport blocks, and, if configured, reception and processing of HARQ ACK/NACK signaling. The receive HARQ operation includes reception of transport

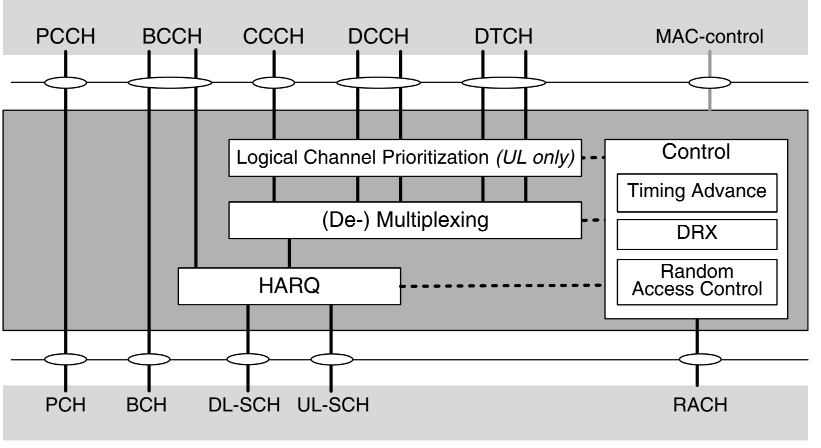

Figure 6.1 MAC architecture. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

blocks, combining and decoding of the received transport blocks and, if configured, generation of HARQ ACK/NACK signaling.

The logical channel prioritization entity decides how much data from each configured logical channel should be included in each MAC PDU whenever radio resource for a new transmission is available. As stated above, this decision is delivered to the multiplexing and de-multiplexing entity.

The control entity is responsible for a number of functions including DRX, resource requests, alignment of the uplink timing, power headroom reporting, and so on. For peer-to-peer communication between the eNB and the UE, control messages called MAC Control Elements (MAC CEs) are exchanged through the MAC PDUs. The Random Access (RA) procedure is used to request uplink radio resource to transmit uplink data when the UE does not have any dedicated uplink transmission resource, or to respond to the network's request.

6.3 MAC Channels and Mapping

The MAC layer exchanges data with the PHY layer through transport channels and with the RLC layer through logical channels. The transport channel used for data transfer is selected according to how it is transmitted over the radio interface. The logical channel used for data transfer is selected according to the characteristics of the data.

Transport channels are classified primarily as the downlink channel or the uplink channel, and logical channels are classified according to whether they transport user traffic data or control information such as RRC messages.

6.3.1 Transport Channels

There are three types of downlink transport channel:

- **Broadcast Channel (BCH):** This transport channel is used to transport the parts of the system information that are essential for the UE to operate properly within the network.

The transport format for this transport channel is fixed, and the amount of information that can be delivered over this channel is limited.

- **Downlink Shared Channel (DL-SCH):** This transport channel is used to transport user data or control information in the downlink, and the radio resources allocated for this channel can be changed dynamically. In addition to user data and control information, this transport channel transports the remaining system information that is not transported via the BCH.
- **Paging Channel (PCH):** This channel is used to transport paging information to UEs to inform them of an incoming call. This channel is also used to inform UEs about updates to system information and updates to Public Warning System (PWS) messages.

There are two types of uplink transport channel:

- **Uplink Shared Channel (UL-SCH):** This transport channel is used to transport user data or control information in the uplink. The radio resources allocated for this channel can be changed dynamically.
- **Random Access Channel (RACH):** This transport channel is used to transmit the random access preamble when the UE does not have any allocated uplink transmission resource for the transmission of available data, or when the UE is ordered to perform the Random Access procedure by the eNB. Information about the random access preamble is transported over this transport channel.

6.3.2 Logical Channels

There are four types of control logical channel:

- **Broadcast Control Channel (BCCH):** This logical channel is used in the downlink to broadcast system information and PWS messages. This logical channel is connected to a Transparent Mode (TM) RLC entity.
- **Paging Control Channel (PCCH):** This logical channel is used in the downlink to notify the UE of an incoming call or a change in system information. This logical channel is connected to a TM RLC entity.
- **Common Control Channel (CCCH):** This logical channel is used to deliver control information both in the uplink and in the downlink when the sender or the recipient of the information cannot be identified uniquely in layers lower than the RRC layer. This logical channel is connected to a TM RLC entity.
- **Dedicated Control Channel (DCCH):** This logical channel is used to deliver dedicated control information in both the uplink and downlink for a specific UE which has an RRC connection with the eNB. This logical channel is connected to an Acknowledged Mode (AM) RLC entity.

There is only one type of traffic logical channel:

- **Dedicated Traffic Channel (DTCH):** This logical channel is used to transmit dedicated user data in both the uplink and downlink. This logical channel is connected either to an Unacknowledged Mode (UM) RLC entity or an AM RLC entity.

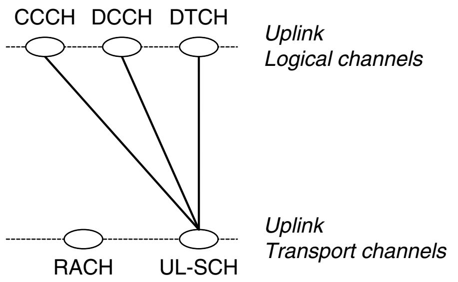

Figure 6.2 Mapping between UL channels. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

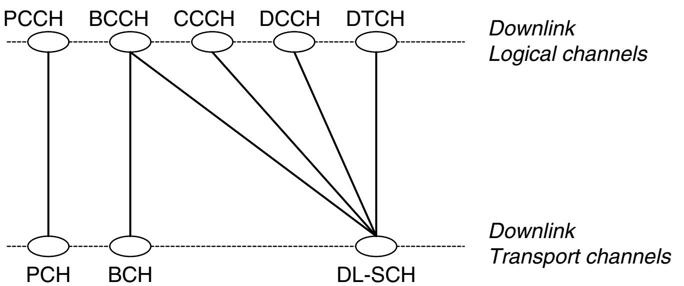

Figure 6.3 Mapping between DL channels. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

6.3.3 Channel Mapping

The MAC layer is in charge of connecting the transport channels and the logical channels. Among all possible combinations between the logical channels and the transport channels, only a few selected combinations are allowed. The allowed mapping between a transport channel and a logical channel is shown in Figures 6.2 and 6.3 for the uplink and the downlink respectively.

6.4 Scheduling

To use the limited radio resource efficiently among multiple UEs, radio resources should be managed carefully so that the radio resources are allocated to UEs only when necessary. Thus, in LTE where Orthogonal Frequency Division Multiple Access (OFDMA) is used, the radio resources are shared dynamically among multiple UEs. Radio resource allocation among the UEs and among the radio bearers is decided and coordinated by the eNB. To make an optimal allocation, the eNB should be aware of accurate information regarding the data generated in the UEs and the data to be delivered to the UEs.

The dynamic scheduling method mentioned above requires transmission of resource allocation information whenever data are transmitted between the eNB and the UE. For services such as Voice over IP (VoIP), where packets of small size are periodically generated, dynamic scheduling generates a lot of signaling overhead. To reduce this overhead, Semi-Persistent Scheduling (SPS), where the radio resource can be semi-statically allocated to UEs, is also used in LTE to increase the voice service capacity in a cell.

Finally, to satisfy different users' different needs, the eNB considers the Quality of Service (QoS) requirements of each configured radio bearer to make a decision on the radio resource allocation.

6.4.1 Dynamic Scheduling

In this mode of scheduling, the eNB allocates to each UE uplink or downlink radio resources that are valid for only one or several subframes. This scheduling mode is suitable for services that generate bursty data. For example, for a Web-browsing service, traffic is generated when the user makes a click on a Web link; the time when the user makes a click is not predictable and is intermittent in nature. Thus, for this service, the eNB allocates radio resource dynamically and adaptively when there is actually traffic. When the eNB makes a decision on downlink radio resource allocation to the UE, the eNB may consider the characteristics and amount of downlink data buffered in the eNB. For uplink radio resource allocation, the eNB may consider the information in the Buffer Status Reports (BSRs) received from the UE.

In dynamic scheduling, the eNB allocates downlink radio resources to the UE using resource allocation information called the *downlink assignment*, and allocates uplink radio resources to the UE using resource allocation information called the *uplink grant*. The downlink assignment and the uplink grant are signaled separately to the UE through the Physical Downlink Control CHannel (PDCCH), masked with the UE's C-RNTI. The downlink assignment indicates the HARQ information as well as information about the allocated radio resources. On the other hand, the uplink grant indicates information about the radio resources only (see Section 6.8).

6.4.2 Semi-Persistent Scheduling (SPS)

As mentioned above, a VoIP service periodically generates many small-sized packets at short intervals. If dynamic scheduling is used for a VoIP service, a lot of the downlink assignment and the uplink grant will also be used, leading to a heavy load on the PDCCH. Because the capacity of the PDCCH is limited, reducing the number of the downlink assignment and the uplink grant is required to increase the number of supportable simultaneous VoIP calls in a cell. Thus, SPS is used to allocate radio resources for a long time period with a minimized load on the PDCCH. Typically, SPS is used for predictable services where similar-sized data packets are generated periodically.

For example, when a UE is in a voice call, the voice packets are generated periodically and the generated packets have a similar or the same size. This means that similar radio resources are likely to be allocated periodically to the UE to transfer the generated data. To allow the UE to use these radio resources with the minimal amount of signaling through the PDCCH, the UE first needs to know the periodicity of the radio resource allocation and the actual transmission attributes, such as coding information or frequency information of the allocated radio resources. Delivery of the information to the UE is made using both RRC layer signaling and PHY layer signaling. Through RRC layer signaling using RRC messages, the UE is informed of the interval with which the radio resources are periodically assigned. Through

PHY layer signaling using the PDCCH, frequency information and transmission attribute information such as the modulation and coding scheme are delivered. In addition, the time reference from which the interval is applied should also be known to the UE. The UE's actual reception timing of PHY layer signaling through the PDCCH is used as the reference time.

Voice packets of a VoIP service are not the only available packets while a VoIP service is ongoing. For example, for connection management purposes, the RRC layer may generate RRC messages which are long in size and unpredictable in timing. To transport these unpredictable packets, the dynamic scheduling method can be used at any time regardless of SPS. In this case, because the resource allocation information for both SPS and dynamic scheduling is delivered through the PDCCH, there is a need to distinguish the resource allocation information for SPS and the resource allocation information for dynamic scheduling. For this purpose, in addition to the C-RNTI used for dynamic scheduling, a Semi-Persistent-Scheduling C-RNTI (SPS-C-RNTI) is used to mask resource allocation information over the PDCCH when the resource allocation information is for SPS.

When the resource allocation information masked by the C-RNTI is received at the subframe when the radio resources for SPS are configured, the UE uses the radio resource indicated by dynamic scheduling at that subframe and does not use the radio resource configured by SPS. This operation is called overriding the SPS resource, where the radio resource for SPS is replaced temporarily by the radio resources allocated by dynamic scheduling.

Figure 6.4 illustrates an example of the overriding of SPS by dynamic scheduling. In this example, the interval of the resource allocation for SPS is 20 ms. At 0 ms, the configuration of SPS occurs and resource A is allocated for SPS. Accordingly, the UE uses resource A with a 20 ms interval. However, if the eNB wants to use radio resource other than resource A at 40 ms, the eNB transmits the DL assignment using dynamic scheduling. Thus, at 40 ms, when the UE receives the DL assignment masked by the C-RNTI, the UE uses resource B indicated by the DL assignment for the subframe. In this case, the UE does not use resource A at 40 ms. At 60 ms, because there is no DL assignment, the UE continues to use resource A configured by SPS.

In essence, the resource allocation information over the PDCCH masked by SPS-C-RNTI indicates:

- the activation of radio resources for SPS;
- the reactivation of radio resources for SPS;

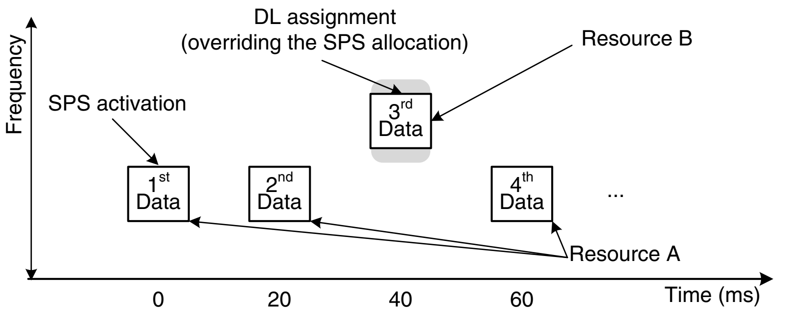

Figure 6.4 Example of SPS overriding

- HARQ retransmission of transport blocks which were originally transmitted over the radio resources for SPS;
- the release of radio resources for SPS.

Among the above four cases, in order to enable the UE to differentiate between the activation of radio resources for SPS and HARQ retransmission for SPS, the New Data Indicator (NDI) field, delivered through the resource allocation information over the PDCCH, is examined.

For activation of radio resources for SPS, the NDI field is set to 0. When the UE receives resource allocation information masked with the SPS-C-RNTI with the NDI field set to 0, the UE stores the received resource allocation information and periodically uses the radio resources indicated by the stored information for the reception or transmission of MAC PDUs.

On the other hand, if the NDI field is set to 1, the resource allocation is for HARQ retransmission for SPS. When the UE receives resource allocation information masked with the SPS-C-RNTI with the NDI field set to 1, it does not use the stored radio resource, but uses the received radio resource for the reception or transmission of MAC PDUs.

The reactivation of radio resources for SPS can be performed when the characteristics of the concerned service change. For example, when the voice codec rate for a VoIP service is increased while SPS is used, if the configured radio resources for SPS remain the same, the newly generated voice packets will not fit into the configured radio resources. To handle this situation, it is necessary to modify the radio resource configuration for SPS so that larger voice packets fit into the allocated radio resource. The reactivation of radio resources for SPS is performed in the same way as the activation of radio resources for SPS; that is, when the UE receives the resource allocation information masked with the SPS-C-RNTI with the NDI field set to 0 over the PDCCH, the UE replaces the stored resource allocation information with the newly received resource allocation information.

For radio resource efficiency, when the allocated semi-persistent radio resources are no longer used, the radio resource for SPS should be released. For example, in the case of a VoIP service, people typically do not talk all the time during the conversation, and the time duration when people do not talk, called the *silent period*, is relatively long. If the configured radio resources for SPS remain the same after transition from the talk spurt period to the silent period, the radio resource for SPS will be wasted due to a lack of data to send. In other words, when there is not so much gain from using SPS, it may be better to release the configured radio resources for SPS. To do so, the resource allocation information masked by the SPS-C-RNTI can include an indication to release radio resource for SPS and this is called *explicit SPS release*. When the UE receives resource allocation information indicating release of the radio resource for SPS, it discards the stored resource allocation information for SPS and stops using the configured radio resources.

When explicit SPS release is used, the eNB needs to make sure that the release command is received correctly by the UE. Because the release command is delivered through the radio interface, the release command can be lost over the radio interface. If the eNB releases the downlink radio resources for SPS when the UE has missed the release command, the mismatch between the eNB and the UE may lead to service quality degradation. To prevent this problem, when the UE receives resource allocation information with an indication to release downlink radio resource for SPS, it has to send a HARQ ACK to the eNB.

In the uplink, as a solution to the problem of loss of the explicit SPS release command, another mechanism is adopted. In fact, the eNB cannot immediately be aware of the need to release the uplink radio resources for SPS. Thus, an implicit SPS release mechanism to release the uplink radio resources for SPS is also adopted in LTE. In this implicit SPS release, when a certain number of MAC PDUs is transmitted consecutively without including any MAC SDU using the radio resource for SPS, the UE releases the uplink radio resources for SPS without informing the eNB of this event. Likewise, by receiving consecutively a certain number of MAC PDUs without any MAC SDU, the eNB can become aware that there is no more need to allocate uplink radio resources for SPS and eventually release the uplink radio resources for SPS.

6.5 Scheduling Information Delivery

6.5.1 Buffer Status Reporting (BSR)

In LTE, the allocated radio resources for the different UEs are orthogonal in time and frequency. Therefore, if the uplink radio resource is allocated by the eNB to a specific UE but is not used by that UE, the allocated radio resource is just wasted because other UEs cannot use it. Thus, to maximize the use of radio resource, the eNB takes full control of it. When a UE wants to transmit some data in the uplink, it needs to request allocation of radio resource from the eNB, except when it transmits the random access preamble.

To assist the eNB's optimal allocation of uplink radio resources to a UE, a Buffer Status Report (BSR) is transmitted from the UE to the eNB. The BSR indicates how much data are buffered in the UE's memory. The BSR is only meaningful in the uplink because for the downlink the eNB knows the buffer status due to the collocation of the eNB's scheduler and the downlink buffer. Thus, for uplink radio resource allocation, a UE needs to send the BSR to the eNB to indicate the amount of data in the UE that need to be transmitted.

Two formats are defined for the BSR: Long BSR and Short BSR. As can be inferred from the names, the Long BSR is longer than the Short BSR and is used to deliver buffer status information for the four Logical Channel Groups (LCGs). The Short BSR is used to deliver buffer status information for only one LCG. Which BSR format is used depends on what triggers the BSR, how many LCGs have data to send, and how much space is available in the MAC PDU.

The above states that the amount of data is indicated not per logical channel but per LCG. Actually, if the UE reported the BSR for each logical channel, the eNB could have more accurate and detailed knowledge about the UE's buffer status. However, since it is obvious that the BSR itself is the signaling overhead, it is preferable to make the size of the BSR as small as possible. In addition, even if many logical channels are configured with the UE, some of them may have similar QoS requirements or similar characteristics. If so, it does not bring much gain to the eNB's scheduler for the UE to report the buffer status for each logical channel. Thus, by grouping logical channels with similar attributes into at most four LCGs and by making the UE report the buffer status per LCG, LTE strikes a good balance between reporting efficiency and reporting accuracy.

A BSR can be triggered in any of the following situations:

- When data arrive for a logical channel which has higher priority than the logical channels whose buffers are not empty. In this case, the triggered BSR is called a Regular BSR. In terms of the QoS requirements, it is sensible to serve data with higher priority ahead of data with lower priority. For example, when an urgent RRC message arrives at the UE's buffer while the UE's buffer is filled with low-priority data such as Web-browsing traffic, the UE should report the arrival of higher priority data to the eNB immediately, regardless of previously transmitted BSRs. Otherwise, the connection between the UE and the eNB may be lost because the eNB may decide not to give any radio resources to the UE.

- When data become available for the UE's buffer, which is empty. This is also called a Regular BSR. If the UE does not inform the eNB of the arrival of new data, there is no way that the eNB can decide to allocate radio resource to the UE.
- When the *retxBSR-Timer* expires and there is still data in the UE's buffer. This is also called a Regular BSR. The *retxBSR-Timer* is used to avoid the stall situation that occurs when the UE thinks it has sent a BSR to the eNB but the eNB has not received the BSR. Because the triggered BSR is included in the MAC PDU as a MAC Control Element and this MAC PDU is transmitted using the HARQ mechanism, this deadlock situation may occur if the MAC PDU containing the Regular BSR is lost over the radio interface, and the HARQ NACK transmitted by the eNB is wrongly interpreted as a HARQ ACK by the UE (HARQ NACK-to-ACK error). To avoid this deadlock situation, the *retxBSR-Timer* is used. The timer is started when the MAC PDU including the BSR is transmitted, and restarted whenever the UE receives the uplink grant for a new transmission. When the *retxBSR-Timer* expires, it means that the UE has not received any uplink grant for some time. In this case, if a new Regular BSR is not triggered, the UE cannot send any more BSRs and there is no way for the UE to inform the eNB of its buffer status. Thus, to prevent this situation, the UE triggers a BSR if there is actually data in the UE buffer when the *retxBSR-Timer* expires.
- When a *periodicBSR-Timer* expires. This BSR is called a Periodic BSR. This is used for the UE periodically to deliver updated buffer status information to the eNB. For example, let's assume that the data belonging to only one logical channel arrive continuously at the UE's buffer and the UE is in the middle of transmission of the data. In this case, the first BSR is triggered when the first data for the logical channel are available. However, provided that the data belonging to other logical channels do not arrive at the UE's buffer, no BSR listed above will be triggered. As time goes by, more data may arrive for the logical channel and some data for the logical channel may be discarded. Thus, to prevent the discarding of data in a logical channel, a BSR is transmitted periodically using the *periodicBSR-Timer*.
- When the remaining space in a MAC PDU can accommodate a BSR. This is called a Padding BSR. When the UE receives an uplink grant, the UE knows the size of the MAC PDU it should compose. Sometimes, the MAC layer at the UE cannot entirely fill a MAC PDU of that size with data from the upper layers and the MAC CEs. For example, when an additional MAC SDU is composed at the RLC layer, the minimum size of the RLC header is 1 byte and the minimum size of RLC payload is 1 byte (see Section 5.2). In addition, to include this extra MAC SDU in the MAC PDU, at least 1 byte of MAC subheader is needed. Therefore, when the available space in a MAC PDU is less than 3 bytes, the space is unusable for the transport of user data. From an efficiency point of view, it is better to include meaningful data such as a BSR instead of padding bits. Thus, a BSR is included in the MAC PDU when the padding space is enough to include any format of BSR.

If a Regular BSR is triggered while the uplink radio resource for the new transmission is available, the BSR is included in a newly composed MAC PDU and sent to the eNB. On the other hand, if a UE does not have the uplink radio resource for a new transmission when a Regular BSR is triggered, a Scheduling Request (SR) procedure is triggered if the PUCCH for SR transmission is configured for the UE. However, if the UE is neither configured with the SR resource on the PUCCH nor allocated with the uplink grant, the UE starts the Random Access (RA) procedure when a Regular BSR is triggered. As a result of the RA procedure or transmission of the SR through the PUCCH, the uplink radio resource will be allocated to the UE and the MAC PDU containing the BSR will be transmitted to the eNB. Based on this BSR, the eNB will further allocate uplink radio resource to the UE.

However, when a BSR other than a Regular BSR is triggered, SR transmission over the PUCCH or the RA procedure is not triggered. While a Regular BSR is triggered when a UE has user data to send, other BSRs can be triggered even when the UE does not have any user data to send. Because the uplink radio resource is allocated as a result of SR transmission on the PUCCH or the RA procedure, only a Regular BSR can trigger the procedures to prevent unnecessary allocation of radio resource.

One special type of BSR is the Truncated BSR. The Truncated BSR has the same format as the Short BSR but with a different LCID field. The Truncated BSR is used when a Padding BSR is triggered and the space in the MAC PDU is not enough to include a Long BSR but when more than one LCG has data in the buffer. From the UE's perspective, there is not so much difference between a Short BSR and a Truncated BSR. However, from the eNB's point of view, the eNB needs to be sure of the meaning of the BSR, in particular, whether the information refers to only one LCG or several. If the eNB receives a Short BSR from the UE, this means that only the LCG included in the Short BSR has data to send and other LCGs do not have any data. If a Padding BSR is triggered and the UE sends a Short BSR, the eNB will conclude that other LCGs not included in the BSR do not have any data. Thus, when the UE cannot include a Long BSR in the MAC PDU and when more than one LCG have data to send, the UE uses the Truncated BSR instead of the Short BSR. When the eNB receives a Truncated BSR, it knows implicitly that other LCGs not indicated in the Truncated BSR also have data to send.

6.5.2 Scheduling Request (SR)

The SR procedure is used by the UE to request radio resource for a new uplink transmission. The SR procedure starts when a Regular BSR is triggered but uplink radio resource to transmit the BSR is not available in the UE. During the SR procedure, the UE performs either transmission of the SR over the PUCCH or initiates the Random Access (RA) procedure, depending on whether the UE is configured with the PUCCH resource for SR or not. The RA procedure is initiated only when the PUCCH resource for SR is not configured.

Performing the RA procedure during the SR procedure is not beneficial, in that it takes quite a long time for the UE to obtain uplink resource. Since the RA procedure is subject to collision, the time spent getting uplink resource may be unnecessarily long. Therefore, from a resource allocation delay point of view, it is good to allocate the PUCCH resource for SR. However, it may not always be possible for the eNB to configure the UE with the PUCCH resource for SR because the PUCCH resource in a cell is limited. Thus, there is a trade-off between resource allocation delay and the PUCCH load.

The PUCCH resource for SR is allocated by the eNB in a periodic manner. The periodicity of the PUCCH resource allocated for SR is called *SR periodicity*. The SR periodicity impacts upon the delay for the UE to obtain uplink resource; that is, short periodicity leads to fast resource allocation. However, short periodicity consumes more PUCCH resource, because the eNB has to set aside a large chunk of PUCCH resource for SR transmission even though the SR is not transmitted. Thus, the eNB should strike a good balance between SR periodicity and PUCCH load.

When the UE transmits the SR over the PUCCH, the UE keeps track of the number of SR transmissions. If the UE has not been allocated any uplink radio resource after reaching the maximum number of SR transmissions, it releases the PUCCH resources for SR and then initiates the RA procedure instead. This unsuccessful transmission of the SR over the PUCCH may have been caused by the wrong configuration, such as inaccurate transmission power. The eNB may identify the wrong SR configuration by detecting the RA procedure initiated by a UE configured with the PUCCH resource for SR.

In Release 9, a timer, *sr-ProhibitTimer*, is introduced to prevent transmission of another SR within a short time period. Since there is a delay between SR transmission and uplink resource allocation (more than 4 ms), it is useless to transmit another SR during that period. Thus, the UE is not allowed to transmit an SR while the *sr-ProhibitTimer* is running. The value of the *sr-ProhibitTimer* is defined in multiples of SR periodicity.

6.5.3 Power Headroom Report (PHR)

Unlike the eNB, which is fixed in location and always plugged into the power outlet, the UE is assumed to be moving constantly and thus limited in power. With a given maximum power, if the UE is allocated more resources than it can support, the decoding error rate at the eNB will increase. Thus, it is important that the eNB has an accurate power status for the UE and allocates a suitable amount of radio resource. The Power Headroom Report (PHR) is used to provide the eNB with information about the difference between the nominal maximum transmit power and the estimated required power for uplink transmission.

RRC signaling is used to deliver the configuration parameters for the *periodicPHR-Timer*, the *prohibitPHR-Timer*, and the *dl-PathlossChange*. The *periodicPHR-Timer* is used to make the UE periodically transmit a PHR. The *prohibitPHR-Timer* is used to prevent frequent transmission of PHRs. When the pathloss fluctuates dramatically, the short-term changes in pathloss are filtered out by the *prohibitPHR-Timer*. The *dl-PathlossChange* parameter is used as a criterion to trigger a PHR. When the measured downlink pathloss is larger than this parameter, a PHR is triggered by the UE.

6.6 Logical Channel Prioritization (LCP)

The finite radio resource should be allocated and used carefully among the UEs and radio bearers. In the downlink, the eNB is the focal point through which all downlink data flows before being transmitted over the radio interface to each UE. Thus, the eNB can make consistent decisions about which downlink data should be transmitted first. However, in the uplink, each UE makes an individual decision based only on the data in its own buffers and the allocated radio resource. To ensure that each UE makes the best and most consistent decisions in terms of using the allocated radio resource, the Logical Channel Prioritization (LCP) procedure is introduced. The LCP procedure is used for MAC PDU construction by deciding the amount of data from each logical channel and the type of MAC Control Element that should be included in the MAC PDU. By using the LCP procedure, the UE can satisfy the QoS of each radio bearer in the best and most predictable way.

In constructing a MAC PDU with data from multiple logical channels, the simplest and most intuitive method is the absolute priority-based method, where the MAC PDU space is allocated to logical channels in decreasing order of logical channel priority. That is, data from the highest priority logical channel are served first in the MAC PDU, followed by data from the next highest priority logical channel, continuing until the MAC PDU space runs out. Although the absolute priority-based method is quite simple in terms of UE implementation, it sometimes leads to starvation of data from low-priority logical channels. Starvation means that the data from the low-priority logical channels cannot be transmitted because the data from high-priority logical channels take up all the MAC PDU space.

In LTE, a Prioritized Bit Rate (PBR) is defined for each logical channel, in order to transmit data in order of importance but also to avoid starvation of data with lower priority. The PBR is the minimum data rate guaranteed for the logical channel. Even if the logical channel has low priority, at least a small amount of MAC PDU space is allocated to guarantee the PBR. Thus, the starvation problem can be avoided by using the PBR.

Constructing a MAC PDU with PBR consists of two rounds. In the first round, each logical channel is served in decreasing order of logical channel priority, but the amount of data from each logical channel included in the MAC PDU is initially limited to the amount corresponding to the configured PBR value of the logical channel. After all logical channels have been served up to their PBR values, if there is room left in the MAC PDU, the second round is performed. In the second round, each logical channel is served again in decreasing order of priority. The major difference for the second round compared to the first round is that each logical channel of lower priority can be allocated with MAC PDU space only if all logical channels of higher priority have no more data to transmit.

A MAC PDU may include not only the MAC SDUs from each configured logical channel but also the MAC CE. Except for a Padding BSR, the MAC CE has a higher priority than a MAC SDU from the logical channels because it controls the operation of the MAC layer. Thus, when a MAC PDU is composed, the MAC CE, if it exists, is the first to be included and the remaining space is used for MAC SDUs from the logical channels. Then, if additional space is left and it is large enough to include a BSR, a Padding BSR is triggered and included in the MAC PDU.

Table 6.2 shows the priority order considered when generating a MAC PDU. Among the several types of MAC CE and data from the logical channels, the C-RNTI MAC CE and data from the UL-CCCH have the highest priority. The C-RNTI MAC CE and data from the UL-CCCH are never included in the same MAC PDU. Unlike data from other logical

Table 6.2 Priority of MAC CEs and data from logical channels

Priority	
Highest	MAC CE for C-RNTI or data from UL-CCCH
	MAC CE for BSR, with the exception of BSR included for padding
	MAC CE for PHR
	Data from any logical channel, except data from UL-CCCH
Lowest	MAC CE for padding BSR

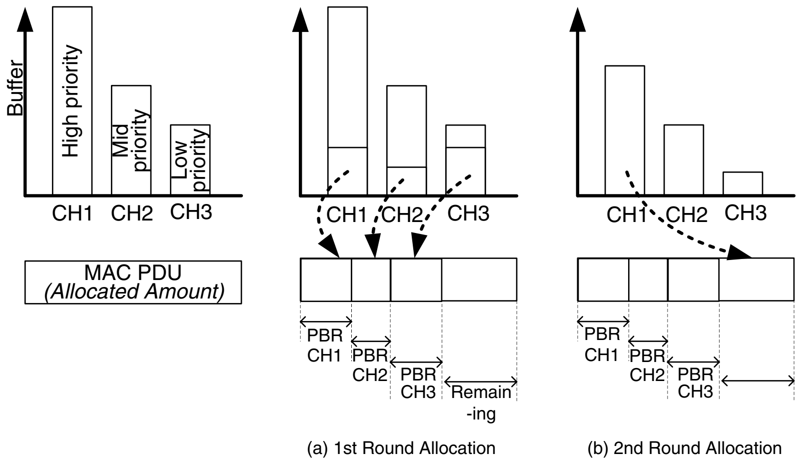

Figure 6.5 Example of logical channel prioritization

channels, data from the UL-CCCH have higher priority than other MAC CEs. Because the UL-CCCH transports an RRC message using SRB0, UL-CCCH data must have higher priority than other data. Typically, data from the UL-CCCH are transported during the RA procedure and the size of a MAC PDU from the UL-CCCH is limited. The C-RNTI MAC CE is used during the RA procedure by a UE whose existence is known by the eNB. Since the RA procedure is subject to collision, it is important to have a means by which the eNB can identify each UE. Thus, the UE is required to include its C-RNTI as its identity as early as possible during the RA procedure.

Figure 6.5 illustrates an example of how LTE MAC multiplexing is performed. In this example, the following are assumed:

- there are three channels: channel 1 is of the highest priority, channel 2 is of middle priority, and channel 3 is of the lowest priority;
- channel 1, channel 2, and channel 3 have been assigned PBR values.

In the first round, each channel is served up to the data amount equivalent to the PBR according to the order of priority. In this first round, a channel without any configured PBR value is not served. In addition, if the amount of data available for the channel is less than the configured value of the PBR, the channel is served up to the data amount that is available in the buffer. As shown in Figure 6.5(a), each channel is allocated space in the MAC PDU up to its configured value of PBR.

In the second round, a logical channel is served only when the following three conditions are met:

- after the logical channels of higher priority than the concerned logical channel have been served;
- there is space remaining in the MAC PDU;
- there are data available in the channel's buffer.

Accordingly, as shown in Figure 6.5(b), channel 1 is served first. Because the remaining data in the buffer for channel 1 are larger than the remaining space in the MAC PDU, all the remaining space in the MAC PDU is allocated to channel 1. Because there is no more space, channels 2 and 3 are not served in the second round.

The description above is the general principle and is not enforced every time a new MAC PDU is composed. Each MAC SDU corresponds to one RLC PDU and one RLC PDU includes at least 1 byte of RLC PDU header. For each MAC SDU, there exists a corresponding at least 1 byte MAC subheader. Thus, whenever a small amount of data from one logical channel is included in a MAC PDU, it will incur at least 2 bytes of header overhead. If the above multiplexing principle was applied in every MAC PDU, the overall overhead caused by the MAC subheader and the RLC PDU header of every logical channel in a MAC PDU would be huge. Thus, rather than applying the above PBR requirements for every subframe, it is better to meet the PBR requirements for a long time period. To reduce the overhead and to prevent too much segmentation, the token-bucket model with PBR is applied.

In the token-bucket model, each logical channel is associated with two parameters: *bucketSizeDuration* and *prioritizedBitRate*. In this model, it is assumed that each logical channel is given a right to transmit a *prioritizedBitRate* amount of data in every subframe. If a certain logical channel has not fully used the right to transmit its *prioritizedBitRate* amount of data in a certain subframe, the remaining right can be used in another subframe. The right to transmit can be accumulated up to a (*prioritizedBitRate* × *bucketSizeDuration*) amount of data. When some data for the logical channel are included in a MAC PDU, the right to transmit is decreased by the amount of data included in the MAC PDU. To prevent a certain logical channel from accumulating too much right to transmit, the parameter *bucketSizeDuration* sets the limit up to which a logical channel can accumulate the right to transmit. Through this token-bucket model, the UE can meet the PBR principle on average for a longer time period, not per subframe.

Figure 6.6 shows an example of logical channel prioritization. Here, for the given logical channel, it was assumed that *bucketSizeDuration* was 4 ms (subframes) and *prioritizedBitRate* was 1 Kb/ms. Thus, the logical channel cannot accumulate more than 4 Kb worth of

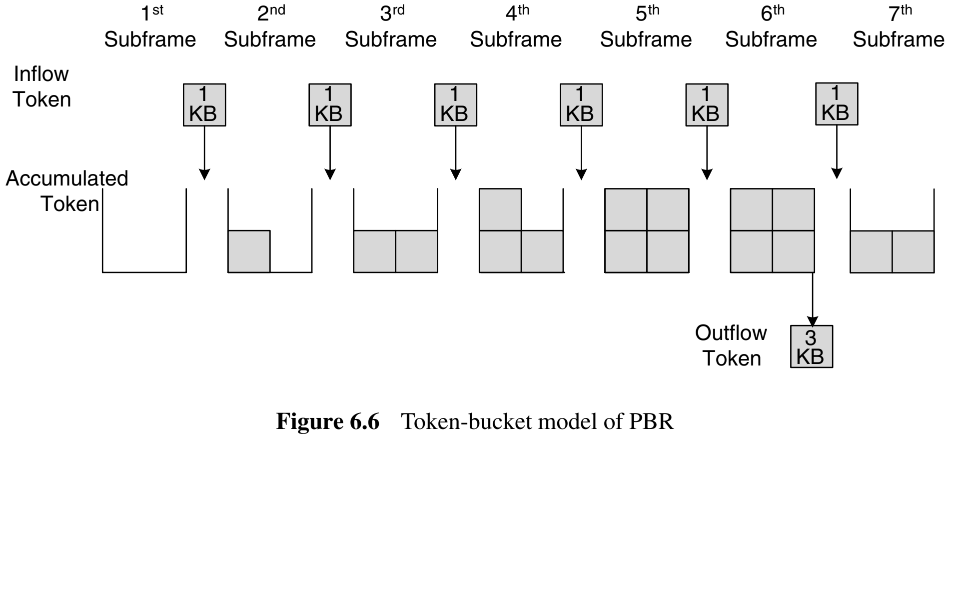

Figure 6.6 Token-bucket model of PBR

right to transmit. In other words, even if data from the logical channel have not been transmitted for a long time, the maximum number of bits that the logical channel can transmit is 4 Kb. In the example, the logical channel has not transmitted any data for the 1st subframe to the 5th subframe. But, because of the limited size of the token bucket, the maximum token accumulated by the logical channel at the 5th subframe is 4 Kb. In the 6th subframe, 3 Kb of data from the logical channel have been transmitted. Because 1 Kb worth of token is accumulated at the 7th subframe, the total accumulated token for the logical channel at the end of the 7th subframe is 2 Kb. Thus, even if the logical channel has not transmitted any data, it can make a lot of transmissions at a later time thanks to the accumulated token, but no more than the maximum token.

6.7 Discontinuous Reception (DRX)

Battery saving is the most important issue in mobile communication. To reduce the battery consumption in the UE, a mechanism to minimize the time that the UE spends monitoring the PDCCH is used; this is called the Discontinuous Reception (DRX) functionality.

The DRX mechanism should meet two conflicting requirements. On the one hand, it should minimize latency in transferring data from one node to another node. On the other hand, to minimize battery consumption, it should minimize the time period for which the UE is mandated to listen to the downlink channel. As a result, the parameterization of DRX involves a trade-off between battery saving and latency reduction.

In LTE, up to two different lengths of DRX cycle – that is, long DRX cycle and short DRX cycle – can be used for a UE in RRC_CONNECTED. Note that the use of short DRX cycle is optionally configured by the eNB.

A long DRX cycle is beneficial for lengthening the UE's battery life. For example, in the scenario where a user is using a Web-browsing service, it may be a waste of UE battery if the UE has to monitor the downlink channel continuously while the user is reading a downloaded Web page. In this case, by using a long DRX cycle, the UE can minimize the time used for monitoring the downlink channel.

On the other hand, a shorter DRX cycle is better when data transfer is resumed quickly. In a packet-switched network, some user packets may arrive at the eNB significantly later than other chunks of the user packets. In this case, if the UE is put into a long DRX cycle before the arrival of the late user packets, the late user packets will be delivered to the UE at the next wake-up time of the long DRX cycle. Because these user packets have already experienced a long delay before the arrival at eNB, the additional delay of a long DRX cycle will badly impact the QoS. Thus, to prevent the UE from entering a long DRX cycle too early, the UE is put into a short DRX cycle. If user packets arrive at the eNB while the UE is using the short DRX cycle, the UE is put back into continuous reception mode by the downlink assignment at the next wake-up time of the short DRX cycle. However, if no more user packets arrive during the short DRX cycle, the UE is put into a long DRX cycle, on the assumption that the packet transfer activity has finished.

Figure 6.7 illustrates the transitions among the long DRX cycle, the short DRX cycle, and the continuous reception mode. If a certain criterion is met, the UE moves from continuous reception mode to a short DRX cycle, from the short DRX cycle to a long DRX cycle, and from the long DRX cycle to continuous reception mode. The transition is controlled either by timers or by explicit commands from the eNB.

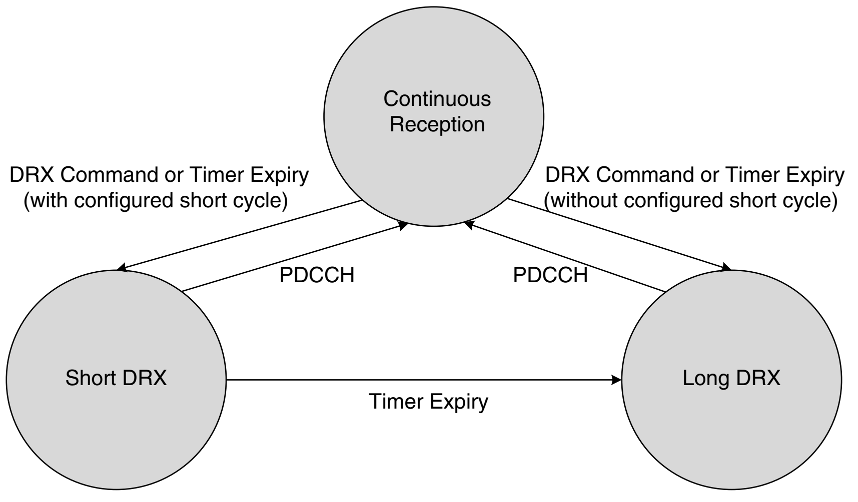

Figure 6.7 State transitions for DRX

The *drx-InactivityTimer* controls the transition from continuous reception mode to a short DRX cycle or a long DRX cycle. If the UE does not receive any new resource allocation information until the expiry of the *drx-InactivityTimer*, it transits to the next level of the DRX cycle; that is, it transits to a short DRX cycle if configured and to a long DRX cycle otherwise.

When the UE moves into a short DRX cycle, it starts *drxShortCycleTimer*. The UE stays in the short DRX cycle until the expiry of the *drxShortCycleTimer*, and moves to a long DRX cycle at the expiry of the *drxShortCycleTimer*. If the UE receives any resource allocation information while the *drxShortCycleTimer* is running, it moves back to continuous reception mode. Actually, whenever the UE receives resource allocation information indicating a new transmission during any of the DRX cycle, it immediately moves back to continuous reception mode.

In addition to timer-based transition, explicit-command-based transition is also used. If the eNB is sure that there is no more data for the UE or if it does not want to schedule any downlink data for the UE for some time, the eNB can send a DRX command MAC CE to the UE. When the UE receives this DRX command MAC CE, it immediately moves either to a short DRX cycle (if configured) or to a long DRX cycle.

In continuous reception mode, the UE is monitoring the PDCCH in all subframes except the subframes for uplink transmission of the half-duplex FDD UE operation and the subframes for the measurement gap. Simply put, continuous reception mode can be matched to the time when the *drx-InactivityTimer* is running.

In the short DRX cycle and the long DRX cycle, the UE monitors the PDCCH for some of the subframes out of all available subframes. Because the UE monitors only a small portion of all possible subframes, the UE's power consumption can be reduced. To indicate the availability of downlink data and to minimize latency in delivering them, the UE and the eNB should have a common understanding about when the UE is expected to monitor the PDCCH and when the eNB can send resource allocation information to the UE.

The mandatory PDCCH monitoring time in each DRX cycle is called the "On Duration" and it is located in the first part of each DRX cycle. More specifically, a DRX cycle consists of an "On Duration" during which the UE should monitor the PDCCH and the "DRX period" during which the UE is allowed not to monitor the PDCCH.

The number of subframes or the length of time that a UE should monitor the PDCCH in one DRX cycle is controlled by the *onDurationTimer*. At the beginning of each DRX cycle, the UE starts the *onDurationTimer* and monitors the PDCCH while the timer is running. The length of the timer controls the scheduling flexibility of the eNB. If the length of the *onDurationTimer* is one subframe, the eNB can send a resource allocation message only during that one subframe. However, if the length of the *onDurationTimer* is more than one subframe, the eNB can select one of the available subframes to send the resource allocation information. This is beneficial to the eNB especially when the PDCCH is heavily loaded. Thus, depending on the length of the *onDurationTimer,* the eNB can have flexibility regarding when to send resource allocation information. However, this comes at a cost to the UE, because monitoring of one more subframe means more consumption of the UE's battery.

Figure 6.8 shows an example of DRX operation. Figure 6.8(a) shows the timing of a short DRX cycle and a long DRX cycle. Regardless of the DRX cycle used, the UE should monitor the PDCCH during the On Duration. By mandating the UE to check the PDCCH at least during the On Duration, the eNB does not need to wait endlessly for UE access. During the period outside of the On Duration, the UE can save its battery by not monitoring resource allocation information.

The Active Time is the time when a UE should monitor the PDCCH for possible resource allocation information. This Active Time includes the time period when timers such as *drx-InactivityTimer*, *drx-RetransmissionTimer*, and *onDurationTimer* are running. When resource allocation information is received during the Active Time, the UE starts or restarts the *drx-InactivityTimer* and monitors the PDCCH in every subframe while the *drx-InactivityTimer* is running. Thus, resource allocation information received during the Active Time effectively extends the Active Time. This is shown in part (b) of Figure 6.8.

At the expiry of the *drx-InactivityTimer* or on receipt of a DRX Command MAC CE, the UE stops the Active Time and moves into a short DRX cycle or a long DRX cycle. The eNB can send a DRX Command MAC CE at any time to put the UE immediately into a DRX cycle. This is shown in part (c) of Figure 6.8.

To maximize UE battery efficiency, HARQ characteristics are also considered in the DRX operation. This is shown in Figure 6.9.

As shown in part (a) of Figure 6.9, the HARQ operation consists of sequential steps: sender's transmission, the propagation delay over the radio interface, the receiver's decoding and transmission of the ACK/NACK, another propagation delay of the ACK/NACK over the radio interface, the sender's reception of the ACK/NACK, and another transmission by the sender. Considering all these steps, there is quite a long time period between subsequent transmissions and reception over one HARQ process.

To utilize this finding, the HARQ RTT timer is used for each DL HARQ process to reduce the UE's power consumption while waiting for HARQ retransmissions. When the decoding of a downlink transport block for one HARQ process fails, the UE starts the HARQ RTT timer for the HARQ process, assuming that the next HARQ retransmission of the process will occur at least after the expiry of the HARQ RTT timer.

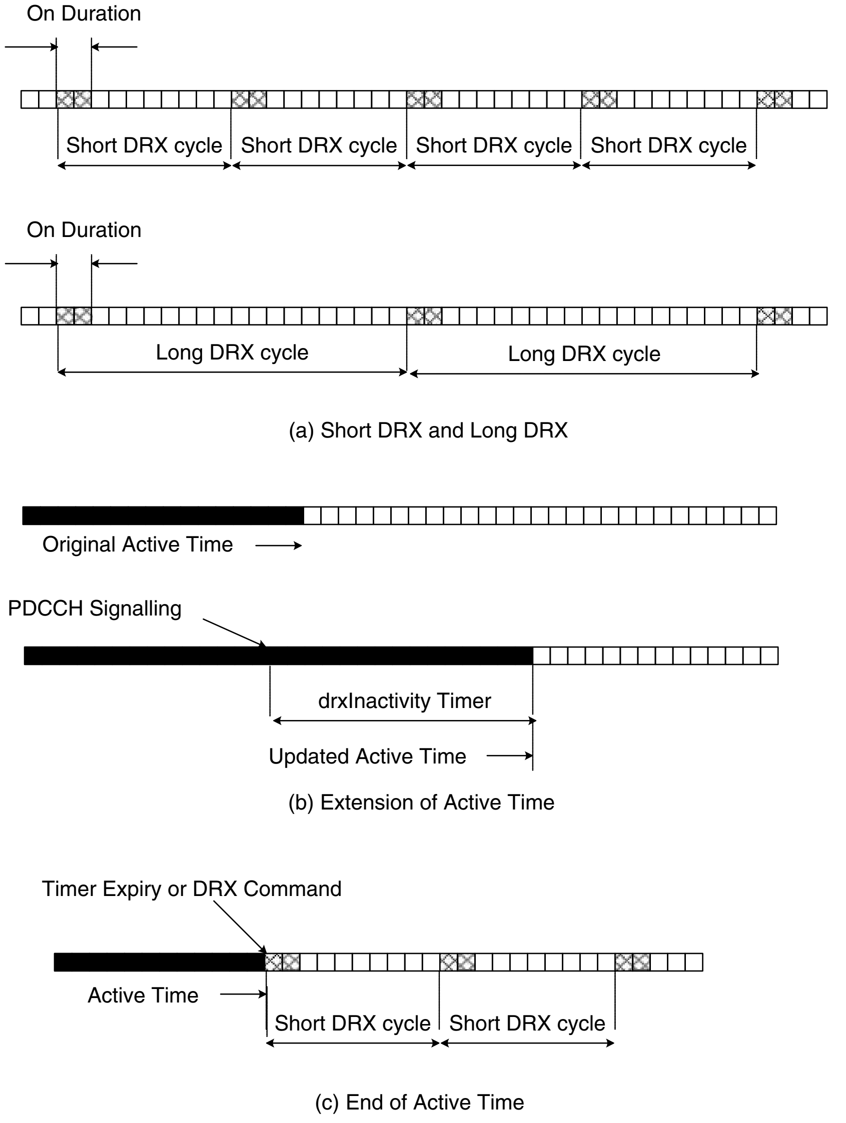

Figure 6.8 Detailed DRX operation

For the HARQ process for which the HARQ RTT timer is running, the UE does not need to monitor the PDCCH.

If a DL HARQ process is not successful in decoding a transport block, the eNB may perform the HARQ retransmission for the HARQ process. Thus, at the expiry of the HARQ RTT timer, the UE should resume monitoring the PDCCH for a possible HARQ retransmission. However, the UE cannot wait endlessly for a HARQ retransmission because a HARQ NACK-to-ACK error might occur or the eNB might just decide not to transmit

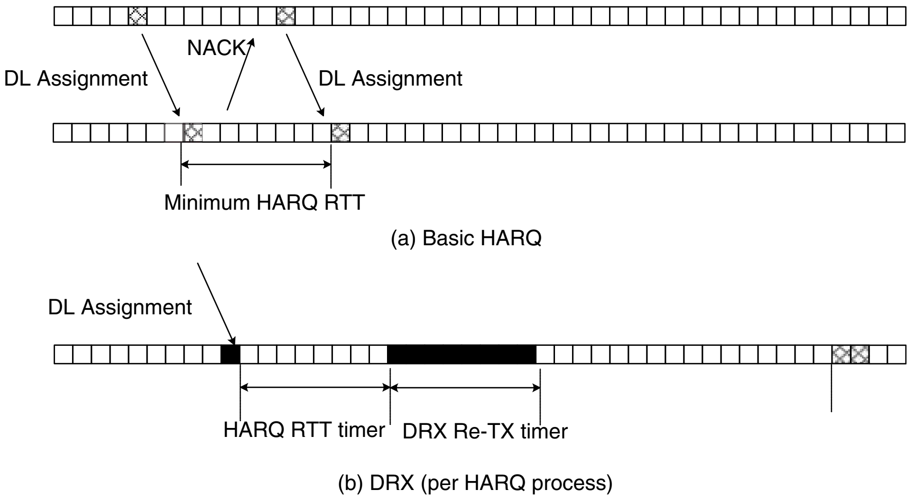

Figure 6.9 HARQ-related DRX operation

anything for the concerned HARQ process. To limit the PDCCH monitoring time for potential retransmission, a *drx-RetransmissionTimer* is used for the DL HARQ process. This timer starts at the expiry of the HARQ RTT timer, and the UE monitors the PDCCH only when the *drx-RetransmissionTimer* is running.

The length of the *drx-RetransmissionTimer* is related to how much eNB scheduling flexibility is required. For optimal UE battery consumption, it is desirable for the eNB to schedule HARQ retransmission as soon as the HARQ RTT timer expires. However, this requires the eNB always to reserve the capacity to transmit resource allocation information for the subframe where the HARQ RTT timer expires. To relax this scheduling restriction and to limit the amount of time that the UE spends monitoring the PDCCH, the length of the *drx-RetransmissionTimer* should be selected carefully.

6.8 Hybrid-ARQ (HARQ)

Due to the innate adverse condition of the radio channel, what a sender has transmitted is usually not what a receiver has received. To overcome this obstacle, the receiver should be able to detect whether it has received everything correctly or not. In addition, the sender should be able to make another transmission if there is something wrong. A communication method using the above two elements is called Automatic Repeat reQuest (ARQ). However,, if a receiver can perform recovery and correction on its own, this will improve transmission efficiency. This capability is called Forward Error Correction (FEC). ARQ combined with FEC is called Hybrid-ARQ (HARQ).

In LTE, HARQ with soft combining is used. In this approach, when a coded bit block for a data unit is received, the receiving side tries to decode the received coded bit block. If the decoding is successful, the recovered data unit is forwarded to the upper layer. However, if the decoding is unsuccessful, the receiving side informs the transmitting side of the decoding failure and the transmitting side makes another transmission for the data unit. Then, the first received coded bit block and the subsequently received coded bit block are combined and decoded to recover the original data unit. Though each received coded bit block is not enough to recover the original data unit independently, sometimes it is possible to recover the original data unit if multiple coded bit blocks are combined for decoding.

In LTE, the N-channel Stop-And-Wait (SAW) method is also adopted. The SAW operation means that upon the transmission of a data unit, the transmitting side stops further transmissions until feedback is received from the receiving side. When a HARQ NACK is received, the transmitting side retransmits the previous data unit. However, due to propagation delay over the radio interface and processing delay on both the transmitting and receiving sides, there is a time period during which no transmission is made by the transmitting side. To use the radio resource efficiently and to increase the data rate, it is essential to minimize this time period. Therefore, multiple independent HARQ processes are interlaced in time so that all transmission time resources can be used by at least one of the HARQ processes. Each HARQ process is responsible for a separate SAW operation and manages a separate decoding buffer. An example of the N-channel SAW operation is shown in Figure 6.10.

In general, the HARQ operation can be categorized according to whether the retransmission for a given HARQ process occurs at a predefined time or at an arbitrary time, and also according to whether the HARQ retransmission is adaptive or non-adaptive. In LTE, asynchronous adaptive HARQ is used for the downlink and synchronous adaptive or synchronous non-adaptive HARQ is used for the uplink.

The New Data Indicator (NDI) field delivered through the PDCCH indicates whether the transmitted data unit is a new one or a retransmitted one. If the value of the NDI field in the downlink assignment changes compared to the value used in the previous downlink assignment, this means that a new data unit is transmitted for the given HARQ process. Likewise, if the value of the NDI field changes in the uplink grant, the UE should transmit a new data unit.

The reason why ARQ or HARQ is used is that the error may occur during transmission and reception. Due to noise in the channel, what has been transmitted by the transmitter is sometimes corrupted over the radio interface. By using ACK or NACK, the transmitting side should know whether its transmission was received successfully by the receiving side. However, ACK and NACK are transmitted over the same medium that the data travel through.

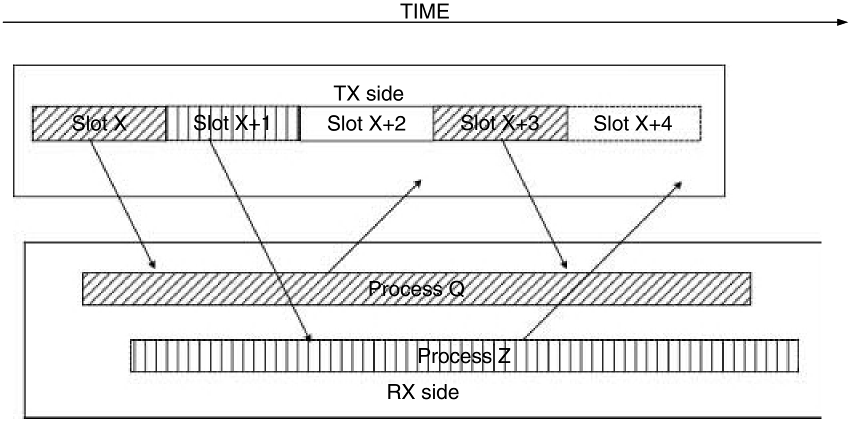

Figure 6.10 N-channel SAW operation

Thus, sometimes a transmitted ACK can be received as a NACK and vice versa. This error is called ACK-to-NACK error or NACK-to-ACK error, respectively, and the system should be ready to handle such error cases.

6.8.1 HARQ in the Uplink

In the synchronous HARQ scheme used in the uplink of LTE, retransmission(s) for each HARQ process occurs at a predefined time relative to the initial transmission. Using the time interval of successive transmissions of the HARQ process and the time duration of each transmission, the maximum number of usable HARQ processes can be calculated. Because HARQ processes are used in a round-robin fashion, the HARQ process being used at a given time can be identified easily based on the transmission timing.

In the uplink, HARQ retransmissions can be either adaptive or non-adaptive. Adaptive HARQ retransmission means that new signaling of the transmission attributes, such as the modulation and coding scheme, and transmission resource allocation in the frequency domain are provided at the time of the HARQ retransmission. Non-adaptive HARQ retransmission means that the previously used transmission attributes are used again without any further signaling of the transmission attributes.

In the non-adaptive retransmission scheme used as the default in the uplink, when no explicit signaling is received over the PDCCH, the UE performs HARQ retransmissions using exactly the same transmission attributes used in the previous HARQ transmission for the concerned HARQ process, or using transmission attributes following the pre-configured rules. However, when new signaling is received at the time of retransmission, the HARQ retransmission is performed using the transmission attributes included in that signaling message.

6.8.2 HARQ in the Downlink

In the downlink for LTE, the asynchronous HARQ scheme is used. HARQ retransmissions for a given HARQ process can occur at any time relative to the initial transmission. Because several HARQ processes are active at the same time, when the eNB assigns radio resources to the UE, explicit signaling to indicate which HARQ process is intended is required in the resource allocation information over the PDCCH. In that explicit signaling, the process identifier of the target HARQ process should be included so that the receiver can correctly associate each HARQ retransmission with the corresponding initial transmission of the same HARQ process.

In the adaptive retransmission scheme used in the downlink, the transmission attributes can be changed at each HARQ retransmission in response to variations in the radio channel conditions. Though the adaptive scheme brings more scheduling gain and flexibility, it also brings increased signaling overhead over the PDCCH. This is because whenever there is HARQ retransmission, the eNB has to send resource allocation information.

6.8.3 TTI Bundling

Due to the use of a finite battery and the limited capacity of power amplifiers, there is a limit in the uplink transmission power of the UE. Thus, when the UE is located near to the cell edge, even a small packet such as VoIP may not be transmittable in a single subframe. In other words, compared to transmission by a UE located in the center of a cell, the UE at the cell edge needs more energy per subframe to achieve the same bit error rate. If the UE cannot increase its transmission power, the UE should alternatively lengthen the transmission time to increase the effective energy per bit. In the end, this can be regarded as decreasing the effective amount of data per subframe.

Decreasing the effective amount of data is also possible in the RLC layer by generating smaller PDUs per subframe. However, this will result in more overhead due to RLC and MAC PDU headers, and thus a MAC-based solution is preferable to an RLC-based solution. TTI bundling is one of the MAC-based solutions.

TTI bundling is effective in extending the uplink coverage and minimizing latency. If TTI bundling is not used, it takes 24 subframes for one HARQ process to make one initial transmission with three retransmissions. However, if TTI bundling is used, it takes only four subframes to make four HARQ (re-)transmissions. Thus, this TTI bundling technique is helpful in delay-sensitive services such as a VoIP service where the user packets are generated every 20 ms.

In a normal dynamic HARQ operation, each uplink HARQ process can have transmission opportunities every eight subframes and each retransmission is controlled separately by HARQ feedback from the eNB. On the other hand, if TTI bundling is used, each HARQ process uses a TTI bundle consisting of four consecutive subframes for uplink transmissions.

Uplink HARQ transmissions made in the TTI bundle share the same HARQ feedback. As shown in Figure 6.11, if TTI bundling is not used, each HARQ retransmission is performed when the UE receives a HARQ NACK from the eNB. However, if TTI bundling is used, even though there is no HARQ feedback, the UE automatically performs HARQ retransmissions in the subsequent three subframes following the first subframe in the TTI bundle.

When the eNB wants to control the HARQ retransmissions of the TTI bundle, it should send the uplink grant considering the timing of the first subframe of the bundle. The UE determines a successful transmission of the TTI bundle based on the HARQ feedback with timing associated with the last subframe of the TTI bundle. In this case, the last subframe is

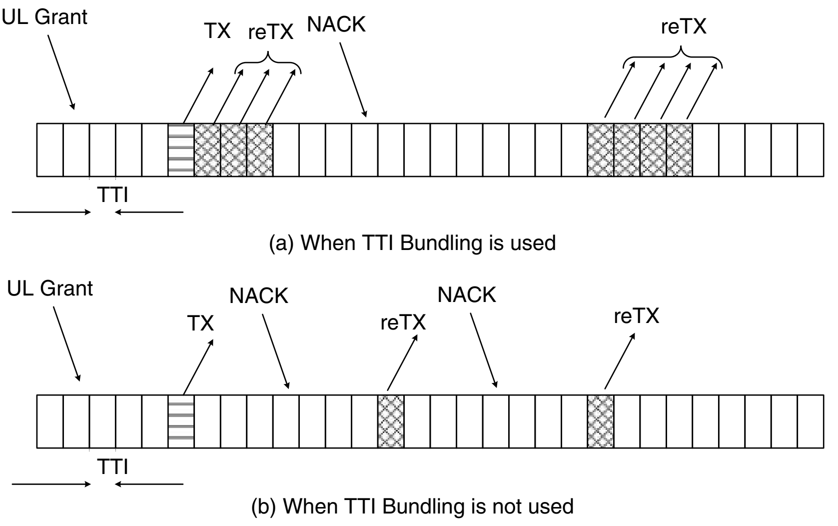

Figure 6.11 TTI bundling

always the fourth subframe of the TTI bundle, regardless of whether actual transmission occurred or not in that fourth subframe in the TTI bundle.

The use of TTI bundling also brings with it a trade-off. Because each TTI bundle consists of four subframes, this requires a lot of scheduling consideration at the eNB. Because four subframes are a big chunk of the radio resource, the eNB should be careful that each subframe of the TTI bundle does not collide with the allocated radio resource for other UEs. HARQ retransmissions in a TTI bundle are non-adaptive, and this will further limit the eNB's scheduling flexibility. In addition, if the eNB decodes the user data successfully using the first few subframes of the TTI bundle, the other subframes of the TTI bundle are just a waste of radio resource. However, as stated previously, TTI bundling reduces the latency of data transmission by a UE at the cell edge and this is beneficial for time-critical services such as a VoIP service. In addition, with TTI bundling, because the last three subframes in a TTI bundle do not require any signaling such as HARQ feedback or the uplink grant, this effectively reduces the overhead of the resource allocation information.

6.8.4 Measurement Gap

Because the UE is moving around, the cell to which the UE is connected should be changed according to the location of the UE and the quality of the radio signal. Although the eNB can measure the quality of the radio signal transmitted by the UE, it cannot measure the quality of the radio signal received by the UE. Thus, to make an accurate decision regarding when to change the serving cell and to which cell the UE should be connected, the UE should measure the downlink signal of the cells and should report the result to the eNB.

Normally, the UE can perform measurement of cells of the same frequency that the UE is using and it can measure cells of frequencies other than the frequency that the UE is camping on while it is in RRC_IDLE. But when the UE is transmitting and receiving user data in RRC_CONNECTED, it may not have time to measure other frequencies or other Radio Access Technology (RAT). For example, when all eight uplink HARQ processes have been scheduled with either initial transmission or retransmission, the UE cannot tune in to other frequencies or other RAT. In this case, to make the UE perform the required measurement, the eNB configures the UE with a measurement gap during which the UE does not need to transmit in the uplink nor listen to downlink channels. Thus, during the measurement gap, the UE can stay away from the serving frequency and ignore HARQ operations.

However, some rules are applied to the HARQ operation in relation to the measurement gap in order to reduce the impact on data transfer activity. When the HARQ feedback for a transmitted transport block cannot be received due to measurement gaps, the UE considers that a HARQ ACK has been received for the transport block. In this case, the UE does not autonomously start HARQ retransmissions at the next transmission opportunity. To resume the HARQ operation on the HARQ process, the UE has to receive new resource allocation information. The UE then performs either initial HARQ transmission or HARQ retransmission depending on the content of the resource allocation information. When an uplink transport block cannot be transmitted due to measurement gaps, the UE considers that a HARQ NACK has been received for the transport block. This is a natural consequence because the eNB cannot even try to decode a transport block which has not been transmitted due to the measurement gap. Accordingly, the UE performs HARQ retransmissions of the transport block at the next transmission opportunities after it comes back from the measurement gap.

In the case of the HARQ operation for SPS, the UE starts initial uplink transmission without receiving any resource allocation information. If the initial transmission timing configured by SPS coincides with the timing of the measurement gap, the UE considers that there is allocated radio resource though there was no actual transmission. After returning from the measurement gap, the UE performs HARQ retransmission at the next available transmission timing or acts according to the received resource allocation information, if any, or the HARQ feedback.

6.9 Random Access (RA) Procedure

The RA procedure is used when UE is not allocated with any uplink radio resources but has data to transmit, or when the UE is not synchronized in the uplink. After a successful RA procedure, the UE's transmission timing is synchronized in the uplink and may be allocated with uplink radio resources. When a dedicated preamble is allocated to the UE by the eNB, the RA procedure is initiated. The dedicated preamble may be allocated when downlink data arrives for the UE which is not synchronized in the uplink, or when access to the target cell is required for handover.

There are two types of RA procedure. The first is the contention-based RA procedure in which the RA preamble is randomly chosen by the UE; the other is the non-contention-based RA procedure in which the RA preamble to be used by the UE is designated by the eNB.

For the non-contention-based RA procedure, the RA preamble can be notified either by RRC signaling using an RRC message or by PHY signaling over the PDCCH. When the eNB wants to trigger the RA procedure for a certain UE but it cannot allocate the dedicated RA preamble for the UE due to limited RA preamble resource in a cell, the eNB can inform the UE of the RA preamble whose index is "000000". When the UE is informed of the RA preamble with index "000000", the UE starts the contention-based RA procedure.

For the contention-based RA procedure, there are two groups of RA preamble: the RA preamble group A and the RA preamble group B. Which RA preamble group is used depends on the pathloss, on the estimated size of the MAC PDU, and on whether this RA attempt is the initial attempt or a re-attempt. If the estimated size of the MAC PDU is big and the measured pathloss is small, the UE will choose one RA preamble among the RA preamble group B. Otherwise, the UE will select one RA preamble among the RA preamble group A. In addition, when the UE performs a re-attempt of the RA procedure, the UE should choose the same RA preamble group as that of the RA preamble used in the first attempt of the RA procedure.

These two RA procedures are shown in the left and right sides of Figure 6.12.

In the contention-based RA procedure, because UEs can choose the RA preamble by themselves, it is possible for more than one UE to transmit the same RA preamble simultaneously. In this case, acknowledgment by the eNB of receipt of the RA preamble is not enough, and the eNB should further perform the contention resolution step, through which the eNB should indicate which UE's transmission has actually been received. Thus, to resolve contention, when a UE transmits a MAC PDU over the uplink radio resource allocated by the RA response, the UE includes its identification information in the MAC PDU. So, if it has a valid C-RNTI, the C-RNTI MAC CE is included in the MAC PDU; if it does not have a valid C-RNTI, for example when a CCCH message is transmitted, the CCCH SDU that contains the UE's identification information is included in the MAC PDU.

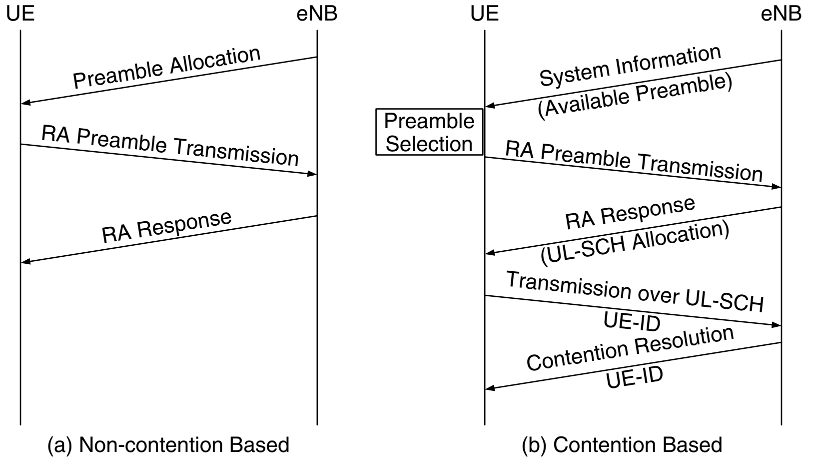

Figure 6.12 Two types of RA procedure

After that, when the UE detects its C-RNTI through the PDCCH or when it receives a UE Contention Resolution Identity MAC CE that is identical to the CCCH SDU previously transmitted, the UE considers the RA procedure successful. However, if the UE fails to detect its matching identifier or the MAC CE, it starts another attempt of the RA procedure.

In the non-contention-based RA procedure, because the designated RA preamble is used by only one specific UE, there is no possibility of collision. As soon as the eNB detects the RA preamble, the eNB knows of the access by the UE and the procedure is terminated by transmission of the RA response. This means that the non-contention-based RA procedure is more efficient and faster than the contention-based RA procedure. Therefore, the non-contention-based RA procedure is used mainly during the handover procedure, which is a time-critical procedure. In addition, if the eNB wants a UE where the uplink timing is not synchronized to be synchronized quickly, the eNB can allocate a designated RA preamble to make the UE perform the RA procedure and become synchronized in the uplink.

Once the RA procedure is initiated, the procedure continues until it is completed successfully or the RRC layer aborts the ongoing RA procedure. Thus, if no RA response is received within a certain time, or if the RA response is received but the matching identifier of the transmitted RA preamble is not included in the RA response, or if the contention resolution was not successful for the UE, the UE starts another attempt of the RA procedure. If attempts of the RA procedure fail consecutively for a certain number of times, the MAC layer reports this failure to the RRC layer. Notification to the RRC layer will eventually stop the ongoing RA procedure.

In the case where a cell is overloaded due to simultaneous RA procedures by many UEs, the eNB can send a backoff parameter through the RA response to cope with the overload situation. This backoff parameter is applicable only in cases where the UE performs the contention-based RA procedure. Once the UE receives the backoff parameter, if no RA response is received or the contention resolution is not successful, the UE will choose a random value between 0 and the value indicated by the backoff parameter. The UE then applies the selected value before the next attempt of the RA procedure; that is, the UE should wait at least for the time equivalent to the selected value before starting the next RA procedure. By delaying some UEs' attempts at the RA procedure with the backoff parameter, the eNB reduces the number of UEs simultaneously attempting the RA procedure. As the number of UEs attempting the RA procedure at a given time decreases, the success rate of the RA procedure goes up and eventually will solve the overload situation.

6.10 Time Alignment

Maintenance of the uplink timing alignment is controlled by the MAC layer and is important for ensuring that the transmission by the UE arrives at the eNB within the defined time window. If the transmission by the UE arrives at the eNB outside the defined time window, the transmission by the UE will act as interference to other UEs. How much the UE should adjust its transmission timing is controlled by the Timing Advance Command (TAC) MAC CE. The eNB calculates how much adjustment is needed by measuring the UE's uplink transmission, and the eNB transmits the adjustment information to the UE via the TAC MAC CE. After receiving the TAC MAC CE, the UE adjusts its uplink transmission timing and starts the *timeAlignmentTimer*. As long as the *timeAlignmentTimer* is running, the UE assumes that its uplink transmission timing is aligned and it can use any uplink radio resource allocated to it. At the expiry of the *timeAlignmentTimer*, the UE assumes that its uplink transmission timing is no longer aligned and releases all configured uplink radio resource such as the radio resources for SPS, the PUCCH, or SRS.

In LTE, the UE does not always need to maintain uplink timing alignment. When user data arrives at the UE's buffer, if the UE's uplink timing is not aligned, the UE should first perform the RA procedure to re-synchronize the uplink timing. Because the RA procedure takes some time to complete, it can be said that there is latency in data transfer for the UE. However, if the UE were to maintain uplink timing alignment even when there was no user data to transfer, the UE would make periodic uplink transmissions for the eNB to decide whether the UE's uplink timing should be adjusted or not. Otherwise, the eNB would have no clue with regard to how much the UE's uplink timing should be shifted. Thus, maintaining timing alignment during the lifetime of connection causes not only a waste of radio resources but also reduces the battery life of the UE. Therefore, when data transfer activity of the UE is inactive for the time being, the UE can be made to lose the uplink timing synchronization even in RRC_CONNECTED. This can be done if the eNB does not transmit a TAC MAC CE to the UE, which eventually leads to the expiry of the *timeAlignmentTimer*.

If the UE's uplink transmission timing is aligned, the UE can use the allocated uplink radio resources. However, if the UE's uplink transmission timing is not aligned, the UE should not use any uplink radio resources except the resource for the RA procedure. This is because the UE's transmission will cause interference to other UEs' transmissions. Note that the UE cannot even transmit a HARQ ACK or HARQ NACK. Therefore, when the eNB wants to transmit user data for a UE where the uplink transmission timing is not aligned, the eNB should first perform the procedure which aligns the uplink transmission timing of the UE. For this purpose, the eNB can order the UE to perform the RA procedure. After a successful RA procedure, the UE will regain uplink timing alignment and the eNB can safely resume the user data transfer toward the UE.

When a UE has user data to send while its uplink timing is not aligned, the UE has to perform the RA procedure to request an allocation of uplink radio resource. During this RA procedure, the UE receives information for the timing adjustment via the RA response. Thus, during the RA procedure, the uplink timing is naturally aligned for the UE, which successfully completes the RA procedure. However, if the UE is not successful at contention resolution in the RA procedure, this means that the received information for the timing alignment via the RA response was not targeted at the UE. In this case, the unsuccessful UE should immediately stop the *timeAlignmentTimer* to invalidate its uplink transmission timing.

Sometimes, uplink data may arrive at a UE that is not allocated any uplink radio resource while the *timeAlignmentTimer* is running. In this case, if the UE uses the contention-based RA procedure to request uplink radio resources, the UE should ignore the received timing alignment information included in the RA response. Otherwise, there is a possibility that the UE may wrongly apply timing alignment information targeted at another UE that performed the RA procedure simultaneously.

6.11 MAC PDU Formats

The MAC Protocol Data Unit (PDU) is a data block unit that is exchanged between the MAC layer and the PHY layer. As a result of multiplexing, the MAC PDU is composed using MAC CEs and MAC Service Data Units (SDUs), and delivered to the PHY layer. Because most CPUs available in the market process data in units of multiples of eight bits, the size of the MAC PDU, the MAC SDU, the MAC CE, and the MAC subheader is set as a multiple of eight bits. The general MAC PDU format is shown in Figure 6.13.

A MAC PDU consists primarily of the MAC header and the MAC payload. The MAC header is further composed of MAC subheaders, while the MAC payload is composed of MAC CEs, MAC SDUs, and padding. Typically, each subheader corresponds to either one MAC CE or one MAC SDU or padding. The appearance order of the MAC CEs, the MAC SDUs, and the padding is the same as the appearance order of the corresponding MAC subheaders. However, for some MAC CEs, there is only subheader and there is nothing in the MAC payload part. To process control information as quickly as possible, MAC CEs are located before any MAC SDUs or padding.

Each MAC subheader consists of the Logical Channel ID (LCID) field, the Length (L) field, the Format (F) field, and the Extension (E) field.

The LCID indicates whether the corresponding part in the MAC payload is a MAC CE, a MAC SDU, or padding. In addition, the LCID field indicates the type of MAC CE or logical channel to which the MAC SDU belongs.

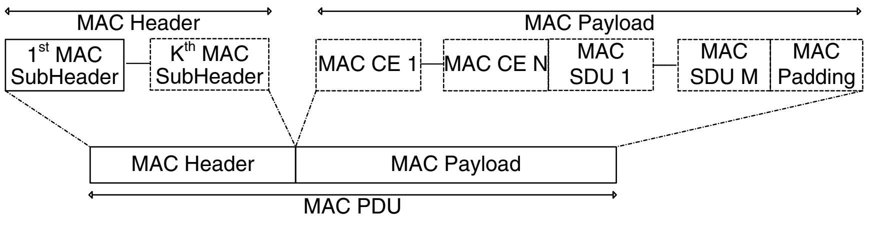

Figure 6.13 Format of a MAC PDU. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

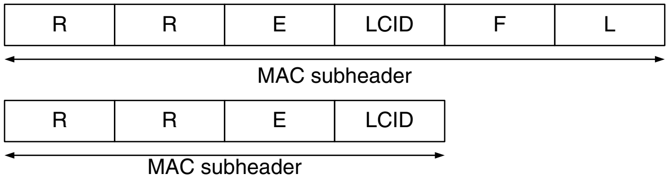

Figure 6.14 Formats of MAC subheader. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

The F field indicates the size of the L field. There are two sizes of the L field: either 7 bits or 15 bits.

The L field indicates the size of the related MAC SDU or the related MAC CE. The size of the L field used depends on whether the size of the corresponding MAC SDU is less than 128 bytes or not.

For the last subheader in the MAC PDU, the L field and the F field are omitted because the size of the last part of the MAC payload can be calculated automatically using the size of the MAC PDU and the size of the other elements in the MAC PDU. In addition, for a MAC CE of fixed size, the L field and the F field are omitted in the corresponding MAC subheader.

The E field indicates whether this subheader is the last subheader or not in the MAC PDU.

Figure 6.14 shows the formats of the MAC subheader. The first is used for a MAC SDU that is not the last element in a MAC PDU or for a flexible-size MAC CE. The second is used for a MAC SDU that is the last element in the MAC PDU or for a fixed-size MAC CE.

When a MAC PDU is used to transport user data from the PCCH or the BCCH logical channels, the MAC PDU includes data from only one logical channel. Because multiplexing is not applied in a MAC PDU for data of logical channels, there is no need to include the LCID field in the MAC header. In addition, for RRC messages transported through the PCCH and the BCCH, the RRC layer composes RRC messages that are the same size as a MAC PDU. Thus, there is no need to indicate the size of the MAC SDU separately in the MAC subheader. The MAC PDU used to transport data for the PCCH or the BCCH is identified easily in the PHY layer because a separate and dedicated identifier is used in the PHY layer signaling. As a result, for MAC PDUs used to transport data from the PCCH or the BCCH, a transparent MAC PDU format is used, as shown in Figure 6.15.

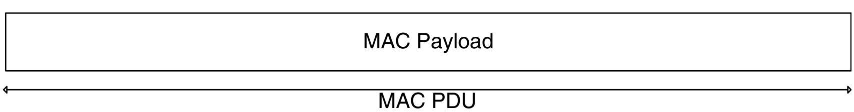

Figure 6.15 Format of transparent MAC PDU. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

Table 6.3 LCID values for MAC CE. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

Index	Direction	MAC CE Type
11010	UL	Power Headroom Report
11011	UL	C-RNTI
11100	UL	Truncated BSR
11101	UL	Short BSR
11110	UL	Long BSR
11100	DL	UE Contention Resolution Identity
11101	DL	Timing Advance Command
11110	DL	DRX Command

6.11.1 MAC Control Elements (CEs)

MAC CEs are used for MAC layer control signaling between the eNB and the UE. For each type of MAC CE, one special LCID value is allocated to identify each MAC CE uniquely. By using the LCID field, there is no need to use a separate field in the MAC subheader to differentiate between MAC SDUs and MAC CEs. LCID values for MAC CEs are shown in Table 6.3.

- **Buffer Status Report MAC CE:** Information regarding how much data is accumulated in the UE's buffer is delivered from the UE to the eNB using this MAC CE. Three LCID values are used to differentiate between a Short BSR, a Long BSR, and a Truncated BSR. The LCG group field in the BSR MAC CE indicates one of the four LCGs to which the following Buffer Size field corresponds. The Buffer Size field actually indicates the amount of data accumulated in the buffer of the concerned LCG. The Short BSR and the Truncated BSR include the buffer status for one LCG, and the Long BSR includes the buffer status for all four LCGs.
- **Power Headroom MAC CE:** Available power headroom information is reported from the UE to the eNB using this MAC CE.
- **DRX Command MAC CE:** This MAC CE is transmitted from the eNB to the UE and used to put the UE into a DRX cycle to save the UE's battery. This MAC CE includes only a MAC subheader, and the size of this MAC CE is fixed.
- **Timing Advance Command MAC CE:** This MAC CE is delivered from the eNB to the UE and informs the UE of the amount of timing adjustment that the UE has to apply for uplink timing alignment. The size of this MAC CE is fixed.
- **C-RNTI MAC CE:** This MAC CE is transmitted from the UE to the eNB and includes the UE's C-RNTI for the purpose of contention resolution during the RA procedure. It is used only when the UE has a valid C-RNTI. The size of this MAC CE is fixed.
- **UE Contention Resolution Identity MAC CE:** This MAC CE is sent from the eNB to the UE and is used for the purpose of contention resolution during the RA procedure. It is used only when the UE does not have a valid C-RNTI. When the UE sends a CCCH SDU during the RA procedure, the eNB includes exactly the same CCCH SDU in this MAC Control

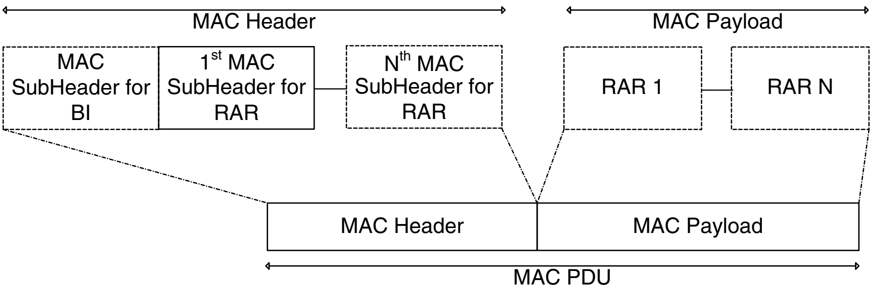

Figure 6.16 Format of MAC PDU for the RA response. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

Element to indicate which UE has won the contention resolution. By comparing what it has sent in the CCCH SDU with the received UE Contention Resolution Identity MAC CE, the UE can decide whether it has finished the RA procedure successfully or not. The size of this MAC CE is fixed.

6.11.2 MAC PDU for Random Access Response

The RA response is used to acknowledge receipt of the RA preamble and to give timing alignment information. Because the RA response message is decoded by multiple UEs, the MAC PDU format for the RA response is different from the normal MAC PDU format, which is decoded by only one specific UE. The MAC PDU for the RA response consists of a MAC PDU header and zero or more MAC RAR elements, as shown in Figure 6.16.

The MAC header consists of one or more MAC subheaders, and the MAC subheader includes either an RA preamble identifier or a backoff indicator. Whether an RA preamble identifier or a backoff indicator has been included is indicated by the value of the Type (T) field.

For each subheader that includes an RA preamble identifier, the corresponding MAC RAR element is included in the MAC PDU. Each MAC RAR element consists of a timing advance command field, a UL Grant field, and a temporary C-RNTI field.

The UL Grant field includes information on the allocated uplink resource that will be used for the UE to transmit the subsequent uplink MAC PDU.

The temporary C-RNTI field includes C-RNTI information used by the UE during the subsequent HARQ operation for the uplink resource allocated by the UL Grant field and during reception of the downlink MAC PDU for contention resolution. For a UE without any valid C-RNTI, the temporary C-RNTI is promoted to a normal C-RNTI once the UE finishes the RA procedure successfully. However, if the UE fails to finish the RA procedure, the temporary C-RNTI included in the MAC RAR element is deleted in the UE. For a non-contention-based RA procedure, the temporary C-RNTI field is ignored.

Reference

1. 3GPP Technical Specification 36.321, "Medium Access Control (MAC) Protocol Specification (Release 10)", www.3gpp.org, accessed on 14 March 2012.

7

Overview of LTE and LTE-Advanced New Features

As described in Section 1.1, the specifications of 3GPP systems are developed in releases. The first specification release for LTE was Release 8, and the focus in the early stage of standardization of the Release 8 specification was the completion of basic functionality. After that, new features for improvement were added in Release 8 and Release 9 until the end of 2009. However, even with these improvements, the Release 9 LTE system did not meet the requirement of a 4G system defined by the International Telecommunication Union (ITU). As a result, additional features were added in the Release 10 specification of LTE to meet the ITU's requirement. In essence, the LTE-Advanced system is a system that is backward-compatible with the LTE system and that meets the performance requirement of the ITU for a 4G wireless system. The specification for Release 10 was completed in the second quarter of 2010, and work on Release 11 is ongoing from February 2012.

This section provides a brief overview of the new features of LTE and LTE-Advanced. The following features are explained in the subsequent sections:

- **Release 8:** Voice over LTE (VoLTE), and Home eNB (HeNB).
- **Release 9:** Public Warning System (PWS), and Multimedia Broadcast/Multicast Service (MBMS).
- **Release 10:** Carrier Aggregation (CA), Relay, Minimization of Drive Test (MDT), enhanced Inter-Cell Interference Coordination (eICIC), and Machine Type Communication (MTC).

7.1 Voice over LTE (VoLTE)

Before the provision of Internet services over wireless networks, the voice service was considered the most important source of revenue by most mobile operators. Even after the introduction of Internet services over wireless networks and the wide penetration of smartphones, the voice service is, and will remain, one of the most essential services that all customers expect. Accordingly, without appropriate support for the voice service, the LTE system cannot succeed.

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

On the other hand, because of the flexibility of the connection, the diversity of the supported media, and the variety of applications, Internet Protocol (IP) based services have been adopted and are used on most network systems, including the LTE system. For example, popular services such as the voice call service, the messaging service, the video call service, the conferencing service, and multimedia services will be provided based on IP. Thus, 3GPP has defined a standardized architecture to support these services over IP, and this architecture is called the IP Multimedia Subsystem (IMS) architecture. The Voice over IP (VoIP) service based on the IMS architecture is called IMS VoIP, and this is the baseline solution to provide voice services within the LTE system.

Although a VoIP service will eventually be provided over the LTE system, the rollout and implementation of IMS will take quite a long time. As a result, until the IMS VoIP becomes fully available on most mobile networks, an interim solution to provide the voice service to users of the LTE system needs to be provided. The Circuit-Switched FallBack (CSFB) mechanism is one of the interim solutions. When a voice call for a UE is started, CSFB directs the UE from the LTE system to a legacy system such as UMTS, GSM, CDMA2000, or 1xRTT where the UE is provided with a conventional Circuit-Switched (CS) based voice call. The Single Radio Voice Call Continuity (SRVCC) mechanism is another interim solution, which enables a switch from the IMS VoIP service to a CS-based voice call over legacy systems.

7.2 Home eNB (HeNB)

There are two types of eNB: Macro eNBs and Home eNBs. A Macro eNB is deployed by mobile network operators, can be used by all subscribers of the operators, and serves wide areas. On the other hand, a Home eNB typically serves small areas of tens of meters in radius and uses low output power. The target of the HeNB is enterprises that want to provide wireless connection services to any user or a group of specific users, or residential customers who want to use wireless connection services in small areas such as their homes. The group of users allowed to access the HeNB is called the Closed Subscriber Group (CSG) and a cell belonging to an HeNB is called a CSG cell. An HeNB can be deployed even within the coverage of Macro eNBs, as illustrated in Figure 7.1.

Functional features for HeNBs have been added to the LTE specifications over several releases. For the Release 8 specification of an HeNB, only essential and basic functions are supported, so that the HeNB can provide preferential services to the group of subscribed users. In other words, for Release 8 HeNB functionality, only RRC_IDLE mobility is

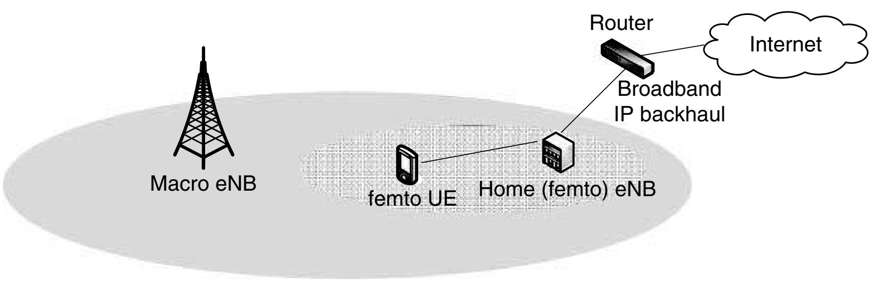

Figure 7.1 HeNB deployment

supported. How users belonging to the CSG should behave in RRC_IDLE is defined in this release. In the Release 9 HeNB specification, support for RRC_CONNECTED mobility is included. Thus, a UE of this release can move from a cell in a Macro eNB to a cell in an HeNB without RRC connection release.

Basically, a CSG cell is used by UEs belonging to the CSG. However, the CSG cell sometimes allows access by UEs that do not belong to the CSG. For example, there are cases where a CSG cell is deployed on the edge of areas covered by cells of Macro eNBs, or in areas where capacities provided by Macro eNBs are not enough to meet the demand. In this case, if a CSG cell can be used to serve UEs not belonging to the CSG, the CSG cell can increase the operator's service area or can offload the burden on the Macro eNB. Thus, from Release 9, a CSG cell can also serve UEs that are not members of the CSG.

When many CSG cells are deployed in an uncoordinated way, problems such as interference between the cells and the impact on measurement performance arise. Work to mitigate these problems is ongoing in 3GPP.

7.3 Public Warning System (PWS)

Japan is one country where disasters like earthquakes and tsunamis frequently inflict huge damage on society. To minimize the losses caused by such disasters, it is important to have a system through which emergency information can be distributed to the people as quickly as possible. For this purpose, the Japanese government set a requirement on the mobile wireless system to deliver information about urgent and important events. To support this requirement, a feature called the Earthquake and Tsunami Warning System (ETWS) was included in the Release 8 specification. With the ETWS feature, information about earthquakes and tsunamis can be delivered to UEs within 4 seconds.

On the other hand, the US government assumed different scenarios for emergency notification. The goal of the Federal Communication Commission (FCC) was to alert citizens of emergency situations such as child abduction, terrorist threats, or severe weather through the use of wireless networks. To support this requirement, a feature called the Commercial Mobile Alert Service (CMAS) was included in the specification from Release 9.

In addition, the Korean government has set similar requirements for the National Emergency Management Agency (NEMA). These requirements will be the basis for the new Release 10 feature called the Korean Public Alert System (KPAS).

Rather than including new features whenever a different country sets new requirements for an alert system, it would be better to have one generic mechanism for a warning system that could be adapted to each region's different requirements. This generic mechanism to deliver emergency information is called the Public Warning System (PWS) and was introduced from Release 9.

7.4 Multimedia Broadcast/Multicast Service (MBMS)

Like any other business, mobile operators seek out new business opportunities in addition to their fundamental business such as voice services and Internet connection services. As the number of handsets with multimedia support increases, operators see multicast and broadcast services as a potential new business area. To meet the operators' demand for providing broadcast and multicast services such as mobile TV services over the mobile wireless network, a new feature called the Multimedia Broadcast/Multicast Service (MBMS) was

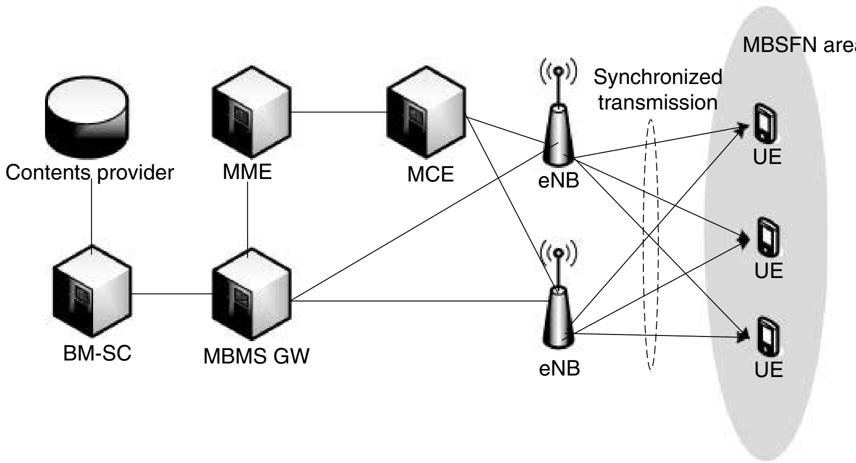

Figure 7.2 MBMS structure and transmission

introduced in the Release 6 specification of UMTS. At the beginning of LTE, the plan was to include the MBMS in Release 8. However, the schedule to complete the Release 8 specification was so demanding that the inclusion of the MBMS feature was delayed until Release 9.

The network architecture for the provision of the MBMS is illustrated in Figure 7.2.

For the delivery of MBMS contents over the radio interface, the MBMS Single Frequency Network (MBSFN) technique is used to transmit identical data simultaneously to multiple cells. A Multicast CHannel (MCH) is used as a transport channel for the delivery of MBMS data and MBMS control information.

In the Release 10 specification, the MBMS counting mechanism, which is used to find the number of UEs wanting to receive an MBMS service, is included to increase radio resource utilization efficiency. In the Release 11 specification, work is ongoing to deliver the MBMS service continuously even if the UE moves across the cell.

7.5 Carrier Aggregation (CA)

To meet the data rate requirement set by the ITU for IMT-Advanced, inevitably support is needed for transmission and reception over a wide bandwidth. Accordingly, 3GPP has set the supportable target bandwidth for LTE-Advanced as 100 MHz, much wider than the supported maximum bandwidth of Release 8, which is 20 MHz. However, because the spectrum resource in the low frequency range is scarce, it is not easy to allocate contiguous 100 MHz bandwidth for a mobile wireless network. As a result, a feature called Carrier Aggregation (CA) has been introduced. As the name suggests, when CA is used, multiple carriers are used simultaneously to provide a wider transmission bandwidth up to 100 MHz. This is illustrated in Figure 7.3.

Since a cell is composed of downlink and uplink carriers, carrier aggregation is also known as cell aggregation. In CA, a UE can be configured with one primary cell called a PCell and up to four auxiliary cells called SCells. The PCell provides basic control of the UE such as RRC connection management, radio bearer management, mobility management,

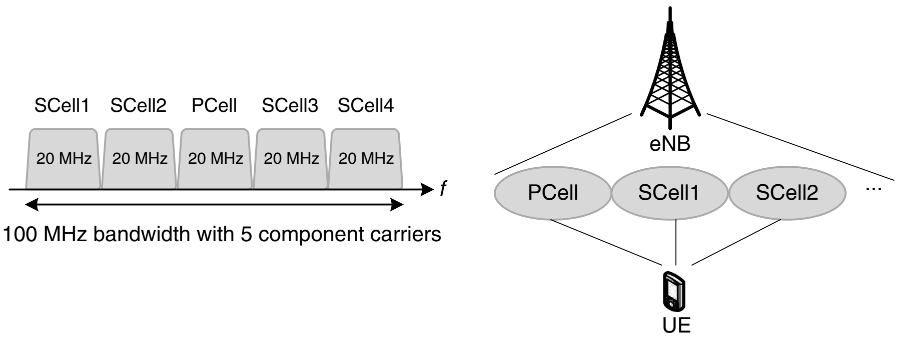

Figure 7.3 Carrier Aggregation

security management, and so on. The SCells serve as additional radio resource to the UE, the addition or removal of which depends on the amount of traffic.

For CA in Release 10, configured carriers for the uplink belong to the same frequency band. However, for CA in Release 11, this restriction is removed. Thus, management of multiple uplink timing is required for CA in Release 11.

7.6 Relay

The Relay technique enables the extension of cell coverage by the use of an additional network node called the Relay Node (RN). Positioned between the eNB and the UE, the RN relays data from the UE to the eNB and vice versa. In addition to CA, Relay is regarded as a key technology introduced in LTE-Advanced.

During the standardization of the RN feature, there was intensive discussion about which protocol layers the RN should support. In the end, it was decided that the RN would be equipped for all radio protocol layers, which means that relaying is performed at the IP packet level. In addition, it was decided that the frequency used for communication between the RN and the eNB would be either the same as or different from the frequency used between the UE and the RN.

The RN plays an eNB role to UEs under its coverage. For the RN to act as an eNB, the RN should be connected to an eNB. The eNB serving the RN is called the Donor eNB (DeNB). The backhaul link between the RN and the DeNB is referred to as the *Un interface*. Figure 7.4 shows a simplified view of the network architecture supporting the RN.

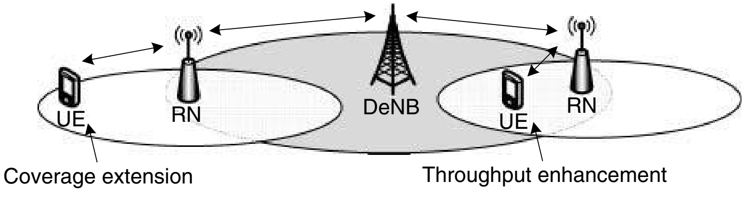

Figure 7.4 Relay

Deploying an RN instead of an eNB is an attractive option for the operator, in that the RN can be deployed at lower cost than an eNB. Because the RN can be repositioned easily and does not require any fixed wire-line connection, operators can install an RN with low CAPEX (Capital Expenditure) and run it with low OPEX (Operating Expenditure). For example, for a hotspot area where additional capacity is required temporarily or for an area where a DeNB cannot be deployed, an RN can be used easily.

For the Release 10 specification, the functionality is limited to a fixed one-hop RN. This restriction will be removed in future releases.

7.7 Minimization of Drive Test (MDT)

When the mobile operator installs new equipment for its network, it needs to verify that the equipment is working as planned. For example, the operator needs to check whether the cell coverage provided by the new equipment is sufficient. In addition, the operator needs to optimize the performance of its network through adjustment of the network setting parameters. To meet these requirements, the operator needs to perform actual measurements to know whether the network is available and reliable, whether the provided capacity is sufficient, and whether the provided QoS is appropriate at each point within its intended coverage area. To perform these measurements, operators run a drive test where vehicles equipped with heavy measurement devices go through every corner of the network coverage. The drive test is an essential part of the network operation process. However, a manual drive test demands huge operating costs from the operator and it cannot be used to check indoor coverage.

With the Minimization of Drive Test (MDT) feature, operators can get lots of measurement results without performing a manual drive test. Because UEs move around, if the UE performs relevant measurement along the path of its movement, the measurement results from each UE can be a valuable input to the network optimization process. Thus, the MDT feature defines what should be measured and reported by the UE to minimize the manual drive test performed by the operator.

In the Release 10 specification, the MDT feature for coverage optimization to detect a coverage hole is included. In the Release 11 specification, work is ongoing to support the QoS verification scenario to check whether the required QoS is being provided to the UE or not.

7.8 Enhanced Inter-Cell Interference Coordination (eICIC)

A Heterogeneous Network (HetNet) is a network where the coverage of pico cells or femto cells overlaps with the coverage of macro cells in a certain geographical area. In that area, if the pico cells and the femto cells use the same frequency as the macro cells, significant interference may occur, as illustrated in Figure 7.5.

For example, for a pico cell, the UE served by the pico cell may suffer heavy interference from a macro cell and sometimes experience RRC connection failure. To prevent RRC connection failure, the pico cell should be deployed with a conservatively small coverage area. For a femto cell (i.e., CSG cell), a UE is connected to a macro cell instead of a femto cell because the UE is not a member of a CSG. In this case, the UE served by the macro cell may experience heavy interference from the femto cell. The coverage area of the femto cell may

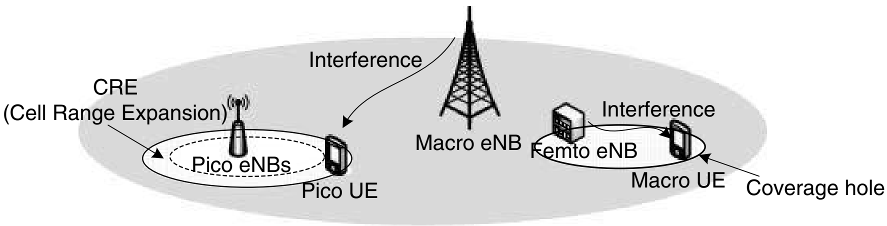

Figure 7.5 Heterogeneous network scenarios

be seen as a coverage hole to a non-member UE. One simple solution to prevent these problems is to allocate different frequencies to pico cells and femto cells. However, this solution is not optimal because frequency resource is very costly.

For the Release 10 specification, two mechanisms for the enhanced Inter-Cell Interference Coordination (eICIC) feature have been introduced. The first is CA-based ICIC and the second is time domain ICIC.

In the CA-based ICIC mechanism, the cross-carrier scheduling method can mitigate PDCCH interference. In other words, when multiple carriers are available for the operation, some carriers may experience severe inter-cell interference on the PDCCH and other carriers may not experience severe interference on the PDCCH. In which case, the PDCCH of the carriers with less interference is used for scheduling of data on the carriers with severe interference.

In the time domain ICIC mechanism, the aggressor cell that causes the dominant interference on other cells does not use the PDCCH in certain subframes, referred to as Almost Blank Subframes (ABSs). Since the interference from the aggressor cell in those subframes is mitigated, a victim cell that otherwise suffers heavy interference from the aggressor cell is able to schedule UEs in those ABSs.

7.9 Machine Type Communication (MTC)

A new communication scenario where a huge number of devices are interacting with one another, referred to as Machine-to-Machine (M2M) communication, is getting a lot of attention from the telecommunication industry. This is because the number of device communications used in M2M is expected to overwhelm Human-to-Human (H2H) communications. To make LTE competitive for the use of M2M, 3GPP has started to study M2M in Release 10 under the feature name of Machine Type Communication (MTC).

Although there are several use cases expected for M2M, 3GPP has limited the scope of the MTC feature to three use cases: smart metering, road security, and consumer electronics. Smart metering covers the scenario where devices collect data automatically about energy consumption or the quality of supply at the customer site and report the results back to the utility company for monitoring and billing purposes. Road security covers the scenario where devices automatically make logs of accidents or collisions and report essential information such as the location of the incident. Consumer electronics relates to the scenario where devices provide the user with services through the devices' automatic connection.

The main challenge of the MTC feature to radio protocols is assumed to be overload control because an enormous number of devices may establish RRC connections almost at the same time. Therefore, dispersing simultaneous access from devices was considered seriously in Release 10. Work is ongoing in Release 11 in order to enable the network to prevent access selectively from devices.

8

Voice over LTE (VoLTE)

In the mobile business, the telephony service has been the major source of revenue for most mobile operators. Even with the proliferation of smartphones, the voice service is still considered important by users. For the success of LTE, it is of crucial importance to maintain support for the voice service over the LTE network.

8.1 Voice Solutions for LTE

8.1.1 Ultimate Voice Solution

The open architecture of the Internet Protocol (IP) created the flexibility that has enabled a thriving multitude of multimedia services to be available over the Internet. To integrate mobile access with the IP multimedia services, 3GPP has defined an architectural framework, called the IP Multimedia Subsystem (IMS), to deliver IP multimedia services over 3GPP systems. The success of the IMS in 3GPP motivated other organizations in the same industry, such as 3GPP2, to adopt the IMS as the architectural framework for IP multimedia services. The IMS also supports integration with the wired network, creating the potential for fixed–mobile convergence.

The mobile industry has been motivated by the success of the IMS standard, and the richness and flexibility of the IMS architecture provides a reliable framework to enable fruitful multimedia services including a voice service with guaranteed quality of service. The IMS is considered by most operators to be the ultimate architecture to provide a voice service over the LTE network. Figure 8.1 illustrates the concept of a voice service based on the IMS architecture deployed over the LTE network. For the sake of convenience, a voice service based on the IMS architecture is referred to as “IMS VoIP” hereafter.

To minimize implementation options in IMS VoIP, the industry collaborated to define a common solution for IMS VoIP by organizing the One Voice Initiative in November 2009. The One Voice Initiative produced a “One Voice; Voice over IMS profile” that captures a minimum mandatory set of IMS and other features for UEs and networks to support for voice and other services, for example, SMS. The work of One Voice was strongly supported by the GSM Association (GSMA) and taken over by the “GSMA VoLTE Initiative” that was

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

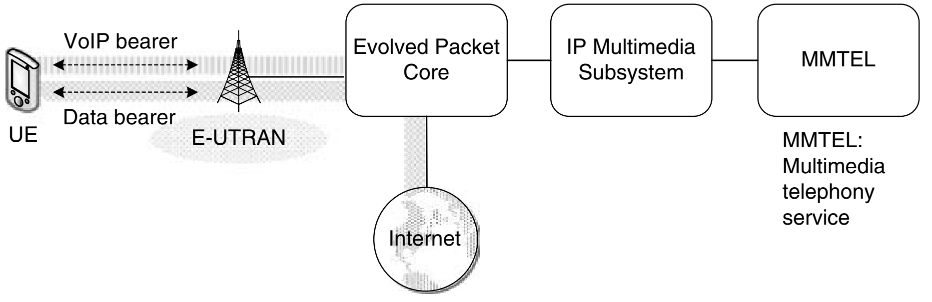

Figure 8.1 Concept of IMS VoIP over the LTE network

established to promote a common voice solution over LTE. The GSMA VoLTE Initiative published an official document (IR.92) which defines the IMS profile for voice and SMS.

8.1.2 Interim Voice Solutions

Given today's mega trend of IP convergence, one would expect that, in years to come, IMS VoIP will be universally available over both the wired network and the wireless network including LTE.

Even though the LTE system (E-UTRAN and EPC) supports only packet-switched (PS) IP services and the IMS VoIP solution has been agreed as the industry-wide solution for the voice service, the industry cannot rely solely on the IMS VoIP solution.

There are legacy networks and legacy mobile devices that can only work on the legacy networks with legacy solutions (i.e., CS voice telephony), and the advent of LTE cannot replace the entire legacy system on a grand scale on the same day. Recalling the enormous investment already made in the legacy networks, some operators may be reluctant to evolve their legacy networks towards an LTE network in a hurry. In general, different operators will have quite different network evolution strategies, which will result in a heterogeneous environment where the LTE network will coexist with legacy networks for quite some time. Even though it is an essential requirement that service continuity between the networks and the quality of service for voice telephony should be guaranteed, the fundamental difference between the LTE network and legacy networks makes satisfying this essential requirement a challenge.

The roaming scenario adds further complexity to the support for only IMS VoIP on the track of network evolution towards LTE. For example, while supporting only IMS VoIP may work on the home LTE network, it may not work on a visited network where IMS VoIP is not supported or only a legacy network is available.

As per IMS deployment, it may take a non-trivial amount of time for IMS infrastructure and IMS-supporting UEs to be widely available in the target market. Nonetheless, LTE deployment should not be delayed until IMS deployment is fully completed over all service areas, as LTE networks are likely to be used to provide only data services in the early phase of the deployment. As a result, a voice service other than the IMS VoIP should be provided over legacy networks in the initial LTE deployment phase.

Even though IMS infrastructure and/or UEs capable of IMS VoIP are widely available, the LTE network of one operator may cover only limited geographical areas in the initial phase of LTE deployment, while legacy 2G/3G coverage will provide access in all areas covered by the operator. Because the voice service should not be suspended as a result of the unavailability of LTE coverage, the issue of how IMS VoIP should be handled at the border area of LTE coverage needs to be resolved.

In short, it should be understood that supporting only a simple voice solution like IMS VoIP is insufficient in practical terms, and interim solutions in addition to the IMS VoIP service will need to be used until the IMS VoIP solution becomes mature on a global scale.

The mobile industry and standard groups considered the following interim solutions:

- Circuit-Switched FallBack (CSFB);
- Voice over LTE via Generic Access (VoLGA).

Figure 8.2 illustrates CSFB and VoLGA. CSFB is the standardized solution that provides a voice service over the existing 2G/3G circuit-switched (CS) network. The principle of CSFB is that the UE is normally camping on LTE, but it is moved to a 2G/3G CS network when a voice service is to be initiated. CSFB is widely supported by many operators as it requires minimum changes to legacy 2G/3G networks, and CSFB fits well with the operators' strategy that 2G/3G systems will provide voice services and LTE systems will provide data services only in the early phase of LTE deployment.

Another proposed solution was VoLGA, which extends the existing Generic Access Network (GAN) to support voice over the LTE network. GAN was designed originally to provide mobile services to the UE via supporting both the 3GPP radio interface and the Wi-Fi radio interface. A GAN gateway provides a secure connection to deliver circuit-switched services to a subscriber via a Wi-Fi access network. The key concept of VoLGA is to replace the Wi-Fi access network with an LTE access network in the GAN architecture. VoLGA did not get much support from the mobile industry and standard groups.

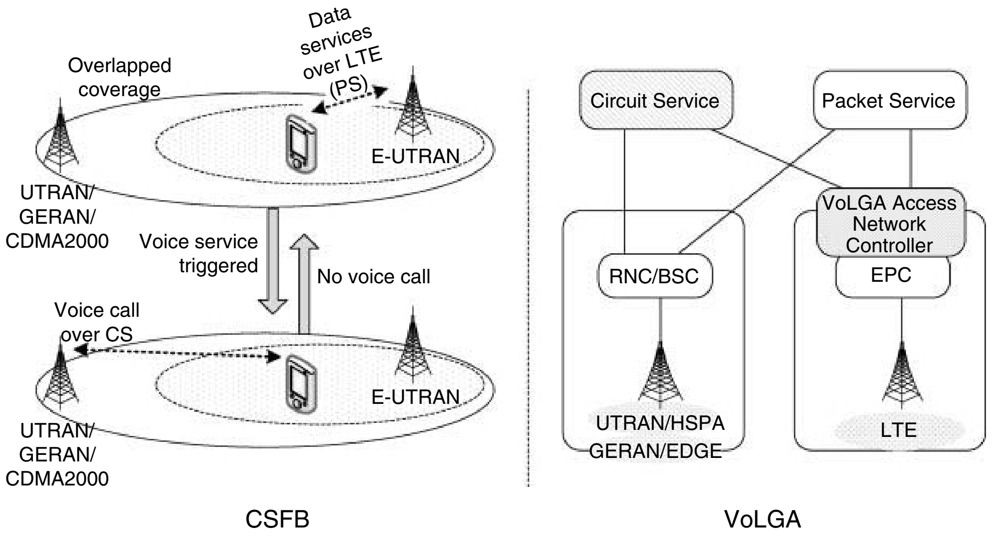

Figure 8.2 CSFB and VoLGA

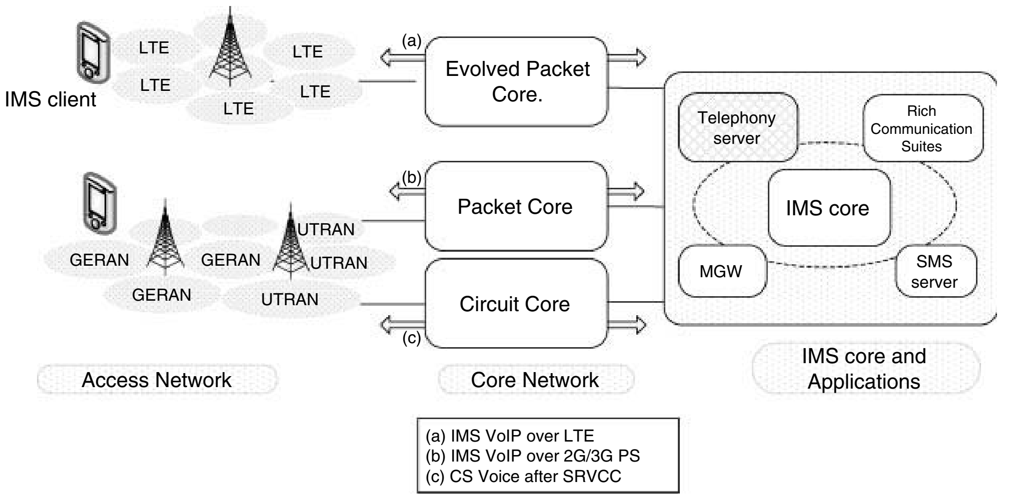

Figure 8.3 Concept of IMS VoIP

8.2 IMS VoIP

The IP Multimedia Subsystem (IMS) is a standardized architectural framework to provide IP-based services. The IMS supports IP multimedia applications including voice, messaging, video, conferencing, and blended multimedia services. There are advantages to using the IMS framework for voice services:

- global roaming with fewer interoperability problems is ensured;
- full performance of the LTE during voice calls is sustained;
- high-definition voice services can be offered readily;
- blended services with video/conferencing/data can be enabled gracefully.

The IMS requires new IMS-specific network elements as part of the dedicated core network architecture. Figure 8.3 illustrates the concept of IMS VoIP over 3GPP networks. Note that IMS VoIP also accommodates 3GPP2 networks, which are not shown in the figure for simplicity.

8.2.1 IMS Profile

Even though the IMS can enable rich multimedia services including a voice service on an IP-based network, the rollout of the IMS has been delayed by many mobile operators. The main reason for the delay in IMS deployment is that the architecture of the IMS is very complex as a result of many architectural options depending on target services and the access/core network in question. If different IMS-based voice solutions were implemented by different operators, it would not be possible to avoid the risk of LTE market fragmentation and interoperability issues. This would also result in a situation where the chance of roaming for subscribers could be quite limited depending on the solution adopted by the operator.

8.2.1.1 One Voice Initiative

Faced with the practical threat of fragmented IMS deployment, the industry recognized the importance of an industry-wide common solution to promote a viable ecosystem for an IMS-based voice service and to enjoy the benefits of economies of scale. The principle of introducing a common standard solution for IMS VoIP is also important to ensure seamless global roaming experiences for subscribers. Motivated by the necessity of a global working standard solution, the industry established the “One Voice” Initiative in November 2009.

The One Voice Initiative adopted the IMS architecture defined by 3GPP for provisioning a voice service and SMS over LTE networks. With the aims of accelerating the deployment of a common solution for an IMS-based VoIP service and avoiding complexity in the initial phase of IMS deployment, the One Voice Initiative produced a “One Voice Profile”. As the word “One” in the name suggests, the objective was to define a single profile for an IMS VoIP service rather than multiples to reduce IOT efforts and implementation costs by defining a de facto solution. The “One Voice Profile” defined a minimum mandatory set of features that should be implemented by UEs and networks to support target IMS services over LTE based on 3GPP specifications.

8.2.1.2 IMS Profile for Voice and SMS

In February 2010, the GSMA VoLTE Initiative was established with the backing of many companies and organizations to promote voice and messaging services over LTE. The companies endorsing the GSMA VoLTE Initiative were supporters of a common IMS-based voice solution for the next generation of mobile networks. The work of the One Voice Initiative was adopted gracefully by the GSMA VoLTE Initiative as the baseline for further work.

The GSMA published the “IMS Profile for Voice and SMS”, which defines a profile specifying a minimum mandatory set of features based on 3GPP specifications for UEs and networks. The IMS Profile is intended to establish a clean ecosystem for IMS-based telephony services over LTE radio access, ensuring high quality and interoperability. Figure 8.4 shows the protocol stacks for the “IMS Profile for Voice and SMS”.

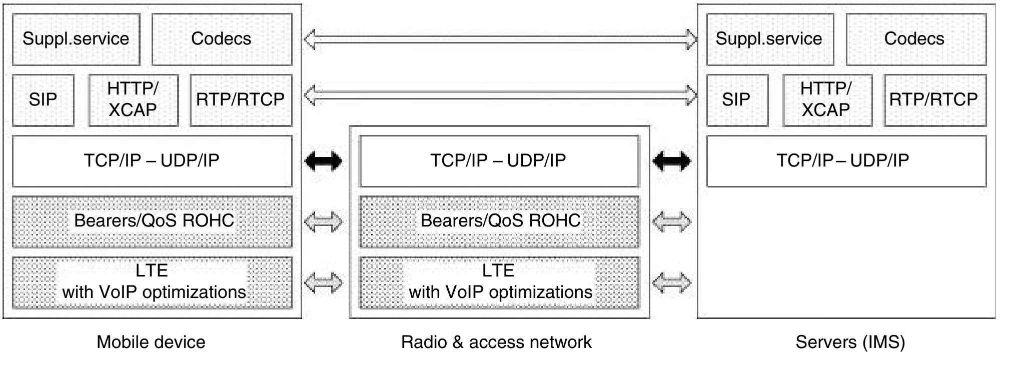

Figure 8.4 Protocol stack for IMS Profile for Voice

The key components of the "IMS Profile for Voice and SMS" are summarized as follows:

- IMS core network with Telephony Application Server (TAS);
- IMS client features at the UE;
- support of Single Radio Voice Call Continuity (SRVCC) for IMS VoIP coverage extension;
- AS features for VoIP support.

The generic IMS functions required for IMS VoIP and IMS VoIP call procedures are beyond the scope of this book.

SRVCC enables the transfer of an IMS voice session from PS to CS. SRVCC is useful for seamless provision of a voice service when IMS voice service delivery over PS is not possible due to PS coverage limitations.

Several AS features help to optimize support for VoIP features, including ROHC, TTI bundling, and SPS.

8.2.1.3 Bearer Characteristics for "IMS Profile for Voice and SMS"

According to the profile, the EPS bearer for the voice service is configured as a Guaranteed Bit Rate (GBR) bearer with a QoS Class Identifier (QCI) value set to 1. The radio bearer for the voice service is configured with UM RLC to reduce overhead while tolerating a small loss of voice data.

8.2.2 Single Radio Voice Call Continuity (SRVCC)

SRVCC provides continuity of a voice call from IMS VoIP over a PS network to CS voice over a CS network for a UE supporting transmission/reception on only one radio at a time. SRVCC is useful for operators with a packet-switched network that covers a limited area and a circuit-switched network that covers the entire service area of the operator. One of the target scenarios for SRVCC is a UE configured with a voice bearer in LTE losing LTE coverage and entering CS coverage. To ensure continuity of the voice call outside of the LTE coverage in this case, the flow of voice data anchored in IMS is transferred from the LTE network to the CS network. SRVCC allows operators to upgrade their 2G/3G networks gradually to LTE networks rather than having to replace their whole networks with LTE all at once.

SRVCC is applicable for voice call continuity from any PS network (e.g., E-UTRAN or HSPA) to any CS network (e.g., GERAN, UTRAN, or CDMA2000). Figure 8.5 shows the network architecture for SRVCC from E-UTRAN to UTRAN/GERAN. The SRVCC procedure is based on the session transfer function defined in [1] for the IMS part, Sv for the interconnection between PS and CS, and the inter-RAT mobility procedure defined in [2] for the access stratum part. The handover from PS voice to CS voice is enabled by utilizing the SRVCC PS to CS procedure via the Sv interface between the MME and the MSC server.

In Figure 8.6, relocation of the voice bearer from a PS network to a CS network through SRVCC is illustrated.

SRVCC from E-UTRAN to UTRAN/GERAN is supported from Release 8. Support for voice continuity in the reverse direction – that is, SRVCC from UTRAN/GERAN to E-UTRAN – will be supported from Release 11.

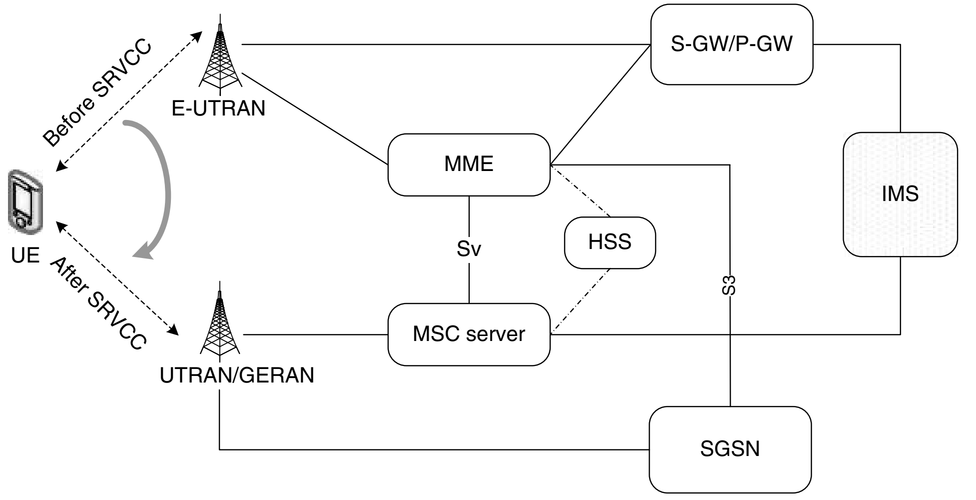

Figure 8.5 Network architecture for SRVCC. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

Figure 8.7 shows a simplified illustration of the SRVCC procedure from E-UTRAN to UTRAN. A UE capable of SRVCC indicates its SRVCC capability as part of the UE capability information during an NAS procedure such as the Attach procedure or the Tracking Area Update procedure. If both the UE and the MME support SRVCC, the MME indicates support for SRVCC to the eNB during the initial context setup over S1 for the UE.

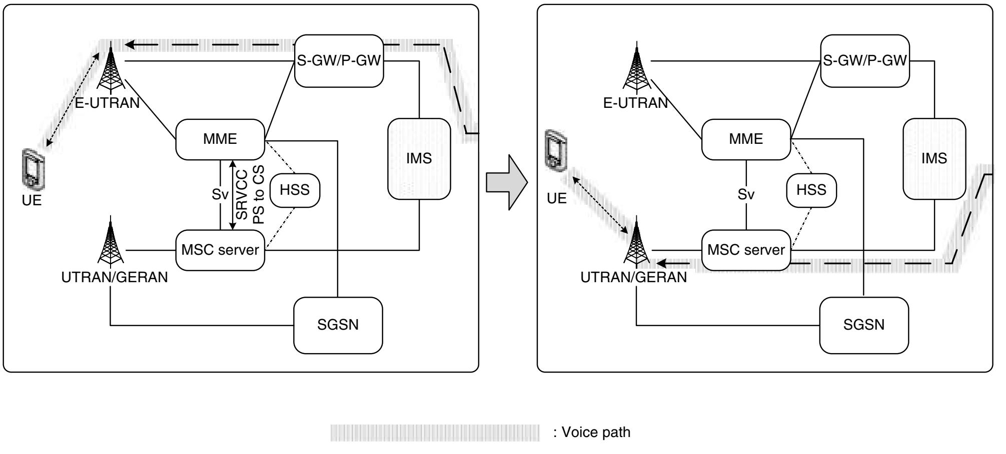

Figure 8.6 Voice path before and after SRVCC. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

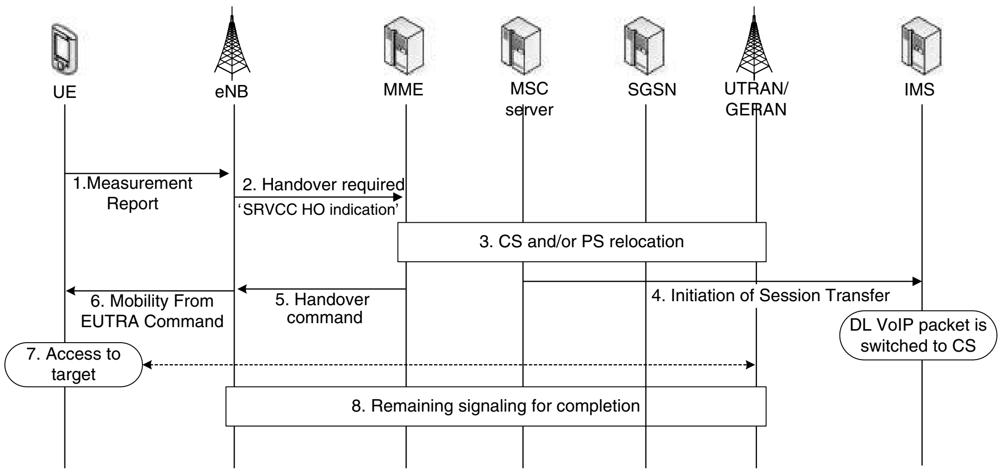

Figure 8.7 SRVCC handover from E-UTRAN to UTRAN/GERAN. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

- **Step 1:** The UE sends a *MeasurementReport* message including measured results of other RAT to its serving cell.
- **Step 2:** The eNB, receiving the *MeasurementReport* message, decides whether SRVCC should be triggered based on whether any voice bearer, for which QCI is set to 1, is established. If SRVCC is triggered, the eNB requests an SRVCC handover to the MME by sending Handover Required including an SRVCC handover indication.
- **Step 3:** The MME, receiving the SRVCC handover indication, splits the voice bearer from other non-voice bearers. The MME then initiates relocation of the voice bearer and other non-voice bearers towards the MSC server and SGSN, respectively. The details of relocation can be found in [3]. During this step, radio resources for the bearers are assigned for the UE by the RNC of the target cell. Upon completion of relocation, the MME receives acknowledgments for the relocations from the MSC server and SGSN.
- **Step 4:** The MSC server, upon receiving the acknowledgment for the relocation request for the voice bearer from the target MSC in step 3, initiates a session transfer towards the IMS. This step is carried out during step 3.
- **Step 5:** The MME, upon receiving acknowledgments both for the relocation of the voice bearer and the non-voice bearers, sends a handover command to the eNB. The handover command includes the information (Target to Source Transparent Container) received from the SGSN and MSC server during step 3.
- **Step 6:** The eNB, upon receiving the handover command, sends to the UE a *MobilityFromEUTRACommand* message.
- **Step 7:** The UE, upon receiving the *MobilityFromEUTRACommand*, accesses the target cell.
- **Step 8:** Further signaling for completion of the SRVCC handover is exchanged, for example, further confirmations between the nodes, release of resources, and so on. Details can be found in [3].

8.3 Circuit-Switched Fallback (CSFB)

Circuit-Switched FallBack (CSFB) is the mechanism, defined in Release 8, to enable a voice service by reusing an existing CS network for UEs normally camping on LTE for IP services. If CSFB is used when a voice service is to be initiated, the UE performs radio switching from LTE to a legacy 2G/3G CS network where a CS voice call can be established. CSFB is applicable to UEs supporting both the LTE network and CS networks such as GERAN, UTRAN, or CDMA2000 1xRTT. From a network point of view, the coverage of a legacy CS network is much greater than the LTE coverage, such that a UE can be directed to one of the available legacy CS networks when CSFB is initiated. The general architectural enhancement to enable CSFB and relevant functionalities/procedures is specified in [4].

CSFB is initiated by a UE sending a particular NAS message, called an extended service request, to the MME. The extended service request may be sent in response to a paging message that was originated in the CS domain and routed to E-UTRAN via the MME in the case of a mobile terminating call. Upon receiving the extended service request for CSFB, the MME requests the eNB to trigger an inter-RAT mobility procedure. The details of CSFB procedures are explained in the subsequent sections.

The inter-RAT mobility procedures applicable for CSFB include redirection, handover, and cell change order:

- **Redirection:** The redirection procedure can be performed by the eNB sending an *RRCConnectionRelease* message including redirection information (see Section 3.12). The redirection information indicates the RAT/frequency to which the UE needs to move. Upon receiving redirection information, the UE releases the RRC connection with E-UTRAN, performs cell selection to camp on a cell on the indicated RAT/frequency, and makes a connection with the selected cell. The redirection procedure is applicable for CSFB to GERAN, UTRAN, and CDMA2000 1xRTT.
- **Handover:** Handover to the other RAT is performed by the eNB issuing a *MobilityFromEUTRACommand* message as a handover command. The handover command includes the radio resource configuration to be used in the target cell. If a UE receives a handover command, it attempts to access the target cell using the radio resource configuration included in the handover command. Handover to the other RAT is applicable for CSFB to GERAN, UTRAN, and CDMA2000 1xRTT. If the *MobilityFromEUTRACommand* message is sent for CSFB to GERAN or UTRAN, it includes a CSFB indicator for use in NAS to handle a handover failure.
- **Cell Change Order:** Cell Change Order (CCO) can also be used for CSFB. CCO is applicable only for CSFB to GERAN. CCO is signaled to the UE via a *MobilityFromEUTRACommand* including the information relevant for CCO.

The mobility options applicable for CSFB to UTRAN/GERAN are summarized in Table 8.1.

Note that the RRC connection in E-UTRAN is not maintained after CSFB. This means that, while the UE is provided with a voice service in the legacy RAT, IP services may be degraded or even suspended, depending on the support for IP services in the legacy RAT. To provide a better quality of IP services to the UE, the legacy RAT may decide to perform inter-RAT mobility to E-UTRAN for the UE as soon as possible

Table 8.1 Possible mobility options for CSFB to UTRAN/GERAN. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

No	Mobility Option (Related RRC Messages)	Target RAT	Release	UE Capability
1	Redirection (*RRCConnectionRelease*)	UTRAN	Release 8	Mandatory for UEs supporting CSFB to the UTRAN
2	Redirection with system information (*RRCConnectionRelease*)	UTRAN	Release 9	Indication of the *e-RedirectionUTRA* in the UE capability
3	Redirection (*RRCConnectionRelease*)	GERAN	Release 8	Mandatory for UEs supporting CSFB to the GERAN
4	Redirection with system information (*RRCConnectionRelease*)	GERAN	Release 9	Mandatory for UEs supporting CSFB to the GERAN
5	PS handover with DRB(s) (*MobilityFromEUTRACommand*)	UTRAN	Release 8	Mandatory for UEs supporting CSFB to the UTRAN
6	PS handover (*MobilityFromEUTRACommand*)	GERAN	Release 8	Indication of the *interRAT-PS-HO-ToGERAN* in the UE capability
7	Cell change order with the NACC (*MobilityFromEUTRACommand*)	GERAN	Release 8	Mandatory for UEs supporting CSFB to the GERAN
8	Cell change order without the NACC (*MobilityFromEUTRACommand*)	GERAN	Release 8	Mandatory for UEs supporting CSFB to the GERAN

after the end of the voice service. No UE autonomous action for reverting to E-UTRAN is defined.

Compared to the setup time for a voice call within a legacy CS network from scratch, CSFB introduces more delay to establish a voice call, mainly due to radio switching. In Release 9, some features were added to enhance CSFB performance.

A UE that is capable of CSFB to UTRAN indicates support for UTRA FDD or TDD and the supported band list in the UE capability signaling. A UE capable of CSFB to GERAN indicates support for GERAN and the supported band list in the UE capability signaling.

8.3.1 CSFB to UTRAN or GERAN

8.3.1.1 Reference Architecture

Figure 8.8 shows the network architecture for CSFB to UTRAN/GERAN. The MME and MSC server are interconnected via the SGs interface, over which signaling exchange and the delivery of CS paging (paging that originated from a CS network) to the LTE network are performed.

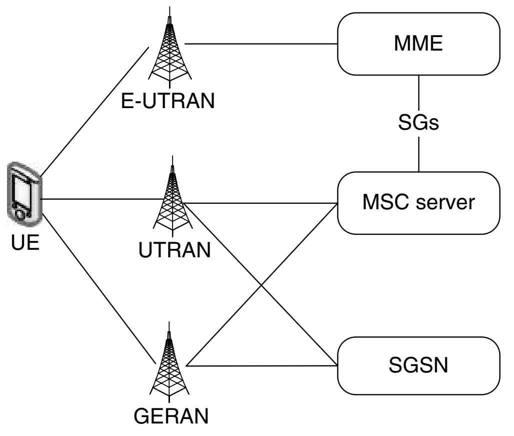

Figure 8.8 Network architecture for CSFB to UTRAN/GERAN

8.3.1.2 Combined Registration to the LTE Network and CS for CSFB

In CSFB, since a voice service is delivered over 2G/3G CS networks, the UE should first be registered to both the LTE network and the 2G/3G CS network to enable CSFB. Registration to both the LTE network and the CS network is done efficiently with a "combined" registration procedure, rather than performing two registration procedures independently. The combined registration can be applied to the Attach procedure and the Location Area Update procedure in the NAS layer.

With regard to the combined registration, the mapping between tracking area and location/routing area is managed in the MME, which enables the MME to trigger CSFB towards the proper cell of the CS network.

8.3.1.3 CSFB for a Mobile Originating Call

Figure 8.9 shows a simplified illustration of a mobile originating call with CSFB procedures. Each step of the procedure is described below:

- **Step 1:** If the UE cannot use IMS VoIP in LTE and the UE is attached to both LTE and CS networks, it triggers CSFB for a mobile originating call by sending an extended service request to the MME.
- **Step 2:** The MME, receiving the extended service request, triggers the eNB to perform the CSFB-related inter-RAT mobility procedure from E-UTRAN to GERAN or UTRAN by using an S1AP message. In this S1AP message, a CSFB indicator and the Location Area Identity (LAI) to which the UE is registered are included.
- **Step 3:** The eNB, receiving the message triggering CSFB, triggers an inter-RAT mobility procedure to GERAN or UTRAN. There are many options for carrying out inter-RAT mobility: PS handover, redirection, and cell change order. Depending on the selected mobility procedure, the eNB sends a relevant command for inter-RAT mobility. A *MobilityFromEUTRACommand* message is sent to the UE for handover or cell change order,

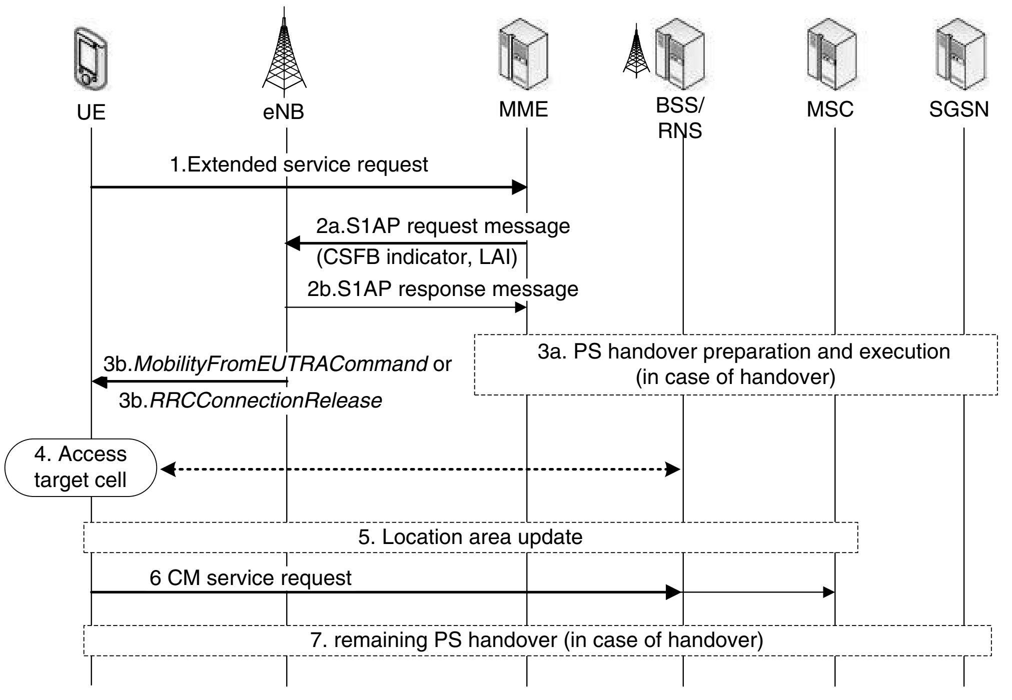

Figure 8.9 Mobile originating call with CSFB to UTRAN/GERAN. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

and an *RRCConnectionRelease* message including redirection information is sent for redirection.

- **Step 4:** The UE, receiving the inter-RAT mobility command from the eNB, performs an inter-RAT mobility procedure. If the UE receives a command for handover, it synchronizes and accesses the target using the radio resource configuration included in the handover command. If the UE receives a command for cell change order or redirection, the UE selects a suitable cell on the RAT/frequency indicated in the command and attempts to access the cell.
- **Step 5:** If the new cell belongs to a location area that is different from the cell the UE has most recently registered to, the UE updates its location to the CS network.
- **Step 6:** If the inter-RAT mobility procedure is successful, the UE continues the CS voice call setup procedure by sending a Connection Management (CM) service request. In Release 10, the CM service request indicates that this is a call establishment as a result of CSFB.
- **Step 7:** In cases where PS handover is needed, the remaining steps to complete the handover are processed.

Figure 8.10 shows the change of routing path for data and voice due to a mobile originating CSFB call.

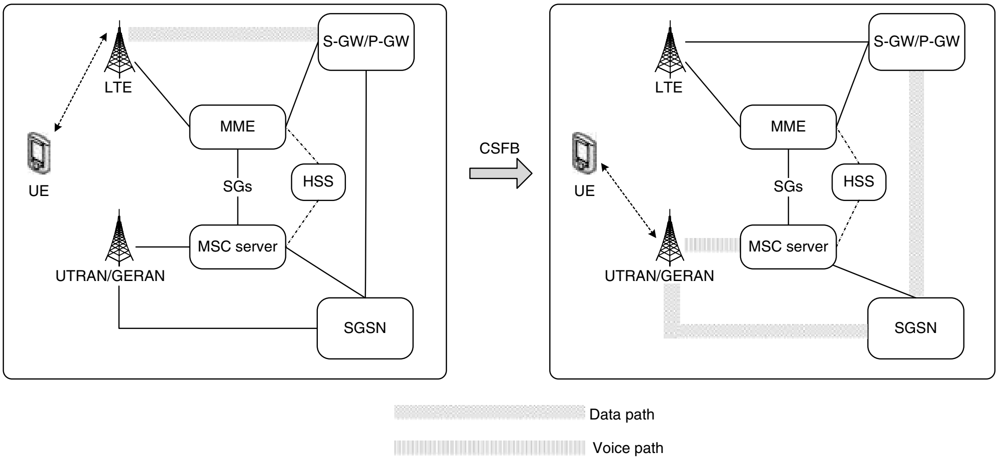

Figure 8.10 Change of voice and data path after CSFB

8.3.1.4 CSFB for a Mobile Terminating Call

The procedures for a mobile terminating call with CSFB to UTRAN/GERAN are shown in Figure 8.11. In the mobile terminating call, a paging procedure is added to the mobile terminating call. A mobile terminating call involving CSFB is initiated by the MSC server receiving an incoming voice call.

- **Step 1:** If the MSC receives an incoming voice call, it sends a paging request to the MME via the SGs interface. The MME, upon receiving the paging request from the MSC, pages the UE unless the UE is configured only for SMS. Note that the *Paging* message includes a CN domain indicator set to “CS”.
- **Step 2:** The UE, receiving the *Paging* message addressed to it, establishes an RRC connection if in RRC_IDLE and sends an extended service request for CSFB to the MME. The MME, receiving the extended service request, sends a service request over the SGs to the MSC. If the MSC receives a service request from the MME via the SGs, it stops retransmitting the paging request to the MME.
- **Step 3:** The MME then triggers the eNB to perform an inter-RAT mobility procedure for the UE by sending an S1AP message including a CSFB Indicator and LAI allocated to the UE. The S1AP message further includes UE capabilities to assist the eNB mobility procedure.
- **Steps 4–8:** For the rest of the process, the UE follows almost the same procedures as those shown in Figure 8.9, except for step 7, where the UE sends a paging response instead of a CM service request. If the MSC receives a paging response, it establishes a voice call.

8.3.1.5 Performance Enhancement of CSFB to GERAN/UTRAN

In Release 8, the basic mechanisms of CSFB were defined. Compared to the call setup time of a native voice call that originates and then terminates purely in the CS domain without radio switching, the CSFB mechanism introduces some additional delay when establishing a voice call. There are several causes contributing to such additional delay during CSFB:

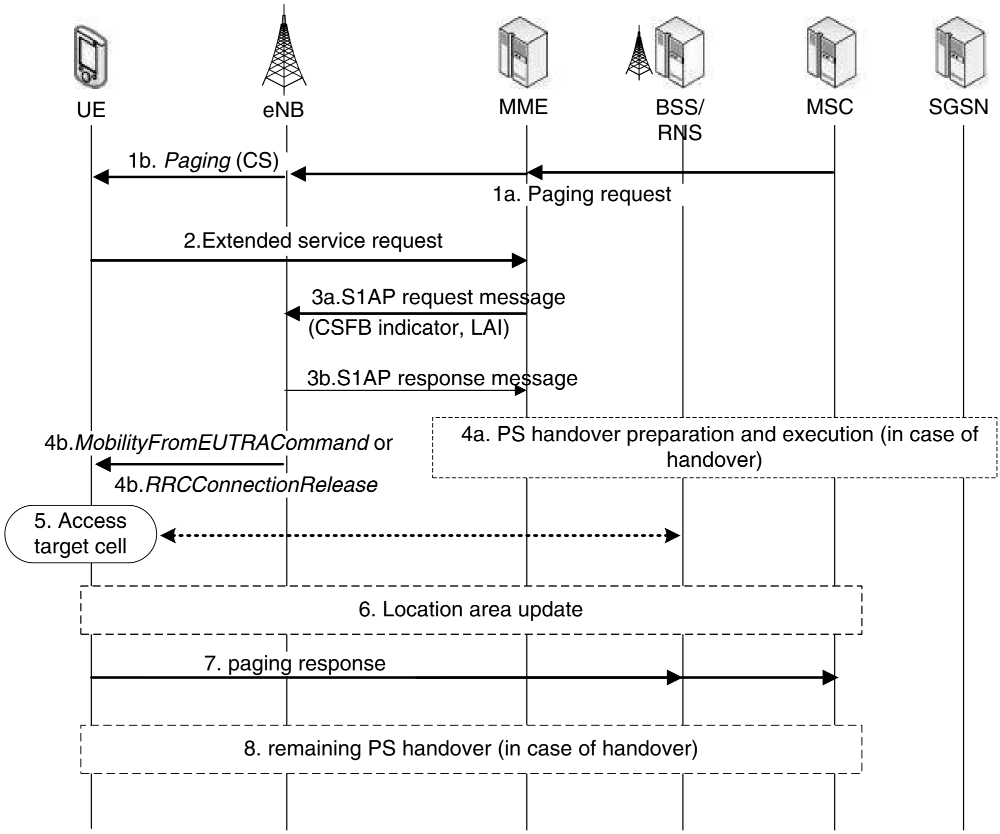

Figure 8.11 Mobile terminating call with CSFB to UTRAN/GERAN. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

- **Inter-RAT measurements:** If PS handover is used for CSFB, the UE may be requested by the eNB to perform inter-RAT measurements and reporting in order to select a suitable target cell. In this case, the inter-RAT measurement procedure introduces non-trivial delay.
- **System information acquisition of inter-RAT cell:** If redirection is used for CSFB, the UE needs to search for a suitable cell on inter-RAT. From the selected cell of the inter-RAT, the UE has to acquire the radio resource configuration parameters and the accessibility parameters from the system information. It takes noticeable time to read the system information of the selected cell, for example, $1 \sim 2$ seconds to acquire the relevant system information of a UTRAN cell, depending on the scheduling of the relevant system information on the radio interface.
- **Location area update at inter-RAT:** After executing inter-RAT mobility for CSFB, the NAS layer of the UE needs to trigger a Location Area Update procedure if the new cell belongs to an LAI that is different from that stored in the UE.

To enhance the performance of the CSFB procedure, Release 9 introduced an optimized redirection mechanism whereby an eNB can include the system information of one or multiple target cells (up to six) in the *RRCConnectionRelease* message. The system information can be utilized by the UE to accelerate the redirection procedure during CSFB, by allowing the UE to skip system information acquisition on the selected cell.

With regard to the possible delay incurred by inter-RAT measurements, the network may perform handover blindly (i.e., without a measurement report from the UE) to reduce the delay. However, blind handover may decrease the rate of handover success, which in turn increases the call drop rate.

The occurrence of a location area update during CSFB can be minimized if the network carefully performs coverage engineering jointly with the tracking area and the location/routing area. One example of coverage engineering is where the cell coverage of E-UTRAN/UTRAN/GERAN and the corresponding tracking/location/routing area are fine-tuned in such a way that whenever the location/routing area is changed, the tracking area is also changed.

8.3.2 CSFB to CDMA2000 1xRTT

3GPP also accommodates CDMA2000 1xRTT CS defined by 3GPP2 as one of the target CS domains for CSFB, and hence both mobile originating calls and mobile terminating calls via CSFB to CDMA2000 1xRTT CS are supported from Release 8.

8.3.2.1 Mechanisms of CSFB to 1xRTT

There are several mechanisms of CSFB to CDMA2000 1xRTT, which include (1) 1xCSFB, (2) enhanced 1xCSFB (e1xCSFB), (3) 1xCSFB with dual receiver, and (4) e1xCSFB with dual transceiver, depending on inter-RAT mobility type and UE capability. For CSFB to CDMA 1xRTT, the eNB provides the 1xRTT-related parameters in SIB8 of the system information.

1. **1xCSFB:** 1xCSFB is the default mechanism defined in Release 8 for CSFB to 1xRTT. Redirection is used for inter-RAT mobility triggered by 1xCSFB; that is, the eNB sends an *RRCConnectionRelease* message with redirection information to the UE for inter-RAT mobility from E-UTRAN to 1xRTT. Upon receiving the mobility command for 1xCSFB, the UE releases the RRC connection, leaves E-UTRAN, and attempts to access the 1xRTT. If 1xCSFB is performed, IP services over EPS bearers are suspended or deactivated.

   The network support for 1xCSFB is indicated by providing 1xRTT registration parameters in *SystemInformationBlock8* (SIB8) of an E-UTRAN cell. Support for 1xCSFB is mandatory for a UE supporting CSFB to 1xRTT. A 1xCSFB-capable UE registers with 1xRTT via the LTE network by using the 1xRTT registration parameters provided in SIB8.
2. **Enhanced 1xCSFB:** Enhanced 1xCSFB (e1xCSFB) is the enhanced version of 1xCSFB defined in Release 9 to enable the network to provide the UE with radio resources that are allocated for the UE by the target 1xRTT cell.

The UE indicates to the network support for e1xCSFB by *e-CSFB-1XRTT* in the UE capability. Based on this capability information, the eNB may decide to perform e1xCSFB for the UE when a voice service is to be initiated. A UE supporting e1xCSFB registers with 1xRTT according to the 1xRTT registration parameters in SIB8.

e1xCSFB allows the option of performing concurrent PS handover to High Rate Packet Data (HRPD), to provide IP service continuity after e1xCSFB. For enhanced 1xCSFB with concurrent PS handover to HRPD, the UE indicates that it supports concurrent 1xRTT and HRPD to the network in the UE capability.

For e1xCSFB, the eNB sends to the UE a *MobilityFromEUTRACommand* message that includes 1xRTT configuration; that is, a 1xRTT channel assignment message. If concurrent PS handover to HRPD needs to be performed, the *MobilityFromEUTRACommand* also includes HRPD redirection information. Upon receiving the mobility command, the UE attempts to access the 1xRTT cell by using the 1xRTT channel assignment information.

3. **1xCSFB with Dual Receiver:** The dual receiver 1xCSFB defined in Release 9 is the approach that makes use of dual radio capabilities of the UE, if supported. The dual receiver capability makes it possible for the UE to receive over E-UTRAN and 1xRTT simultaneously. In this case, the UE normally transmits and receives on E-UTRAN while the UE is also camping on 1xRTT performing 1xRTT dormant mode operations by using 1xRTT receiving capability.

   Note that registration to 1xRTT is done directly over 1xRTT without signaling through the LTE network. This means that simpler network implementation is possible in 1xCSFB with dual receiver because coordination between the 1xRTT and the LTE network for 1xRTT registration is not necessary. When the UE needs to leave E-UTRAN for registration to 1xRTT, the UE sends an extended service request to the MME, similar to the normal CSFB MT/MO call procedures, which results in the release of the RRC connection to the eNB and an S1-U bearer to the S-GW.

   Mobility to 1xRTT is performed by the RRC connection release procedure. The *RRCConnectionRelease* message does not include redirection information because the UE has already been camping on 1xRTT after registration to 1xRTT by using the CDMA2000 radio receiver. It should be noted that with this dual receiver 1xCSFB, UE battery consumption will be higher because both E-UTRAN and 1xRTT radios should be kept running.
4. **e1xCSFB with Dual Transceiver:** If the UE is capable of receiving/transmitting over both LTE and 1xRTT radios simultaneously, the simplest way of providing CS voice services would be to make both radios operate independently. The benefit of independent operations is that the IP service over LTE is not impacted by the status of 1xRTT radio operations. However, such approach would considerably shorten UE battery life as two radios need to be running all the time.

   In Release 10, a dual transceiver e1xCSFB mechanism is introduced to utilize the dual transceiver capability whilst avoiding higher battery consumption. The idea behind the dual transceiver e1xCSFB mechanism is to turn on the 1xRTT radio only when a CS voice call or SMS are to be served, and to turn it off otherwise. While the 1xRTT radio is switched on, the 1xRTT signaling for, for example, registration to 1xRTT is tunneled between the UE and 1xRTT via the LTE network. Note that even while the 1xRTT radio

Table 8.2 Possible mobility options for 1xCSFB. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

No	Mobility Option (Related RRC Messages)	Tx/Rx	Release	UE Capability
1	Redirection (*RRCConnectionRelease* with redirection information)	Single/Single	Release 8	Mandatory for UEs supporting CSFB to 1xRTT
2	Enhanced 1xCSFB (*HandoverFromEUTRA PreparationRequest, ULHandoverPreparationTransfer, MobilityFromEUTRACommand*)	Single/Single	Release 9	Indication of the *e-CSFB-1XRTT* in the UE capability
3	Enhanced 1xCSFB with concurrent HRPD handover (same as enhanced 1xCSFB)	Single/Single	Release 9	Indication of the *e-CSFB-ConcPS-Mob1XRTT* in the UE capability
4	1xCSFB with dual receiver (*RRCConnectionRelease* without redirection information)	Single/Dual	Release 9	Indication of the *rx-Config1XRTT* set to *dual* in the UE capability
5	Enhanced 1xCSFB with dual transceiver (*HandoverFromEUTRA PreparationRequest, ULHandoverPreparationTransfer, DLInformationTransfer*)	Dual/Dual	Release 10	Indication of the *e-CSFB-dual-1XRTT* in the UE capability

is running, the UE can keep connected with E-UTRAN to continue IP services with dual transceiver capability.

The eNB advertises support for e1xCSFB with dual transceiver by broadcasting *csfb-DualRxTxSupport* in SIB8. The UE indicates support for e1xCSFB with dual transceiver by *e-CSFB-dual-1XRTT* in the UE capability. For e1xCSFB with dual transceiver, the eNB may send to the UE a *DLInformationTransfer* message including the 1xRTT configuration. Upon receiving the *DLInformationTransfer* message, the UE with dual transceiver turns on the 1xRTT radio while the UE maintains the operations of the LTE radio.

Table 8.2 summarizes the available options for 1xCSFB.

8.3.2.2 Reference Architecture

Figure 8.12 shows the network architecture for CSFB to 1xRTT. The MME is connected to the 1xCS Interworking System (IWS) via the S102 interface, over which tunneled 1xRTT messages are communicated between the 1xRTT part of the UE and the 1xRTT network. 1xRTT CS paging is also transferred to the LTE network via S102 and is then delivered to the UE.

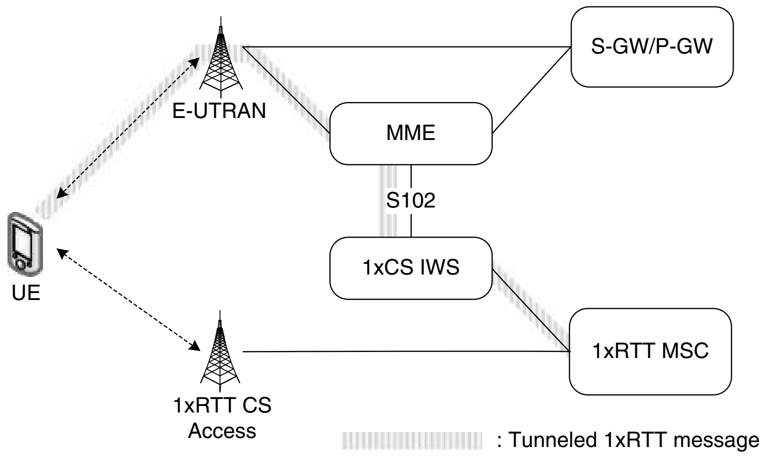

Figure 8.12 Network architecture for CSFB to 1xRTT CS. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

8.3.2.3 Pre-Registration to 1xRTT over the LTE Network

A UE needs to be registered on a 1xRTT network for CSFB to 1xRTT. In Release 8, a mechanism was introduced to allow a UE to register with 1xRTT via the LTE network. This 1xRTT registration procedure over the LTE network is often called "pre-registration" and is performed by exchanging 1xCS messages between the UE and the 1xRTT MSC. Pre-registration can be applied for 1xRTT CS and/or 1xRTT HRPD.

Figure 8.13 illustrates the pre-registration procedure. The E-UTRAN can advertise the pre-registration parameters in SIB8. Network support for pre-registration to 1xRTT and HRPD is indicated by *csfb-RegistrationParam1xRTT* and *preRegistrationInfoHRPD*, respectively, in SIB8.

The RRC layer of the UE, upon receiving a SIB8 indicating support for pre-registration, forwards the relevant parameters included in SIB8 to the CDMA2000 upper layer to determine whether (re-)registration is needed. If the CDMA2000 upper layer of the UE requests registration to 1xRTT CS, the UE performs the registration procedure by sending a 1xRTT CS registration request message to the eNB. The 1xRTT CS registration request message is forwarded to the MME and then tunneled via the S102 interface to the 1xCS IWS responsible for interworking between the LTE network and the 1xRTT CS domain.

Upon receiving the 1xRTT CS registration request message, the 1xCS IWS performs registration of the UE with the 1xRTT MSC. The result of 1xRTT CS registration is sent back to the UE through a 1xRTT CS registration response message.

8.3.2.4 1xRTT CSFB for a Mobile Originating Call

Figure 8.14 shows the simplified procedures of a mobile originating call with CSFB to 1xRTT using the *RRCConnectionRelease* message with redirection information.

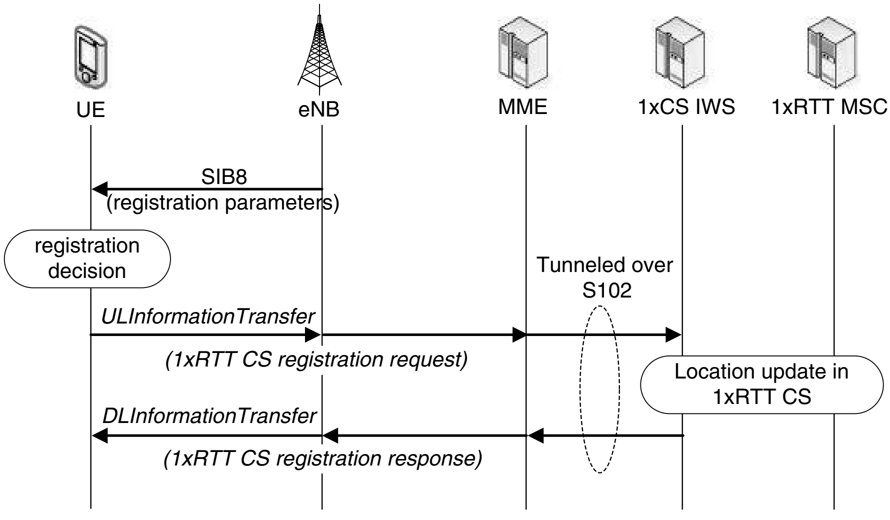

Figure 8.13 Pre-registration to 1xRTT. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

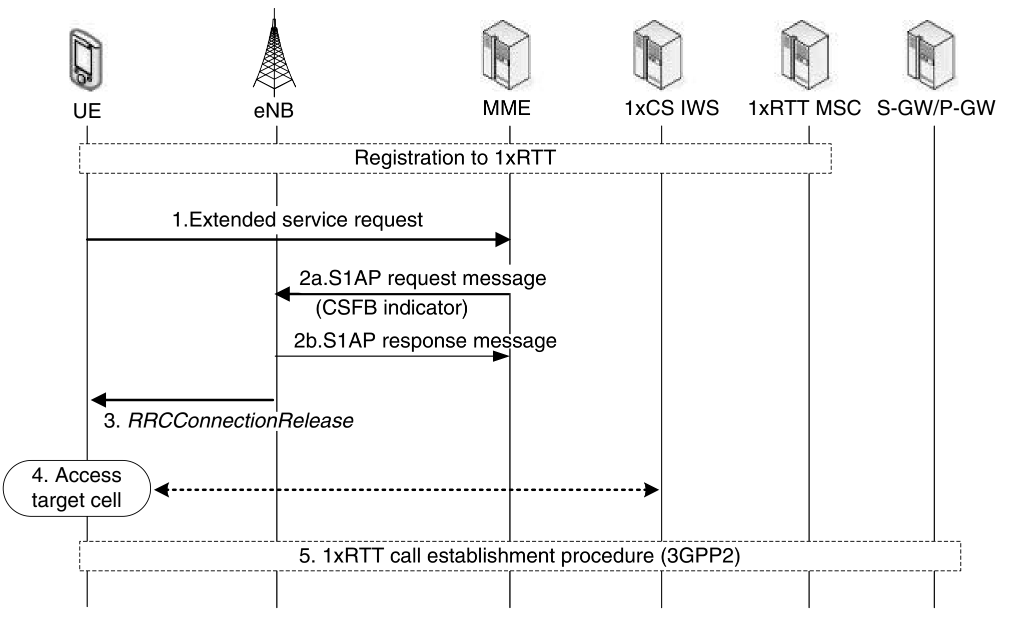

Figure 8.14 Mobile originating call with CSFB to 1xRTT (1xCSFB). Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

- **Step 1:** If the UE cannot initiate IMS VoIP and the UE is attached to both LTE and 1xRTT, the UE decides to invoke 1xRTT CSFB for a mobile originating call by sending an extended service request to the MME.
- **Step 2:** The MME, receiving the extended service request, triggers the eNB to perform the CSFB-related inter-RAT mobility procedure from E-UTRAN to 1xRTT by using an S1AP message (UE context modification request). The S1AP message includes a CSFB indicator.
- **Step 3:** The eNB, upon receiving the S1AP message including the CSFB indicator, triggers inter-RAT mobility to 1xRTT by sending an *RRCConnectionRelease* message to the UE including redirection information towards 1xRTT. Then, the eNB requests that the MME release the S1 UE context stored in the MME.
- **Step 4:** The UE, upon receiving the *RRCConnectionRelease* message from the eNB, releases the RRC connection and accesses the 1xRTT network according to the redirection information included in the *RRCConnectionRelease*.
- **Step 5:** For 1xRTT call establishment, the UE follows 3GPP2 specifications.

8.3.2.5 Enhanced 1xRTT CSFB (e1xCSFB) for a Mobile Originating Call

Figure 8.15 shows the simplified procedures for e1xCSFB for a mobile originating call. Note that in step 7, a *MobilityFromEUTRACommand* is sent to a UE supporting a single transceiver while a *DLInformationTransfer* is sent to a UE supporting a dual transceiver.

8.3.2.6 1xRTT CSFB for a Mobile Terminating Call

For a mobile terminating call with CSFB to 1xRTT, paging reception/response signaling is added to the procedure for a mobile terminating call. Figure 8.16 shows the 1xRTT CSFB procedures for a mobile terminating call.

- **Step 1:** The 1xRTT MSC sends a 1xRTT CS paging request to the 1xCS IWS. The paging request is then tunneled to the MME via the S102 interface. The 1xRTT CS paging request is forwarded to the eNB and then forwarded to the UE via a *DLInformationTransfer*.
- **Step 2:** The UE responds to the received 1xRTT CS paging request by sending an extended service request to the MME.
- **Step 3:** The MME, receiving the extended service request, invokes the eNB to perform CSFB-related inter-RAT mobility by using an S1AP message including a CSFB indicator.
- **Steps 4–7:** The same steps 3 to 5 in Figure 8.14 are applied. In addition, the UE sends a 1xRTT CS paging response to 1xRTT. Upon receiving the paging response, the 1xRTT MSC stops retransmitting the 1xRTT CS paging request.

8.4 Service Domain Selection

For a mobile originating call, if the UE supports both CSFB and IMS VoIP, it has to decide which voice solution should be used out of preference. This decision is equivalent to selecting either the CS or PS domain for voice services. The UE desirably considers (1) the voice domain preference provisioned to the UE by the HPLMN operator's preference, and (2) the network capabilities indicated by the network [5].

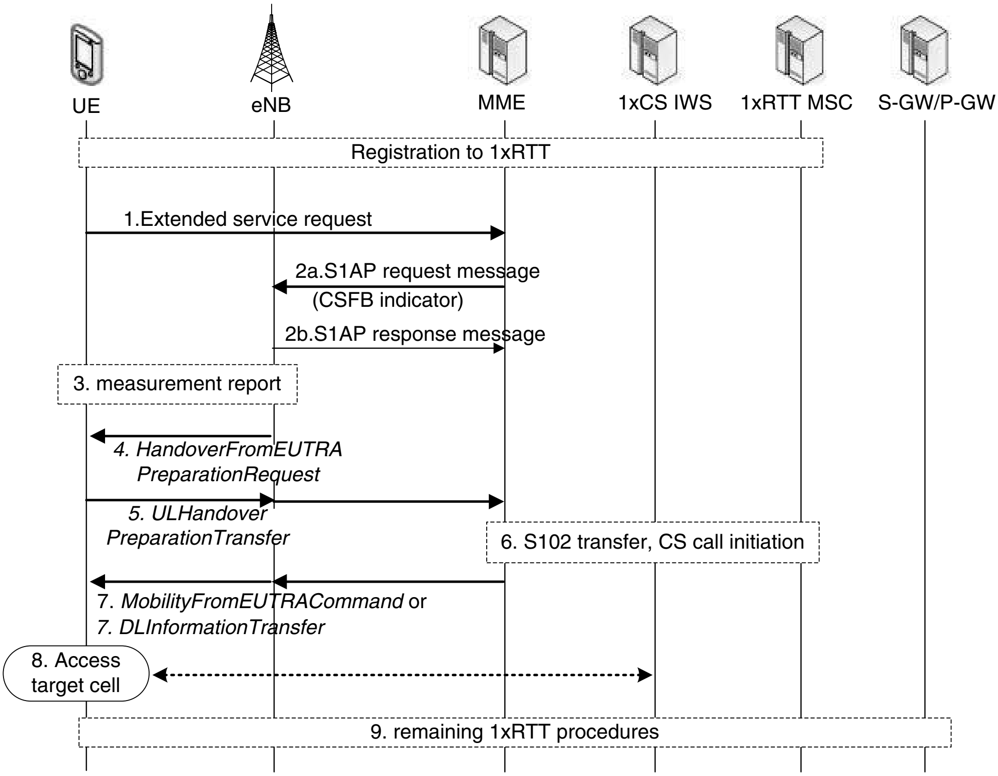

Figure 8.15 Mobile originating call with enhanced CSFB to 1xRTT. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

- **Voice Domain Preference:** The voice domain preference indicates the domain (PS domain or CS domain) preferred by the UE for a voice service.
  - **CS Voice only:** The UE configured with this value does not attempt to use IMS voice with an EPS bearer.
  - **CS Voice preferred, IMS PS Voice as secondary:** The UE configured with this value "preferably" attempts to use the CS domain for a voice service.
  - **IMS PS Voice preferred, CS Voice as secondary:** The UE configured with this value "preferably" attempts to use IMS voice with an EPS bearer.
  - **IMS PS Voice only:** The UE configured with this value attempts to use IMS voice with an EPS bearer.

  The UE provides this information to the network, and the network can use the information to select a RAT/Frequency Selection Priority (RFSP) for the UE. The selected RFSP may be configured to the UE in order to steer the UE towards the RAT of the preferred domain.

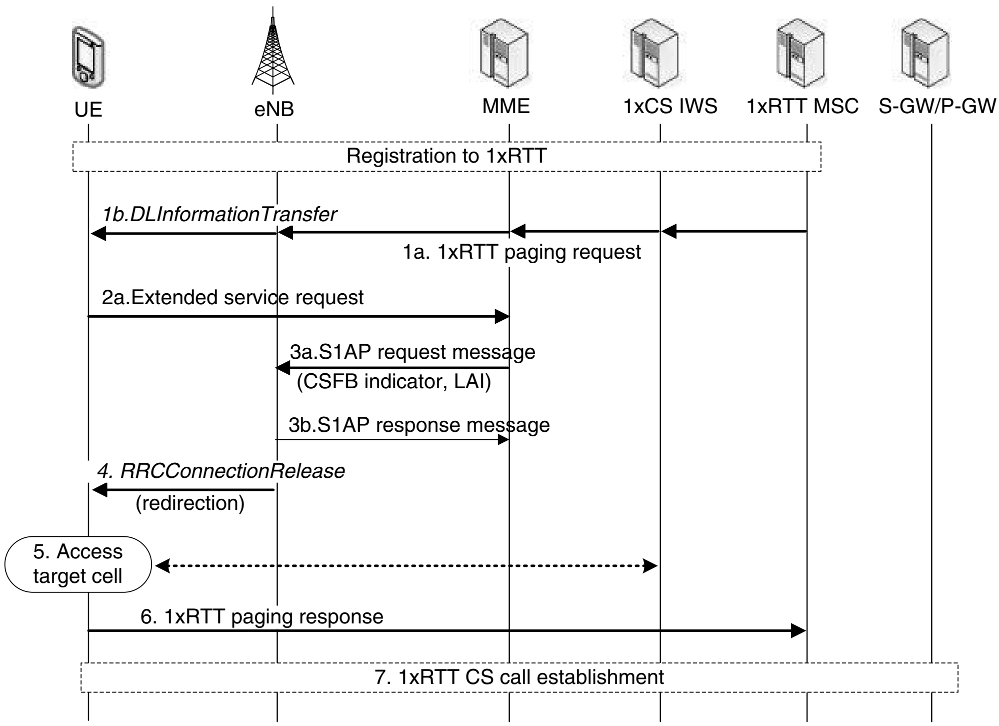

Figure 8.16 Mobile terminating call with CSFB to 1xRTT (1xCSFB). Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

- **Network Capabilities:** During the Attach or Tracking Area Update procedures, the network provides the UE with the following information on its supported features. This information may affect the selection of voice solution by the UE:
  - IMS voice support;
  - SMS only;
  - CSFB not preferred.

8.4.1 UE Decision between IMS VoIP and CSFB

In principle, the choice between IMS voice and CSFB is left to UE implementations, but the implementation guidelines for UEs supporting both CSFB and IMS are presented in [6]. In general, the UE knows, during the Attach procedure(s), whether its preferred domain for a voice service can actually provide a voice service. If a voice service is possible on the preferred domain, the UE will use the preferred domain when a voice call is to be serviced, otherwise it will use another domain. For example, let us assume that one UE is set to "CS Voice preferred, IMS PS Voice as secondary". According to the UE's voice domain preference, the UE first attempts to initiate combined registration to be ready for CSFB, as described in Section 8.3. If the combined attach procedure is successful, and if the network does not indicate “CSFB not preferred” or “SMS only” in the attach accept message, the UE will desirably select CSFB for a voice service. Otherwise, the UE will use IMS voice. Other examples can be found in [6].

8.5 Comparison between IMS VoIP and CSFB

IMS VoIP can fully employ the benefits of a flat IP architecture, for example, maintaining the LTE data rate during a voice call and accommodating fruitful multimedia attributes that can be blended with a voice service or other services. Investment on the network is focused largely on deploying the IMS core rather than upgrading a legacy network. Since the voice service is integrated with the generic IMS framework within all IP structures, operational benefit from simplified service provision is maximized. IMS VoIP is fully 3GPP-compliant, and thus fewer interoperability issues are expected.

CSFB induces additional call setup delay in comparison to a generic voice call that originates and terminates in a legacy network. The contributors to the additional delay include the RRC Connection Establishment procedure on LTE and the subsequent mobility procedure from LTE to 2G/3G, which are not required in a legacy voice call setup procedure. If the UE needs to perform a location area update after moving to legacy RAT during CSFB, further delay is added. During a voice call enabled by CSFB, it is unavoidable that the data service will be degraded or even suspended as the UE is served by legacy RAT. CSFB also requires an upgrade of a legacy network to support, for example, CS paging over E-UTRAN and mobility from LTE to 2G/3G for CSFB.

Even with its shortcomings, CSFB is a viable and appealing solution because it can readily provide voice services by reusing the existing voice solution over the legacy network. The changes required to the network are not significant, so operators are willing to promote the use of CSFB as an interim voice solution. Since CSFB is a 3GPP standard, interoperability issues can be mostly avoided.

In Table 8.3, a summary of comparison results between IMS VoIP and CSFB is provided.

8.6 RAN Optimization for VoIP

From an access stratum perspective, the bearer carrying voice packets has the following characteristics:

- the payload size of voice packets is small;
- the arrival of IP packets is regular, for example, with 20 ms periodicity.

From these characteristics, one would expect the signaling overhead due to a relatively larger packet header and frequent scheduling to result in a significant increase of signaling overhead in the radio interface when VoIP is used extensively. In some situations, VoIP coverage may be limited by a power-limited uplink, which cannot be resolved by conventional retransmission mechanisms due to the stringent delay requirement for voice. To resolve these issues, Release 8 defined some optimizations including Robust Header Compression, TTI bundling, and Semi-Persistent Scheduling.

Table 8.3 Comparison between IMS VoIP and CSFB

Comparison Class	Comparison Factor	IMS VoIP	CSFB
Performance	Call setup delay	Comparable or shorter than native call setup delay	Longer than native call setup delay
	Concurrent voice and data during voice call	Yes	Yes/No (depending on target)
	Maintaining LTE data rate during voice service	Yes	No
Service	Support of advanced services	Yes	No
UE complexity	Features to be supported	Medium complexity (IMS client, SRVCC)	Low complexity (CSFB-related enhancement)
NW operational benefit/cost	New network equipment installation	Yes (IMS core)	None, but small upgrades to MME
	Legacy network equipment upgrade	Limited	Yes (all MSCs overlaid on LTE networks)
	Coverage engineering between LTE and legacy coverage	Not required	Required
	All IP structure	Yes	No
Interoperability issue	Standard compliance	3GPP compliant	3GPP compliant

8.6.1 Robust Header Compression (ROHC)

ROHC is a method to compress IP/UDP/RTP or IP/TCP headers. The IMS Profile for Voice and SMS mandates that the UE and the network should support Robust Header Compression (ROHC) to reduce overhead. ROHC is performed in the PDCP layer. The details of ROHC are described in Section 4.2.

8.6.2 TTI Bundling

In cases where the coverage of VoIP is limited by a power-limited uplink at, for example, a cell edge, conventional methods such as HARQ retransmission in the MAC layer or segmentation in the RLC layer could be used. Such methods cannot avoid introducing additional delay and signaling overhead, which definitely impacts negatively upon the quality of real-time service and efficient usage of scarce radio resources.

As an alternative solution, the concept of TTI bundling has been introduced. If TTI bundling is configured, one transport block containing a VoIP packet is transmitted repeatedly over a bundle of four consecutive TTIs. In each TTI, a different redundancy version is applied to the transport block. Downlink HARQ feedback for uplink transmission with TTI bundling is sent by the eNB once per bundle. For more information, see Section 6.8.

8.6.3 Semi-Persistent Scheduling for HARQ

Since a speech codec typically generates a voice packet every 20 ms, a scheduling grant needs to be provided to the UE in a periodic manner during a voice service. Considering that the size of a VoIP packet is relatively small, dynamic scheduling of each voice packet would result in significant overhead in terms of control signaling and scheduling.

To reduce such overhead, Semi-Persistent Scheduling was introduced in Release 8. With Semi-Persistent Scheduling, predefined scheduling grants for reception and/or transmission are configured, and the UE can receive downlink traffic or transmit uplink traffic at predefined occasions without explicit signaling. For more information, see Section 6.4.2.

References

1. 3GPP Technical Specification 23.237, "IP Multimedia Subsystem (IMS) Service Continuity; Stage 2 (Release 10)", www.3gpp.org, accessed on 14 March 2012.
2. 3GPP Technical Specification 36.331, "Radio Resource Control (RRC); Protocol Specification (Release 10)", www.3gpp.org, accessed on 14 March 2012.
3. 3GPP Technical Specification 23.216, "Single Radio Voice Call Continuity (SRVCC); Stage 2 (Release 10)", www.3gpp.org, accessed on 14 March 2012.
4. 3GPP Technical Specification 23.272, "Circuit Switched (CS) fallback in Evolved Packet System (EPS); Stage 2 (Release 10)", www.3gpp.org, accessed on 14 March 2012.
5. 3GPP Technical Specification 24.301, "Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS); Stage 3 (Release 10)", www.3gpp.org, accessed on 14 March 2012.
6. 3GPP Technical Specification 23.221, "Architectural Requirements (Release 10)", www.3gpp.org, accessed on 14 March 2012.

9

Home eNB (HeNB)

The HeNB is a low-power base station to provide mobile devices with connections to a mobile operator's network via, for example, residential broadband IP connections. Due to the low-power transmission capability, the coverage of HeNB cells is typically quite small in comparison to that of macro cells. For this reason, it is often said that HeNBs are base stations for femto cell services. Figure 9.1 illustrates a conceptual view of HeNB deployment.

In 3GPP specifications, low-power base stations serving femto cells are called the Home Node B (HNB) for UTRAN and the Home eNode B (HeNB) for E-UTRAN, respectively.

"Low power" can also be interpreted as "smaller and low-priced equipment", which is favorable to scalability and/or ubiquitous deployment of femto cells. From many operators' experience, femto cells have been proven to be a cost-efficient means of capacity enhancement and coverage extension. Deep penetration of broadband IP connections also encourages wider femto cell deployment.

While demand for femto cell deployment has been growing strongly, a standard solution involving femto cells that is favorable to a multi-vendor environment does not exist. Different vendors with different implementations formed fragmented markets, and the lack of interoperability between different vendors prohibited the economies of scale of femto cell markets.

To enable high-volume production of femto cells with full interoperability between different vendors, 3GPP has made strenuous efforts in cooperation with the Femto Forum and the Broadband Forum. As a result, support for the femto cell was enabled from Release 8 throughout a number of specifications to ensure standard-based femto cell production and deployment.

It is interesting that, in contrast to what the name "Home"(e)NB may suggest, the H(e)NB deployment environment is not limited to the residential home but also includes enterprises, urban hotspots, and rural areas. In the initial phase of studying femto cell support in 3GPP, femto cell deployment was considered mainly for the home environment [1]. Even though 3GPP has identified many other scenarios in which HeNBs are employed outside the home, 3GPP specifications still stick to the original terminology and therefore the term "H(e)NB" has been used officially throughout the 3GPP specifications.

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

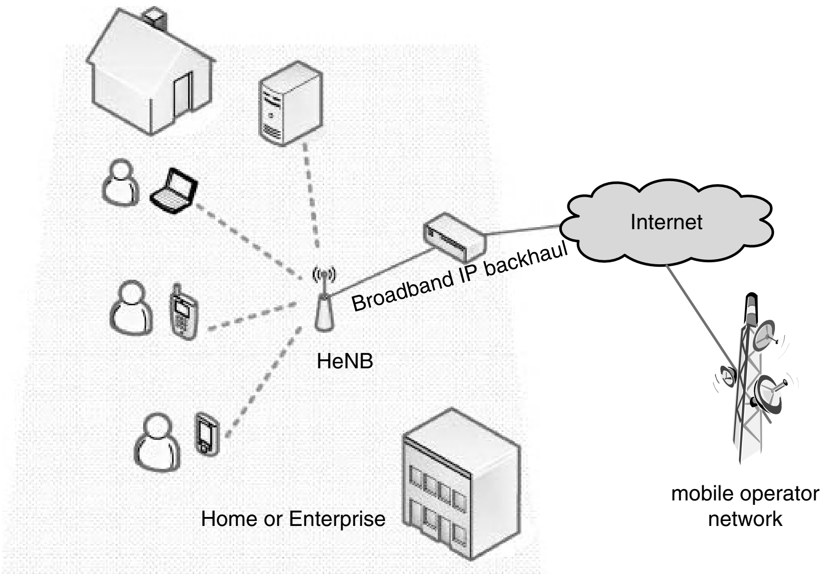

Figure 9.1 Conceptual view of HeNB

There are many terms defined in [2,3] which are relevant to HeNBs. Some key terms and definitions are given below:

- **Home Node B (HNB):** An HNB is a piece of equipment on the customer premises that connects a UE over UTRAN to a mobile operator's network using a broadband IP backhaul.
- **Home eNode B (HeNB):** An HeNB is a piece of equipment on the customer premises that connects a UE over E-UTRAN to a mobile operator's network using a broadband IP backhaul.
- **H(e)NB:** HNB and HeNB.
- **Closed Subscriber Group (CSG):** A Closed Subscriber Group identifies subscribers of an operator who are permitted to access one or more cells for which access is restricted (CSG cells). All H(e)NBs serving the same CSG share the same identity, called the CSG Identity.
- **CSG cell:** A CSG cell is a cell (part of the PLMN) broadcasting a specific CSG Identity. A CSG cell is accessible by the members of the closed subscriber group for that CSG Identity. All the CSG cells sharing the same identity are identifiable as a single group.

9.1 Architectural Framework

9.1.1 Access Mode

In Release 8, the HeNB operates in either open access mode or closed access mode. If the HeNB operates in open access mode, it is not associated with any CSG and thus operates as a normal eNB.

Table 9.1 HeNB access modes and UE access permissions (Release 8)

CSG membership	HeNB Access Mode	
	Open access mode	Closed access mode
CSG member	UE Access is allowed	Access is allowed
CSG non-member	UE Access is allowed	Access is not allowed

If the HeNB operates in closed access mode, it is associated with a CSG broadcast by the HeNB, and access to the HeNB is only allowed for UEs qualified as CSG members. Table 9.1 shows the two HeNB access modes and the corresponding UE access permissions.

9.1.2 Use Cases

There are many use cases where HeNBs may provide distinctive benefits, as shown in Figure 9.2:

- **Closed access to home members:** An HeNB is installed at home and operates in closed access mode. CSG membership is shared only by family members. Family members are permitted to access the HeNB while others are not. As all scheduling opportunities

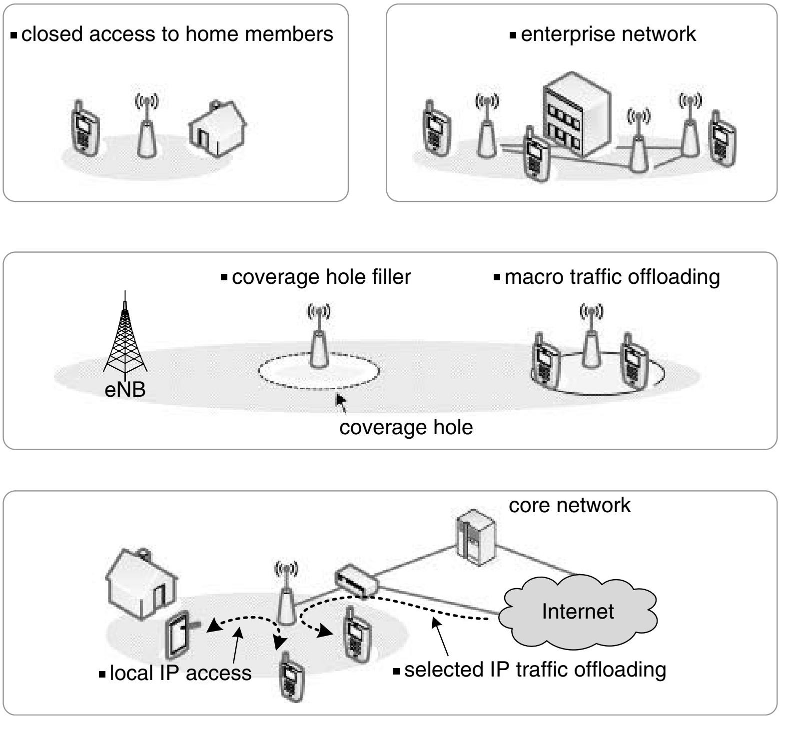

Figure 9.2 HeNB use cases

provided by the HeNB are shared only among family members, a better QoS is experienced; for example, a higher data rate and shorter delays.
- **Enterprise network:** A cluster of HeNBs is installed in an enterprise and operates in closed access mode. An employee can access groupware with his own mobile device. The HeNB may be interconnected with an enterprise telephony system. Employees can make/receive a call to/from a fixed line via mobile devices connected to the HeNB.
- **Macro cell traffic offloading:** An HeNB is installed in a metropolitan hotspot and operates in open access mode. As the UE traffic under the HeNB service area is traversed via the HeNB rather than a macro cell, macro cell traffic is offloaded to the HeNB and the macro cell is less congested.
- **Coverage hole filler:** An HeNB is installed by a home owner in his residential area where an mobile operator's public network coverage is not available. The user can then be provided with rich multimedia services over radio interface.
- **Home-GW:** An HeNB is installed at a residence and operates in closed access mode. Home-based services, for example, a home media server, a home multimedia player, a wireless printer, and IP TV, are concentrated at the HeNB. A user can enjoy home-based services via a mobile device within the HeNB coverage and/or in the mobile operator's coverage beyond the HeNB coverage. A user can transfer a multimedia session running on a mobile device to another device, for example, a TV, via the H(e)NB.
- **Local IP Access (LIPA):** An HeNB is installed in a residence/enterprise. IP-enabled UE is capable of being connected to other IP-enabled devices via the HeNB without traversing the mobile operator's network except the operator's network component, for example, a Local-GW collocated in the HeNB. IP-enabled UE connected to the HeNB is still reachable by the operator in terms of control signaling.
- **Selected IP Traffic Offloading (SIPTO):** An HeNB is installed in a residence/enterprise. UE is provided with an IP service (e.g., an Internet service) without traversing the mobile operator's network.

9.1.3 High-level Requirements

The following are high-level requirements that should be supported in the HeNB and services using the HeNB [4]:

- **Access Control:** If allowed by the agreement between the operator and the HeNB Hosting Party, the access mode of the HeNB can be configured by the operator.
- **CSG Provisioning:** CSG membership should be able to be added or removed, under the supervision of the operator. The UE and the network should maintain a list of CSGs for which the UE is allowed access.
- **Display:** The CSG type of a cell – that is, open or closed (or hybrid in Release 9) – should be shown/indicated to a user when the UE is camping on a CSG cell such that a user can distinguish the cell from a normal cell. If the CSG type associated with the CSG has not been configured for the UE, the “HNB name” broadcast by the HeNB is indicated instead.
- **Operations, Administration, and Maintenance (OAM):** OAM procedures should be supported by the HeNB such that the operator can configure the parameters of HeNBs.
- **Mobility Management:** Normal PLMN selection procedures can be used by the UE to register with a PLMN via the HeNB. In mobility in RRC_IDLE, a CSG cell should be preferred if the relevant criteria are fulfilled. Mobility in RRC_CONNECTED – that is,

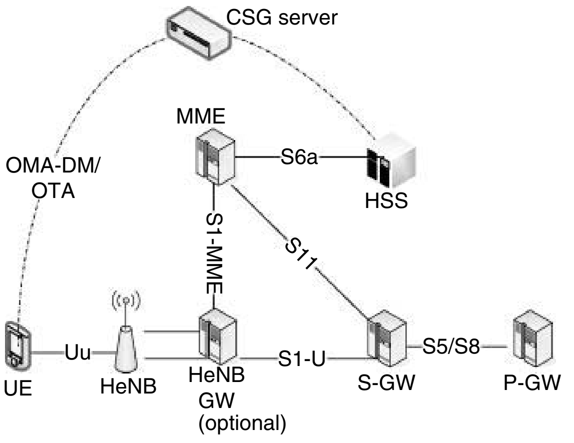

Figure 9.3 Network architecture supporting HeNB. Reproduced by permission of 3GPP, © 2010. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

handover – is supported in the case where the source and/or target cell is an HeNB cell. A scan for available CSGs is required to support mobility.

- **Security:** The HeNB should provide a security level that is at least equivalent to the Release 8 3GPP system.

9.1.4 Network Architecture

In Figure 9.3, the network architecture supporting HeNBs is illustrated [5]. An HeNB can be connected directly to the MME/S-GW. A Home eNB Gateway (HeNB GW) may be used to establish the S1 interface between the HeNB and the EPC. The HeNB GW facilitates the deployment of a large number of HeNBs by reducing signaling burden that could otherwise be propagated to the MME.

For the control plane, the S1-MME interface from the HeNB may be terminated at the HeNB-GW, or a direct connection between the HeNB and the MME may be established. For the user plane, the S1-U interface from the HeNB may be terminated at the HeNB GW, or a direct connection between the HeNB and the S-GW may be established.

The CSG server is connected to the UE and the Home Subscriber Server (HSS) for CSG provisioning purposes. Note that Release 8/9 HeNBs do not support the X2 interface. In Release 10, an HeNB can support the X2 interface with neighboring HeNBs with some restriction as described in Section 9.5.2.

9.2 CSG Provisioning

It should be possible for a network to manage CSG membership, for example, by adding or removing a particular subscriber UE from a CSG list that is under the network's control. A UE should be able to provision the CSG list that is under the UE's control. The management of CSG membership for the UE and the network is called *CSG provisioning*.

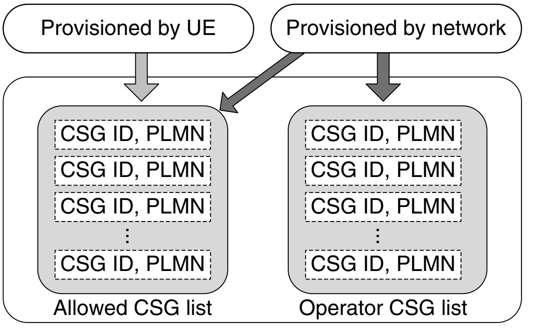

Figure 9.4 UE CSG white-list

From the UE perspective, the list of CSGs of which the UE is considered a member should be managed. The list of CSGs is referred to as the CSG white-list for the UE. The operator should also manage the CSG subscription data of its subscribers.

9.2.1 CSG Subscription Data

The CSG subscription data of UEs are stored in the HSS. The CSG subscription data are transferred to the MME when a UE registers with a network. For a UE, the CSG subscription data are stored in the USIM of the UE. Figure 9.4 illustrates an example of a CSG white-list structure. As shown in this figure, a CSG white-list consists of an "Allowed CSG list" and an "Operator CSG list". The Allowed CSG list can be provisioned by both the UE and the network, while the Operator CSG list is only provisioned by the network. CSG provisioning can be carried out by Open Mobile Alliance Device Management (OMA DM) procedures or by Over-The-Air (OTA) technologies. NAS procedures are also used for CSG provisioning in the case of manual CSG selection, where the CSG white-list can be updated during, for example, Attach or Tracking Area Update procedures.

Both the "Allowed CSG list" and the "Operator CSG list" consist of a list of entries each comprising a CSG identity and a PLMN identity associated with the CSG identity in the same entry. The UE considers that the CSG identity stored in the CSG white-list is valid only within the scope of the associated PLMN.

9.2.2 CSG Member Status

An E-UTRAN cell broadcasts information relevant to accessibility of the cell in SIB1. The PLMN identity of the cell is one of the pieces of information broadcast in SIB1. A CSG cell also broadcasts a CSG identity identifying the CSG to which the cell belongs.

A UE is said to be a member of a CSG cell if the paired CSG identity and PLMN identity broadcast by the CSG cell match at least one entry in the CSG white-list. Otherwise, the UE is a non-member of the CSG cell.

9.3 System Information Related to CSG

When a UE detects a cell, it should be able to identify the characteristics of the cell, for example, type or accessibility. If the cell is identified as a CSG cell, the UE should be able to

Table 9.2 HeNB-related broadcast information

HeNB Information	Description	Format	SIB
CSG Identity (CSG ID)	Identity of CSG associated with the HeNB Mandatory broadcast by only CSG cells	27 bitstring	SIB1
CSG Indication	Indicates whether the cell is a CSG cell or not Mandatory broadcast by both CSG cells and normal cells	Boolean	SIB1
HNB name	Human-readable name of the HeNB in free text format Mandatory broadcast by only CSG cells	Coded in UTF-8 with a variable number of bytes per character	SIB9
PCI split info for CSG cells	Reserved PCI range for CSG cells Mandatory broadcast by CSG cells Optional broadcast by non-CSG cells	PhysCellIdRange	SIB4

perform evaluation with regard to accessibility of the cell. To this purpose, HeNB-related information is broadcast by cells. The HeNB-related information includes CSG Identity, CSG Indication, HNB name, and PCI split information for CSG cells. Such information can be categorized into two types: CSG identification information and CSG cell deployment information. Table 9.2 lists the HeNB-related broadcast information.

9.3.1 CSG Identification Information

A CSG cell broadcasts CSG identification information so that a UE can identify whether the cell is a CSG cell. The CSG identification information is also used by the UE to verify whether the CSG cell is accessible. In the following, the parameters for CSG identification and the corresponding attributes of the parameters are described.

- **CSG Identity (CSG ID):**
  - The CSG ID identifies the closed subscriber group to which the cell provides services.
  - The CSG cell broadcasts the CSG ID.
  - The UE uses the CSG ID to verify authorized access to the CSG.
  - SIB1 carries the CSG ID.
  - The format of the CSG ID is a 27 bit-long bitstring.
- **CSG Indication:**
  - A CSG Indication indicates whether the cell broadcasting the CSG Indication is a CSG cell or not.
  - All E-UTRAN cells broadcast a CSG Indication.
  - The UE uses the CSG Indication to identify whether the cell is a CSG cell or not.
  - SIB1 carries the CSG Indication.
  - The CSG Indication is Boolean in nature. Only a CSG cell can set the CSG Indication to TRUE.
- **HNB Name:**
  - The HNB Name is a broadcast string in free text format that provides a human-readable name for the HeNB.

Table 9.3 HeNB cell type/access mode identification by UE (Release 8)

Cell Type	HeNB Access Mode	CSG Indication	CSG Identity
Normal cell	Open access mode	FALSE	Absent
CSG cell	Closed access mode	TRUE	Present

Different types of cells broadcast different sets of CSG identification information. A CSG cell should broadcast a CSG Indication that is set to TRUE and a CSG Identity to which it provides service. A normal cell only broadcasts a CSG Indication that is set to FALSE. Note that if the HeNB operates in open access mode, a cell of the HeNB is a normal cell. Table 9.3 summarizes CSG identification information broadcast by the different types of cells.

9.3.2 CSG Cell Deployment Information

The cell reselection procedure requires the UE to perform several evaluations, for example, ranking of the cell reselection candidates by comparing the measured quality, and a suitability check by reading the system information of the reselected cell. For a UE that does not support CSG, it is desirable to ignore CSG cells during the cell reselection procedure to avoid unnecessary cell reselection attempts for CSG cells.

In order to enable such a UE to ignore CSG cells without reading the CSG Indication of those cells, 3GPP adopted the operational assumption that a range of Physical Cell Identities (PCIs) should be reserved for CSG cells for a frequency on which CSG cells are deployed. This means that the whole PCI range (0, . . . ,503) is divided into a CSG range and a non-CSG range, and any CSG cell on the concerned frequency should be configured to use one of the PCIs within the PCI range reserved for CSG cells.

The usefulness of this PCI split stems from the fact that the PCI of a concerned cell can be known during the cell detection phase – that is, at the acquisition of a primary sync channel and a secondary sync channel for the cell – which is significantly earlier than reading the system information of the cell. Given that PCI split information is available to a UE that has a non-existent or empty CSG white-list, the UE can ignore any cell using a PCI within the PCI range for CSG cells from among the cell reselection candidates, such that it can avoid unnecessary reselection attempts/evaluations.

The PCI split information between CSG cells and non-CSG cells is provided to the UE in SIB4, and is referred to as the *csg-PhysCellIdRange* [6]. Broadcasting PCI split information is mandatory for CSG cells but optional for non-CSG cells, for example, macro cells.

The PCI split information is associated with a frequency. This association means that if PCI split information for a frequency has been obtained once by a UE, it is applicable when the UE is camping on other cells on the concerned frequency until new PCI split information is obtained for the concerned frequency. The association also means that if the UE receives PCI split information from a cell on a certain frequency, the information applies only for that frequency. For another frequency, the corresponding PCI split received for that frequency should be applied. The PCI split information is considered valid up to a maximum of 24 hours once it has been obtained.

9.4 Identification of CSG

Mobility to a CSG cell should start with the identification of a CSG cell in the vicinity of the UE. The identification process can be driven by either of the following two means:

- an autonomous CSG search function;
- a manual CSG selection function.

An autonomous CSG search function searches for and identifies CSG cells autonomously, and a manual CSG selection function performs the task manually based on NAS-level intervention, for example, a user's selection or preference.

9.4.1 Autonomous CSG Search

An autonomous CSG search function searches for and identifies member CSG cells in the proximity of the UE without the user's intervention. This function should be applicable to member CSG cells on intra-frequency, inter-frequency, and even other RAT. The autonomous search function is applicable only when the UE's CSG white-list is non-empty. The identified member CSG cell is then considered a cell selection/reselection candidate for a UE in RRC_IDLE or a possible handover target cell for a UE in RRC_CONNECTED.

Despite the importance of the autonomous CSG search function, the details of this function are not specified in a standard; when and where the function is enabled is totally up to the UE. Note that in [7], minimum requirements regarding the performance of reselection from a non-CSG to an inter-frequency CSG cell are specified only to ensure the testability of the autonomous CSG search function.

In terms of the implementation of autonomous CSG search, a periodic search of CSG cells could be considered. For more efficient implementation, location information can be used as a trigger to activate the search function. The location information could be radio measurements of macro cells at the moment when the UE is camping on a member CSG cell, or it could be more accurate geographical information, for example, the Global Navigation Satellite System (GNSS) location of the member CSG cell. To enable such smart implementation, whenever the UE camps on a CSG cell as a member, it needs to store the current location. If radio measurements of macro cells are used as location information, the cell identity and/or the measured quality of serving/neighboring macro cells should be stored. Based on the stored location information, whenever a UE experiences a similar radio fingerprint to one it has stored previously, the UE will assume that the previously visited member CSG cell is located nearby and thus will activate a search for the CSG cell.

Note that the autonomous search function is independent of normal measurements. This implies that if the autonomous search function is in use, the UE should be able to measure and identify the presence of a CSG cell, if present, even when the UE is allowed to skip normal neighboring cell measurements due to sufficiently good quality of the serving cell.

9.4.2 Manual CSG Selection

A manual CSG selection function is a complementary function to the autonomous CSG search function. A manual CSG selection can be used when the autonomous CSG search is not sufficient, for example, when the UE wants to select a certain CSG manually.

Upon a request from the NAS, the AS should find available CSG cells by scanning all radio frequencies in the supported E-UTRA band, and report all available CSGs together with the corresponding HNB names and PLMNs to the NAS. On receipt of the reported CSGs from the AS, the NAS may select one of the reported CSGs. The AS should then attempt to camp on a suitable or acceptable CSG cell belonging to the selected CSG.

It should be noted that, as the CSGs reported by the UE AS to the NAS are not restricted to member CSGs, the NAS may select a non-member CSG. If this happens and the UE has selected a CSG cell of that CSG ID, the UE should send a request for an Attach or Tracking Area Update in order to indicate to the network that the UE has selected a CSG cell with a CSG ID that is not included in the UE's CSG white-list. The network may accept the request even if the UE is not a member of the CSG, for example, to provide a temporary access grant to the UE. If the request is accepted, the UE should add the CSG ID to the Allowed CSG list, unless already present. When accepting access for a UE that is not a member of the CSG, the network may configure the validity time for the added CSG. In this manner, a UE can gain temporary access to a non-member CSG.

9.5 Mobility with CSG Cells

9.5.1 Mobility in RRC_IDLE

For cell selection regarding a CSG cell, the normal cell selection rule applies except with an extended definition of a "suitable cell" – CSG membership status is additionally taken into account for suitability. This means that during cell selection a CSG cell is not prioritized over a normal cell.

In contrast, for cell reselection, a member CSG cell is prioritized over normal cells. Figure 9.5 shows the sequential UE behaviors for RRC_IDLE mobility to a CSG cell.

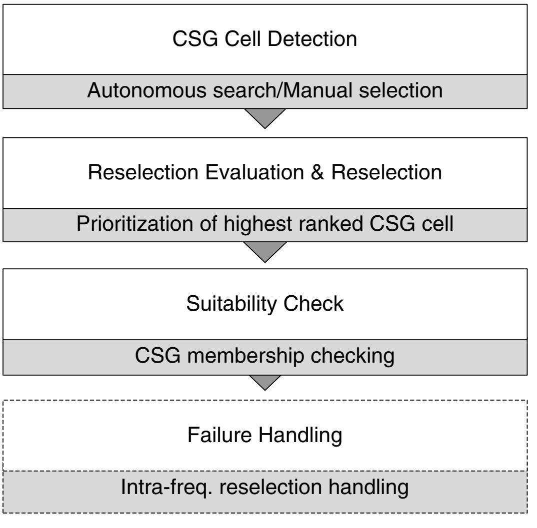

Figure 9.5 RRC_IDLE mobility to a CSG cell

As shown in the figure, prioritization of a member CSG cell introduces some changes of reselection behavior on the following aspects:

- detection of a CSG cell for reselection candidate;
- cell reselection evaluation and reselection to a CSG cell;
- suitability check on the reselected cell;
- failure handling of cell reselection to a CSG cell.

9.5.1.1 Step 1: Detection of a CSG Cell

For a normal reselection, a UE detects the neighboring cells by performing measurements according to the measurement rules described in Section 2.7. This is not always true for CSG cell detection, as detection of CSG cells is enabled by the autonomous CSG search or a manual CSG selection. Remember that autonomous search and manual selection operations are independent of normal measurement rules.

9.5.1.2 Step 2: Reselection to/from a CSG Cell

If the UE detects one or more suitable CSG cells, the UE considers those CSG cell(s) to be cell reselection candidate(s). The subsequent cell reselection behaviors are slightly different depending on the frequency of the detected CSG cell. Note that the following behaviors consistently apply irrespective of the current serving cell type, for example, a CSG cell or a normal cell.

- **Intra-frequency CSG cell reselection:** If the CSG cell is detected on the serving frequency, the UE applies the normal reselection rule; that is, the highest ranked cell among the cells on the serving frequency is selected.
- **Inter-frequency CSG cell reselection:** If the CSG cell is detected on a non-serving frequency, the UE should attempt to reselect it if the CSG cell is the highest ranked on that frequency, irrespective of the priority of that frequency. Note that this prioritization mechanism for a CSG cell is equivalent to assuming that the frequency of a suitable CSG cell has the highest priority as long as the CSG cell is the highest ranked on that frequency. If the CSG cell is not the highest ranked on that frequency, the UE does not assume it has the highest priority but a signaled priority for the frequency. This is to prevent interference that is otherwise supposed to be generated from/to the UE camping on a non-best-ranked cell on the frequency.

9.5.1.3 Step 3: Suitability Check on a Reselected CSG Cell

Once the UE has reselected a new cell as described in step 2, the UE checks whether the CSG cell is suitable. The general conditions for a cell to be deemed suitable are described in Section 2.2. For a CSG cell to be deemed suitable, one more condition is imposed: the UE should be a CSG member for the CSG cell, that is, the CSG ID and the PLMN identity advertised by the CSG cell should be present in the UE CSG white-list. If the UE is a CSG non-member for a CSG cell, it considers the cell unsuitable.

During the search for CSG cells, the UE may detect multiple CSG cells that are suitable. As all CSG IDs in the UE's CSG white-list have the same priority, as specified in [3], UE implementation should determine which CSG the UE selects.

9.5.1.4 Step 4: Failure Handling of Cell Reselection to a CSG Cell

If the reselected cell turns out to be unsuitable due to the UE's non-member status of the CSG, the UE should exclude the cell from its reselection candidates, but other cells on the same frequency should still be considered as reselection candidates. Note that this behavior is different from the handling of reselection attempts to normal cells, where, in the case of reselection failure due to the unsuitability of a normal cell, all cells on the concerned frequency are excluded from reselection candidates to prevent interference that could otherwise be generated from/to the UE camping on those cells. The difference in the CSG case is to ensure that, given that the UE may frequently encounter non-member CSG cells, any non-member CSG cell that the UE faces should not deprive the UE of reselection chances to other accessible cells on the same frequency.

If the reselected cell turns out to be "barred", similar handling to that described in Section 2.8 applies. That is, when the reselected cell is considered a barred cell, the UE still considers the other cells on the same frequency of the barred cell as cell reselection candidates. Note that the *intraFreqReselection* in the system information of the barred cell is ignored in this case. This is different from the handling of a barred cell that is not a CSG cell, where intra-frequency reselection to other cells is strictly controlled by *intraFreqReselection*.

9.5.2 Mobility in RRC_CONNECTED

It is essential to provide a user with a seamless service experience when a user moves to/from an HeNB cell with ongoing service. Connected mode mobility to a CSG cell is supported from Release 9.

From the network architecture point of view, there is some limitation on connected mode mobility with an HeNB. The limitation includes:

- In Release 9, it is assumed that a CSG cell has no X2 interface; that is, an X2 interface does not exist between a CSG cell and another normal/CSG cell. This means that handover to/from a CSG cell is always S1 handover in Release 9.
- In Release 10, an X2 interface can be established between CSG cells. However, X2 handover between CSG cells can be done only in cases where the CSG ID of both cells is the same, where no access control related to the CSG ID is necessary.

Figure 9.6 shows the procedures for handover to a CSG cell. The details of each step are described below.

Note that among steps 1 to 6, focusing on the radio interface, a normal handover requires only step 4, but the other steps are only necessary for a handover to a CSG cell. Considering the general RRC procedure processing delay and additional delay induced by system information acquisition and reporting, as shown in step 6a and step 6b, handover delay is much longer toward a CSG cell than a normal cell. If the target CSG cell is inter-frequency and thus steps 3 and 4a are triggered, more delay will be added. However, such an additional delay is considered less critical in the sense that the UE is still within the coverage of the source cell.

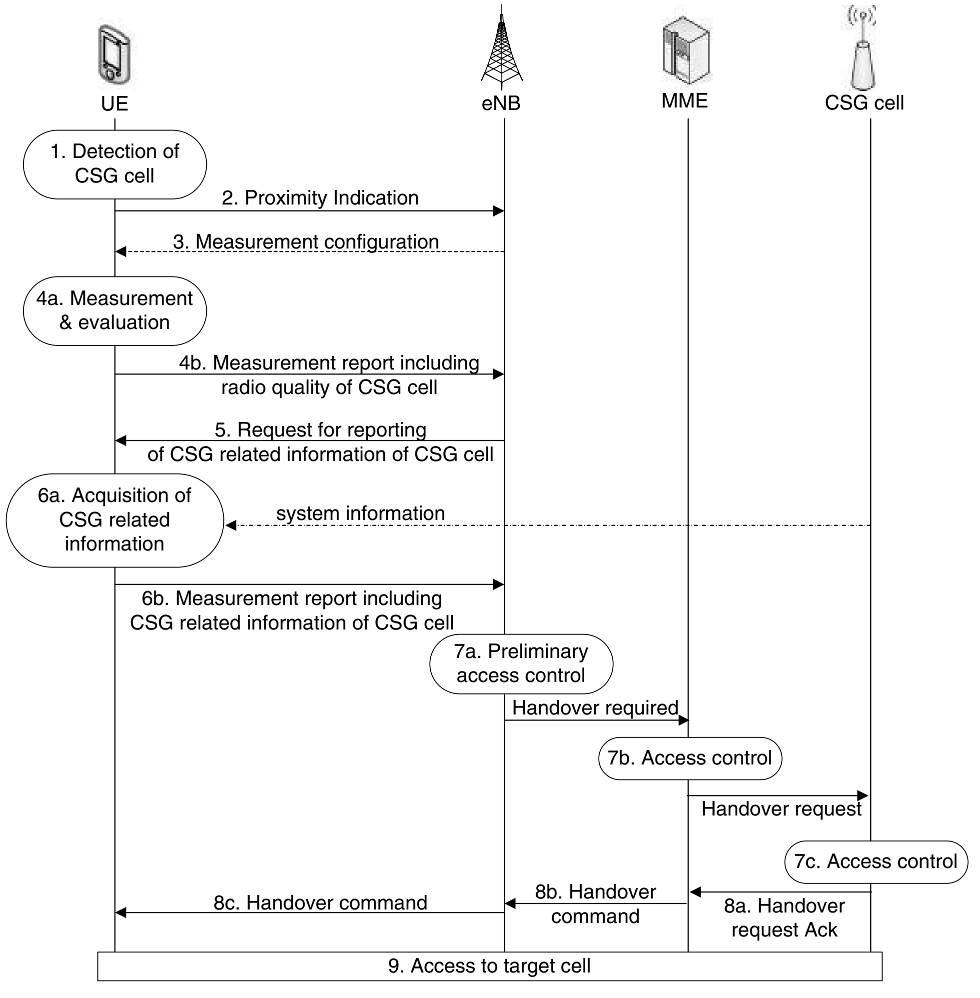

Figure 9.6 Handover to a CSG cell. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

9.5.2.1 Step 1: Detection of a CSG Cell

The handover to a CSG cell can be initiated by detection of a CSG cell in the vicinity of the UE. Detection of a CSG cell can be made by using the autonomous CSG search or a manual CSG selection.

9.5.2.2 Step 2: Proximity Indication

If the UE detects one or multiple member CSG cells in the vicinity, as described in step 1, the UE indicates the presence of the CSG cell(s) to the source cell by sending an RRC message called *ProximityIndication*. Upon receiving the proximity indication from the UE, the source cell becomes aware that the UE is in close proximity to a CSG cell. Note that at this step the details of the target CSG cell are not known to the source cell.

The *ProximityIndication* message includes the following information:

- RAT and the corresponding frequency;
- type of indication, either “entering” or “leaving”.

The *ProximityIndication* message is sent per frequency. The message also indicates whether it has been triggered by the “entering” or “leaving” condition, where “entering” and “leaving” mean that the UE is entering or leaving the proximity of the member CSG cell, respectively.

If the UE detects multiple member CSG cells on multiple frequencies, the UE is allowed to send multiple proximity indications with the limitation that any subsequent proximity indication is sent at least five seconds after the previous indication was sent. The five-second limitation is to suppress the number of proximity indications.

Note that triggering of a proximity indication is solely up to UE implementation, and thus it may be that non-smart UE implementation triggers proximity indications more often than necessary. If that occurred, radio resource would be wasted unnecessarily. To avoid such an undesirable situation, the UE is allowed to send a proximity indication only if the serving cell indicates *reportProximityConfig* via an *RRCConnectionReconfiguration* message. In this manner, undesirable transmissions of proximity indications can be suppressed.

9.5.2.3 Step 3: Measurement Configuration

If the serving cell receives an “entering” proximity indication towards a CSG cell(s) on inter-frequency, it may configure to the UE the proper inter-frequency measurement configuration including the measurement gap, if not configured, by referring to the RAT and frequency information included in the proximity indication. If the serving cell receives a “leaving” proximity indication for a CSG cell on inter-frequency, it may release the measurement configuration that was configured to the UE on entering the proximity of the cell. Note that if the proximity indication is towards a CSG cell(s) on intra-frequency, the measurement configuration step and even the proximity indication step are not necessarily needed.

9.5.2.4 Step 4: Measurement Report Including Radio Quality of a CSG Cell

In this step, the UE performs measurement-related behaviors according to the normal measurement configuration for RRM, as described in Section 3.9. If the measured result of the CSG cell satisfies the reporting criterion, the UE sends a *MeasurementReport* message including the measured result of the CSG cell.

9.5.2.5 Step 5: Request for Additional Information for the Target CSG Cell

Upon receiving the measurement report, the serving cell should check if the target cell indicated in the measurement report is a CSG cell by referring to the PCI. If the measurement report was triggered by a non-CSG cell, the target is identified uniquely by the PCI of the target cell, and the serving cell requests handover preparation to the non-CSG cell.

However, if the measurement report is toward a CSG cell, before triggering handover, the serving cell should be aware of more information on the CSG cell, which is not available at this phase in the serving cell, due to (1) PCI confusion, and (2) preliminary access control.

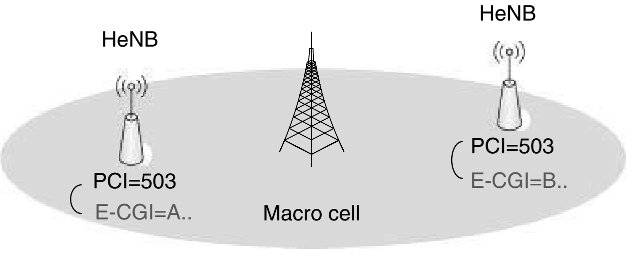

Figure 9.7 PCI confusion for CSG cells

PCI confusion refers to the case where two or more small cells are using the same PCI in the same macro cell area, as shown in Figure 9.7. Due to the nature of possibly uncoordinated deployments and limited PCI resources reserved for CSG cells, it may be that CSG cells are in a state of PCI confusion. In such a case, a CSG cell is not identified uniquely by the PCI. To resolve PCI confusion during handover to a CSG cell, the source cell should be informed of the Global Cell Identifier of the target cell. Since a Global Cell Identifier consists of a PLMN Identity and an E-UTRAN Cell Identifier (28 bits long) that is unique within the PLMN, it can identify the cell uniquely even under PCI confusion.

Preliminary access control is the decision at the source cell regarding whether or not the source cell should continue the handover, depending on the UE's CSG membership. If the target cell is a member CSG cell for the UE, further procedures required for handover should be performed. However, if the target cell is a non-member CSG cell, subsequent procedures for handover should desirably be stopped. If the source cell becomes aware of the membership status of the UE for the CSG cell, the source cell can perform preliminary access control.

To resolve PCI confusion and to enable preliminary access control, the source cell can request the UE to provide additional information on a certain CSG cell. The request for CSG-related information is sent by measurement configuration including *si-RequestForHO* in the reporting configuration of the concerned target RAT. The cell for which the UE is being requested to obtain the Global Cell Identifier is indicated by *cellForWhichToReportCGI* in the measurement object of the concerned target RAT.

9.5.2.6 Step 6: Report of Additional Information on the Target CSG Cell

Upon receiving the measurement configuration for the request of additional information, the UE attempts to acquire the system information by reading the concerned system information from the indicated neighboring CSG cell. The presence of *si-RequestForHO* in the measurement configuration indicates that the UE is allowed to suspend the connection with the serving cell during acquisition of the requested information, which is referred to as the "autonomous gap" [6]. Note that the maximum length of the autonomous gap is 0.15 seconds for an E-UTRA CSG cell, and 2 seconds for a UTRAN FDD CSG cell. The following shows the information that the UE attempts to acquire from the target CSG cell:

- **CSG Identity (csg-Identity):** CSG Identity identifies a closed subscriber group.
- **CSG member status (csg-MemberStatus):** The CSG member status indicates whether the UE is a member of the CSG or not.

- **Target cell identification information (cgi-Info):** The information includes the Global Cell Identifier and the Tracking Area Code, which are used by the source cell to identify the target CSG cell correctly and to prepare handover to the target CSG cell.

Upon completion of the acquisition of all the information above, the UE sends the *MeasurementReport* message including the acquired information to its serving cell.

9.5.2.7 Step 7: Access Control

Access to a CSG cell should be restricted only to CSG member UEs. For this reason, access control based on CSG member status is essential for handover to a CSG cell. CSG-related access control is performed based on the UE CSG white-list and the CSG of the target cell.

In the case of handover to a CSG cell, access control can be carried out in the source cell, the MME, and then in the target CSG cell.

As described in step 5, the source cell can perform preliminary access control based on the CSG member status reported by the UE. If the CSG member status indicates a member CSG cell, the source cell sends to the MME a message requesting handover to the CSG cell by including in the message the Global Cell Identifier and the CSG ID of the cell. Note that the UE or the MME does not transfer CSG white-list information directly to a CSG cell in any case, because the UE CSG white-list is considered to be under the regime of user privacy and thus should not be accessible by customer-premises equipment like an HeNB.

The MME is, in principle, responsible for CSG-related access control because the CSG subscription information of the UE, including the CSG white-list, is available in the MME. Upon receiving the Handover Required message including the CSG ID from the source cell, the MME performs access control. If the UE is verified as being eligible for access to the target CSG cell, the MME sends a Handover Request message to the target HeNB, possibly via the HeNB-GW, and includes the CSG ID in the message. If the access control indicates that the UE is not accessible to the CSG cell, the MME replies to the source cell with a Handover Preparation Failure message.

The target CSG cell, upon receiving the Handover Request message, again verifies that the CSG ID included in the Handover Request message is the same as that broadcast by itself. This verification in the target cell is to prevent access attempts from a fake UE providing a fake CSG ID that can pass the MME membership check but is different from that broadcast by the target CSG cell. In such an invalid case, the target CSG cell would reject the handover.

9.5.2.8 Steps 8–9: Performing Handover

If the validation is successful, the target CSG cell prepares radio resources for handover and sends an *RRCConnectionReconfiguration* message with *MobilityControlInfo* as a handover command to the UE via the source cell (step 8). The UE, upon receiving the handover command, performs normal handover behaviors as described in Section 3.8 (step 9).

9.6 Support for Hybrid Cells

9.6.1 Motivation

The nature of a CSG cell, allowing access only to CSG members, may generate virtual coverage holes to CSG non-member UEs in a macro service area. If there is no other frequency readily available than the frequencies of the CSG cell, and the CSG cell is installed in an area loosely separated in terms of radio propagation, then the CSG non-member UEs near CSG cells will be strongly interfered with by the signals from the CSG cells, and may experience radio link failure.

To preserve the benefit of a CSG cell while avoiding access block issues for CSG non-member UEs, a new mode of operation for an HeNB, called *hybrid mode*, was introduced in Release 9. A cell managed by an HeNB operated in hybrid mode is called a hybrid cell. Access to hybrid cells is allowed for both CSG member and CSG non-member UEs, but the QoS policy may be different depending on the CSG membership status. For CSG member UEs, the hybrid cell may provide prioritized services over CSG non-member UEs.

One possible use case of a hybrid cell is a coffee house chain deploying its own cell in hybrid access mode. VIP customers of the coffee house chain can get preferential access to the cell installed in the coffee house and enjoy a service with higher QoS while others are just granted normal access without any special treatment.

9.6.2 Features

Some features that were specified in Release 8 for support of CSGs have been modified to support hybrid cells in Release 9 and onwards, as described below.

9.6.2.1 Access Mode for Hybrid Cells

Table 9.4 shows the combinations of CSG identification information broadcast by the different types of cell, including hybrid cells. A hybrid cell sets the CSG Indication to FALSE to allow access to CSG non-member UEs, but broadcasts a CSG Identity to allow access for CSG member UEs as CSG members.

9.6.2.2 PCI Range of Hybrid Cells

The PCI of a hybrid cell is set to a PCI value outside the PCI range reserved for CSG cells to ensure that CSG non-member UEs can search the hybrid cell and perform measurements on it for cell reselection.

9.6.2.3 Mobility with Hybrid Cells

The introduction of hybrid cells has only a minor impact on UE behaviors for the mobility procedure. Since a hybrid cell is either a member CSG cell or a non-member CSG cell to the UE, the UE simply follows the behaviors already defined for each case.

Table 9.4 HeNB cell type/access mode identification by UE (Release 9 and onwards)

HeNB Access Mode	Cell Type	CSG Identity	CSG Indication	HNB Name	PCI Split Info for CSG Cells
Open access mode	Normal cell	Absent	FALSE	Absent	Optional
Closed access mode	CSG cell	Present	TRUE	Present	Present
Hybrid access mode	Hybrid cell	Present	FALSE	Present	Optional

- **Mobility in RRC_IDLE:** If the CSG ID broadcast by the hybrid cell is in the UE's CSG white-list with the proper PLMN requirements taken into account, the handling of the hybrid cell in cell reselection should be the same as that of a member CSG cell, because the hybrid cell is a member CSG cell for the UE. This implies that, for hybrid cells that are member CSG cells for the UE, the autonomous CSG search function is used to detect such cells, and reselection rules for CSG cells should be applied. If the CSG ID of the hybrid cell is not present in the UE's CSG white-list, then the hybrid cell should be considered a normal cell, and the normal cell reselection rule should be applied. This implies that the suitability condition applied to CSG cells regarding CSG membership is not applicable for hybrid cells.
- **Mobility in RRC_CONNECTED:** The UE behaviors applicable for member CSG cells for handover are all applicable for hybrid cells for which the UE is a CSG member. For example, if the UE is a member of a hybrid cell, the autonomous search function should be able to detect the hybrid cell according to the requirements specified for member CSG cells, and a proximity indication toward the hybrid cell should be sent to the source cell, if detected.

9.6.2.4 Access Control for Hybrid Cells

Figure 9.8 shows the signaling flow for handover to a hybrid cell. The access control in the case of handover to a hybrid cell is slightly different from that for handover to a CSG cell.

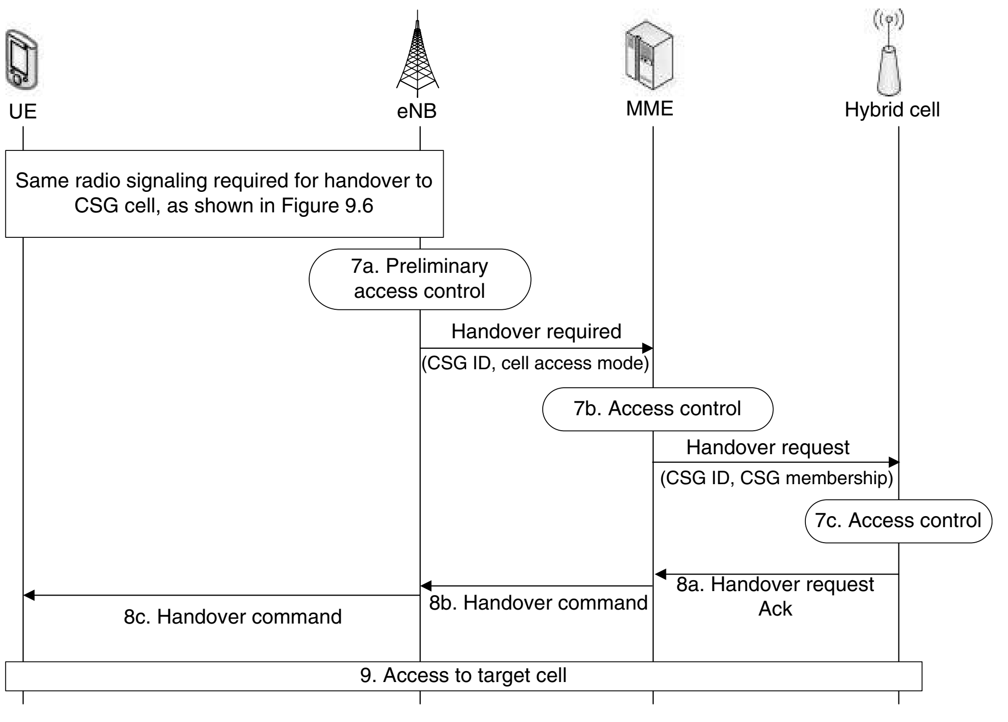

Figure 9.8 Handover to a hybrid cell

For example, after step 7a, the source cell further includes a cell access mode of the target cell in the message requiring handover, if the target cell is a hybrid cell. At step 7b, the MME considers the presence of the cell access mode for access verification towards the hybrid cell. The MME then sets the CSG membership status information of the UE towards the hybrid cell and sends a Handover Request message including the CSG membership information to the target hybrid cell. At step 7c, the target hybrid cell utilizes the CSG membership information received in the Handover Request message to decide whether the UE should be treated as a CSG member or a non-member.

References

1. Nokia, Siemens Networks, Ericsson, Motorola, Alcatel-Lucent, Samsung, Huawei, NEC, “RP-070257: Proposed Study Item on 3G Home Node B”, 3GPP TSG RAN, meeting 35, Lemesos, Cyprus (March 2007)
2. 3GPP Technical Specification 21.905, “Vocabulary for 3GPP Specifications (Release 10)”, www.3gpp.org, accessed on 14 March 2012.
3. 3GPP Technical Specification 24.285, “Allowed Closed Subscriber Group (CSG) list; Management Object (MO) (Release 10)”, www.3gpp.org, accessed on 14 March 2012.
4. 3GPP Technical Specification 22.220, “Service requirements for Home Node B (HNB) and Home eNode B (HeNB) (Release 10)”, www.3gpp.org, accessed on 14 March 2012.
5. 3GPP Technical Specification 23.830, “Architecture aspects of Home Node B (HNB)/Home enhanced Node B (HeNB) (Release 9)”, www.3gpp.org, accessed on 14 March 2012.
6. 3GPP Technical Specification 36.331, “Radio Resource Control (RRC); Protocol Specification”, www.3gpp.org, accessed on 14 March 2012.
7. 3GPP Technical Specification 36.133, “Requirements for support of radio resource management (Release 10)”, www.3gpp.org, accessed on 14 March 2012.

10

Public Warning System (PWS)

3GPP networks support the Public Warning System (PWS). The PWS is used to alert the public to events such as disasters. For instance, when earthquakes, tsunamis, hurricanes, or wild fires occur, the PWS can be used to notify people to evacuate impacted areas within a certain time. In addition, the PWS can be used to notify people of a Child Abduction Emergency (e.g., AMBER alert). PWS notifications should be delivered accurately and in a timely manner to the public in order to help people prepare sufficiently for events.

Different countries may have different requirements on delivery of warning messages, as specified in [1]. Thus, 3GPP networks provide different warning systems for different countries. However, warning systems supported by 3GPP use a common system architecture and common signaling procedures. As listed in Table 10.1, 3GPP standardized different warning systems in different releases, according to requests from corresponding countries.

The Earthquake and Tsunami Warning System (ETWS) was the first warning system to be standardized in Release 8 for Japan. The ETWS was designed based on Japanese requirements focusing on earthquakes and tsunamis. One important requirement of the ETWS is that warning notification providers should be capable of providing Primary and Secondary Notifications to UEs. The Primary Notification should be delivered to UEs within 4 seconds in a Notification Area even in a congestion situation.

The Commercial Mobile Alert Service (CMAS) was standardized in Release 9 for the United States of America. The CMAS was designed to meet the requirements of the Federal Communication Commission (FCC). One important requirement of the CMAS is that multiple notifications should be allowed to be broadcast concurrently on the radio. According to the requirements, the CMAS supports three classes of warning notification: Presidential, Imminent Threat, and Child Abduction Emergency.

Following the ETWS and the CMAS, the Korean Public Alert System (KPAS) was standardized in Release 10 for South Korea, and the EU-ALERT in Release 11 for European countries. The KPAS and the EU-ALERT utilize the procedures defined for the CMAS in both the UE and the network. Warning messages for the CMAS, the KPAS, and the EU-ALERT are delivered to UEs in the same way. Thus, the KPAS and the EU-ALERT have only a minor impact on the PWS – for example, some message identifiers are allocated to the KPAS and the EU-ALERT [2].

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

Table 10.1 The Public Warning Systems supported by 3GPP

Warning System	Target Region	Release
ETWS	Japan	Release 8
CMAS	United States of America	Release 9
KPAS	South Korea	Release 10
EU-ALERT	European countries	Release 11

It should be noted that receipt of PWS Notifications is optional on the UE side. However, UEs may be required to implement the PWS according to regional regulatory requirements.

10.1 Warning System Architecture

Figure 10.1 illustrates the warning system architecture supporting the PWS over the E-UTRAN. The warning system architecture accommodates the Cell Broadcast Center (CBC) and the Cell Broadcast Entity (CBE) that stem from the Cell Broadcast Service (CBS) network architecture in GSM and UMTS. Instead of the interfaces with the Base Station Controller (BSC) and the Radio Network Controller (RNC), the CBC provides a new interface with the MME for warning message delivery over the E-UTRAN.

The CBE is a source of warning messages. Warning messages are formatted in the CBE. The CBE can send information on warning messages to the CBC.

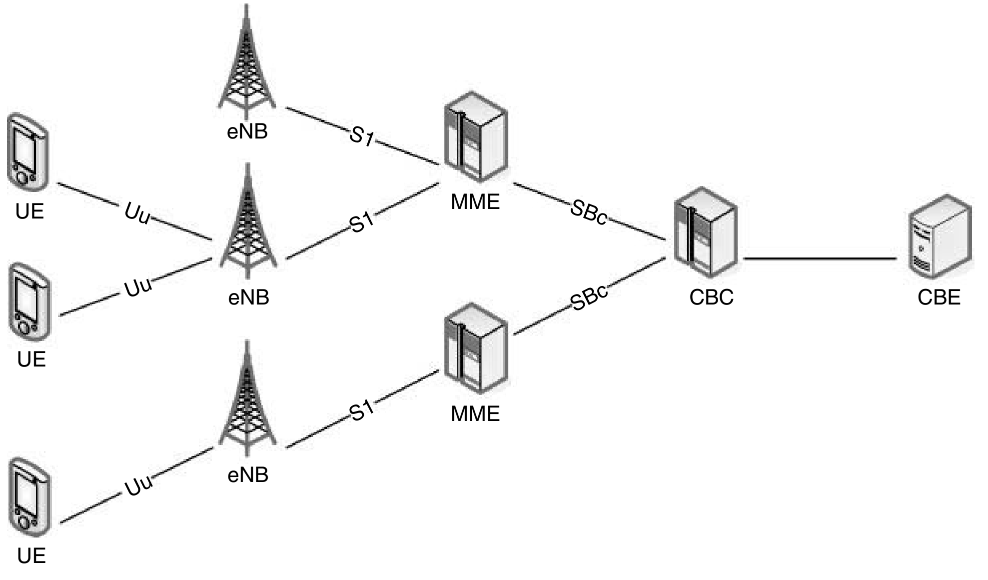

Figure 10.1 Warning system architecture in 3GPP for E-UTRAN. Reproduced by permission of 3GPP, © 2011. 3GPP^{TM} TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

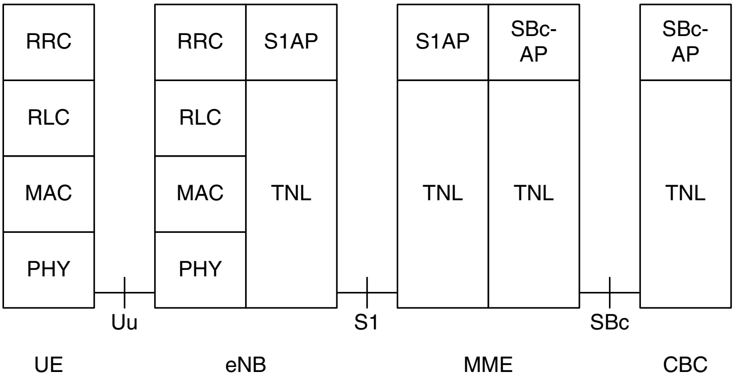

Figure 10.2 Protocol stack for warning message delivery. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

The CBC is responsible for the management of warning messages. The CBC allocates serial numbers to warning messages and initiates broadcast of warning messages by sending the warning messages to the MMEs. The reference point between the CBC and the MME is the SBc interface. The SBc interface is used for warning message delivery and control information signaling.

One or more MMEs and one or more eNBs participate in the delivery of a warning message to UEs. The MME forwards the warning message received from the CBC to the appropriate eNBs via the S1 interface. Then, the eNB broadcasts the warning message via the control plane on the radio interface.

Figure 10.2 explains the protocol stack for warning message delivery. The SBc Application Part (SBc-AP) protocol specified in [3] and the S1 Application Part (S1AP) protocol specified in [4] are used for warning message delivery and control information signaling between the CBC and the MME, and the MME and the eNB, respectively.

The RRC layer is used to deliver a warning message between the eNB and the UE. The RRC layer uses system information to convey the warning message to one or more UEs. Since the system information is transmitted on the BCCH, the warning message bypasses the PDCP layer, utilizes the Transparent Mode RLC layer, and passes the MAC layer transparently. Hence, no AS security is applied and no layer 2 overhead is attached in warning message delivery.

10.2 Warning Messages

Warning messages contain information on events such as disasters. When a UE receives a warning message, it informs the user of the concerned event, such as a tsunami, for example, by displaying the warning message.

A warning message includes some of the following information elements:

- **Message Identifier:** This information indicates the source and the type of a warning message. The coding of the message identifier is shown in [2].
- **Serial Number:** This information is used to identify a change in a warning message with a given message identifier.

- **Warning Type:** This information indicates the type of disaster, either an earthquake or a tsunami. It also indicates how to present this warning message to users.
- **Warning Message Contents:** This information corresponds to user information.
- **Digital Coding Scheme:** This information is used to inform UEs about the alphabet and coding employed for message characters, and message handling at the UE.

A single warning message should include a message identifier and a serial number. Both the message identifier and the serial number are used to identify the warning message. The eNB may repeatedly broadcast the same warning message multiple times, and so UEs may receive the same warning message several times. In this case, UEs can detect duplication of received warning messages by using the message identifiers and the serial numbers. On the UE side, duplicate detection is performed not in the RRC layer but in a layer above the RRC layer. The RRC layer just forwards the received PWS Notifications with the associated message identifiers and serial numbers to the upper layer.

The warning type is used only for an ETWS Primary Notification. The warning type simply indicates to UEs whether the concerned disaster corresponds to an earthquake, a tsunami, or something else. Since the intention of the ETWS Primary Notification is prompt delivery of emergency information, warning messages delivered as ETWS Primary Notifications do not contain detailed information on the concerned event.

The size of the warning message contents ranges from 1 byte to 9600 bytes. However, 9600 bytes is too large to be transmitted on the radio. Thus, the RRC layer is allowed to perform segmentation of a single warning message.

Different notifications may have different sets of information elements in their warning messages. Table 10.2 shows what is included in a warning message for an ETWS Primary Notification, an ETWS Secondary Notification, and a CMAS Notification.

It should be noted that warning security information was optionally included in the ETWS Primary Notification for the security of warning messages. This was specified from

Table 10.2 Information elements of PWS Notifications broadcast on the radio

PWS Notification	Information Elements
ETWS Primary Notification	Message Identifier Serial Number Warning Type
ETWS Secondary Notification	Message Identifier Serial Number Warning Message Contents Digital Coding Scheme
CMAS Notification	Message Identifier Serial Number Warning Message Contents Digital Coding Scheme

Release 8. However, 3GPP decided to invalidate the warning security information because it provided a low level of security. Thus, warning security information is no longer valid in an ETWS Primary Notification.

3GPP plans to enhance the security of PWS messages with a higher level of security in Release 11 or later. This enhanced security is expected to be applied not only to an ETWS Primary Notification, but also to other types of PWS Notification.

10.3 Delivery of Warning Messages on a Network

A warning message is delivered from the CBE to the eNB over the SBc-AP and the S1AP protocols via the CBC and the MME on the network side. The network can update or stop an ongoing warning message using the same protocols.

The following SBc-AP procedures are used for warning message delivery between the CBC and the MME:

- Write-Replace Warning procedure;
- Stop Warning procedure.

The following S1AP procedures are used for warning message delivery between the MME and the eNB:

- Write-Replace Warning procedure;
- Kill procedure.

10.3.1 Warning Message Delivery Procedure

The Write-Replace Warning procedures over the SBc interface and the S1 interface are used to start or overwrite broadcast of a warning message for PWS Notification, as illustrated in Figure 10.3.

Upon a request from the CBE, the CBC initiates the Write-Replace Warning procedure by sending a Write-Replace Warning Request message to the MME. The Write-Replace Warning Request message from the CBC may include a list of tracking areas. When the MME receives the Write-Replace Warning Request message from the CBC, if a list of tracking areas is available, the MME forwards the Write-Replace Warning Request message to the eNBs belonging to the listed tracking areas. Otherwise, the Write-Replace Warning Request message is forwarded to all the eNBs connected to the MME.

The Write-Replace Warning Request message includes information elements, as listed in Table 10.2. In addition, the CBC can determine the following information elements for delivery of the concerned warning message, based on information received from the CBE. The CBC can include them in the Write-Replace Warning Request message sent to the eNB via the MME. The eNB uses them to broadcast the concerned warning message to UEs.

- Repetition Period;
- Number of Broadcasts Requested;
- Warning Area List;
- Concurrent Warning Message Indicator.

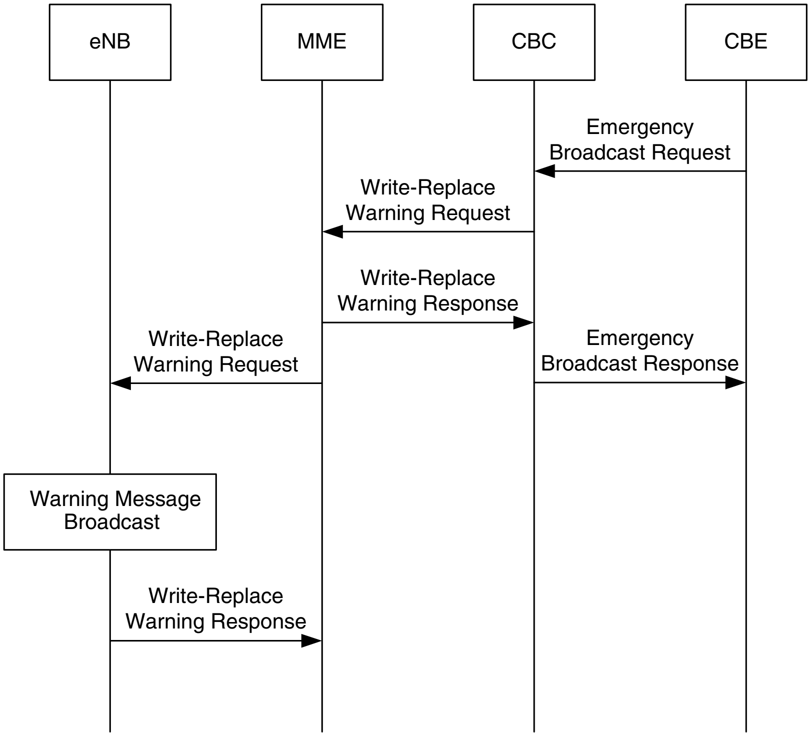

Figure 10.3 Warning message delivery procedure. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

The Repetition Period informs the eNB of the periodicity of the concerned warning message to be broadcast.

The Number of Broadcasts Requested informs the eNB how many times the same warning message should be broadcast. The eNB periodically and repeatedly broadcasts the concerned warning message according to the Repetition Period and the Number of Broadcasts Requested.

The Warning Area List informs the eNB where the concerned warning message should be broadcast. The Warning Area List can indicate the list of emergency areas defined by the operator, the list of tracking areas, or the list of cells. The eNB uses the Warning Area List to determine the cells where the concerned warning message is to be broadcast.

The Concurrent Warning Message Indicator informs the eNB whether or not the concerned warning message replaces an ongoing old warning message. If the Concurrent Warning Message Indicator is not present in the received Write-Replace Warning Request message, the eNB replaces the ongoing old warning message with the concerned warning message newly received from the MME. If the Concurrent Warning Message Indicator is present in the message, the concerned warning message is broadcast with the ongoing old warning message.

The eNB can receive more than one Write-Replace Warning Request message for the same warning message from different MMEs if the eNB is connected to more than one MME. To avoid duplicated broadcast, the eNB detects duplication of the same warning message by checking the message identifier and the serial number within the concerned warning message.

The eNB sends the Write-Replace Warning Response message to the MME in response to the Write-Replace Warning Request message. The Write-Replace Warning Response message may include the Broadcast Completed Area List that indicates the areas where the concerned warning message has been broadcast successfully. The Broadcast Completed Area List uses the list of emergency areas defined by the operator, or the list of tracking areas, or the list of cells.

10.3.2 Warning Message Cancel Procedure

The CBE can request the CBC to stop broadcasting warning messages, for example, when the concerned disaster settles down. When this request is triggered, the network initiates the Stop Warning procedure over the SBc interface together with the Kill procedure over the S1 interface to stop broadcasting warning messages, as illustrated in Figure 10.4.

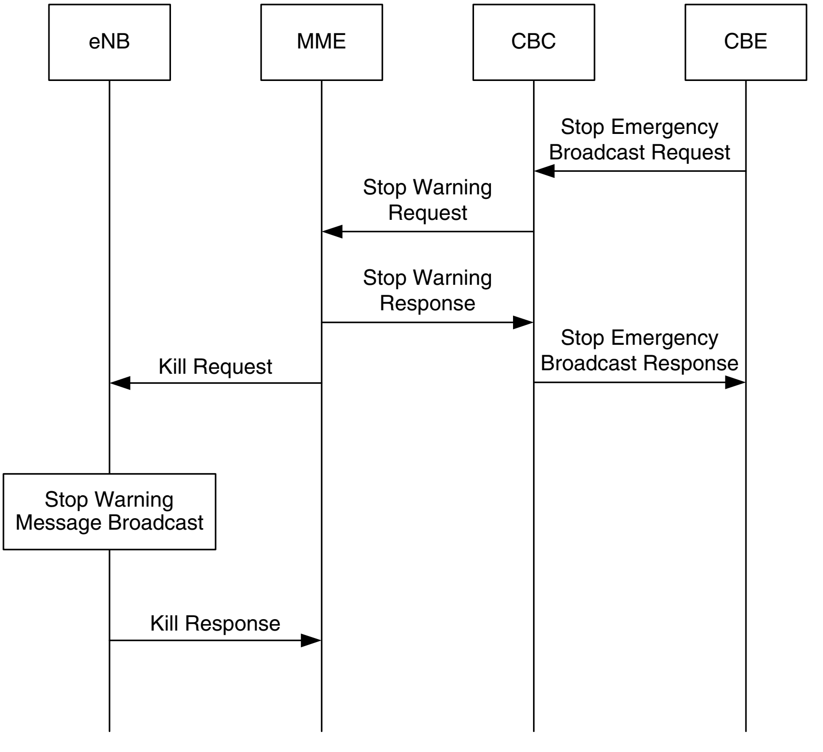

Figure 10.4 Warning message cancel procedure. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

10.4 Delivery of Warning Messages over the Radio Interface

Figures 10.5 and 10.6 illustrate examples of ETWS Notifications and CMAS Notifications in E-UTRA, respectively. Upon receiving the Write-Replace Warning Request message from the MME, the eNB starts to broadcast the warning message to UEs on the radio interface.

When the UE is capable of receiving PWS Notifications, the UE is allowed to receive PWS Notifications both in RRC_IDLE and in RRC_CONNECTED. In RRC_IDLE, the UE supports reception of PWS Notifications at acceptable cells with limited service as well as at suitable cells.

The ETWS reception capability and CMAS reception capability are independent. In other words, the UE should be capable of receiving ETWS Notifications to enable receipt of ETWS Notifications, and the UE should be capable of receiving CMAS Notifications to enable receipt of CMAS Notifications.

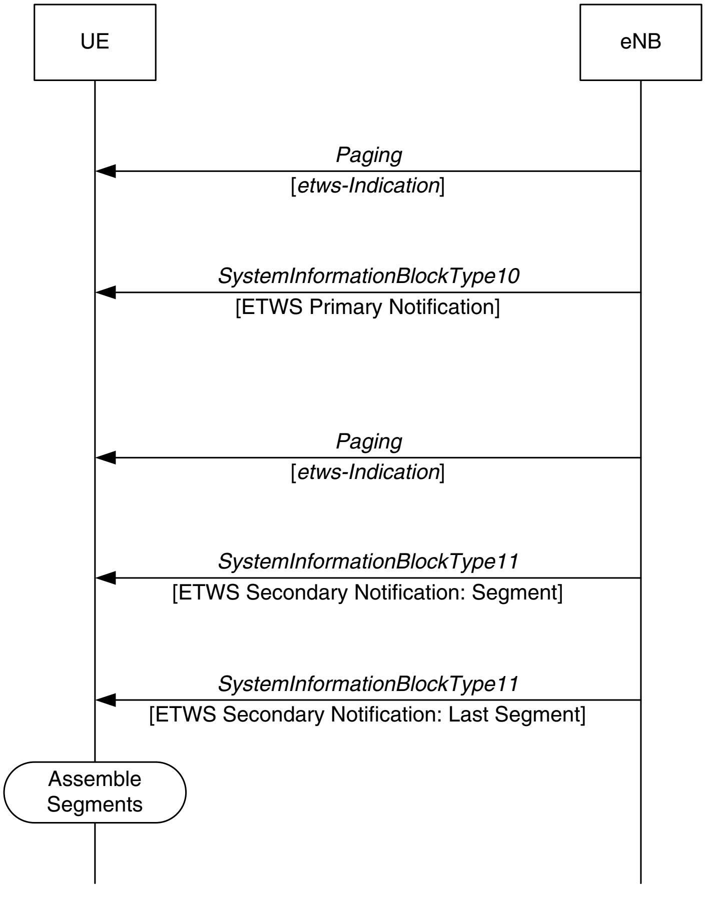

Figure 10.5 An example of ETWS Notifications in E-UTRA. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

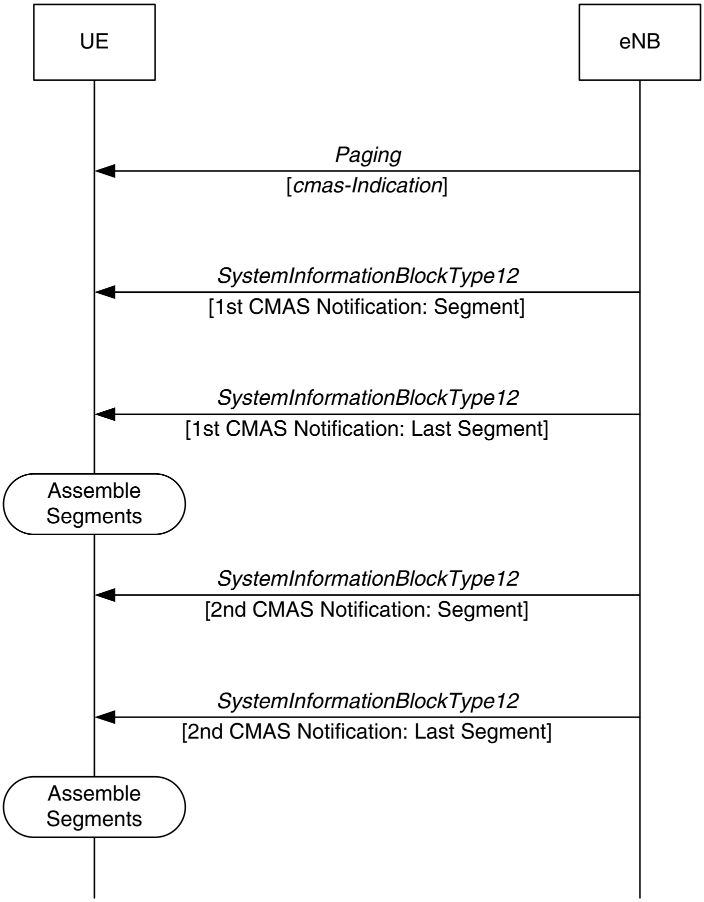

Figure 10.6 An example of CMAS Notifications in E-UTRA. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

10.4.1 PWS Notifications in System Information

The warning message is transmitted not on the user plane, but on the control plane of the radio interface, because the E-UTRAN does not have any common traffic logical channel except for MBMS-related logical channels.

On the control plane of the radio interface, the E-UTRAN uses system information on the BCCH mapped to the DL-SCH to broadcast a warning message. Different types of PWS Notification use different types of System Information Blocks for warning message delivery from the eNB to the UEs, as specified in [5].

- **SystemInformationBlockType10:** The ETWS Primary Notification is contained in this type of system information block.

- **SystemInformationBlockType11:** The ETWS Secondary Notification is contained in this type of system information block.
- **SystemInformationBlockType12:** The CMAS Notification is contained in this type of system information block.

It is important that PWS Notifications should be delivered quickly to the UEs because they concern warning situations. In particular, it is one of the requirements to deliver an ETWS Primary Notification to the UEs within 4 seconds. However, the network is allowed to configure a value longer than 4 seconds as the length of the BCCH modification period used for periodic broadcast of system information. Thus, if such a value were to be configured for some cells, the requirement for delivery of an ETWS Primary Notification within 4 seconds could not be met.

Accordingly, the concept of the BCCH modification period is not applied to SIB10, SIB11, and SIB12, unlike the other SIBs. When a new PWS Notification arrives at the eNB, the eNB is allowed to update SIB10, SIB11, and SIB12 at any point in time within a BCCH modification period. The eNB does not update the value tag of SIB1 for delivery of the PWS Notification on system information.

10.4.2 Indication of PWS Notifications in Paging

The eNB indicates broadcast of PWS Notifications to UEs by using the Paging procedure. The Paging procedure is used to alert UEs quickly to PWS Notifications. When a UE recognizes an indication in the *Paging* message, the UE immediately receives the PWS Notification, which enables users to be alerted promptly. It is required that a UE in RRC_CONNECTED attempts to read the *Paging* message at least once every default paging cycle to check whether or not an indication is present in the *Paging* message. A UE in RRC_IDLE monitors the *Paging* message on its paging occasion every paging DRX cycle. Thus, the length of the paging cycle will determine how promptly users obtain the warning message.

Different indications are provided in the *Paging* message for ETWS Notifications and CMAS Notifications. The *Paging* message includes an *etws-Indication* for indicating broadcast of an ETWS Primary Notification and/or an ETWS Secondary Notification, and a *cmas-Indication* for indicating broadcast of a CMAS Notification.

When receiving a *Paging* message including an *etws-Indication*, only ETWS-capable UEs immediately receive the ETWS Notification via system information, as shown in Figure 10.5. When receiving a *Paging* message including a *cmas-Indication*, only CMAS-capable UEs immediately receive the CMAS Notification via system information, as shown in Figure 10.6. Upon receipt of the *Paging* message, UEs do not need to wait until the next modification period to receive PWS Notifications in the system information, because UEs should acquire warning messages as soon as possible.

10.4.3 Segmentation of Warning Messages

The RRC layer can perform segmentation of a warning message contained in *SystemInformationBlockType11* and *SystemInformationBlockType12* because the size of the warning message may be too large to be carried within one system information block. However, no segmentation of *SystemInformationBlockType10* is applied because the size of an ETWS Primary Notification is small enough to be carried within one system information block.

When a warning message is segmented into more than one segment, the same segment size should be applied to a given segment of the warning message that is broadcast repeatedly with the same message identifier and the same serial number at the cell. Thus, when the UE loses a certain segment of the warning message from the first transmission of the warning message, the UE can acquire the lost segment from a retransmission of the warning message and then successfully assemble the received segments to make the complete warning message.

Note that different cells can apply different sizes of segment for a warning message with the same message identifier and the same serial number. This means that UEs cannot combine segments across different cells for the same warning message.

The ETWS does not support concurrent broadcasts of multiple warning messages. In other words, a cell is allowed to broadcast at most one ETWS Primary Notification and at most one ETWS Secondary Notification. When a new ETWS Primary Notification arrives at the eNB, it replaces an old ETWS Primary Notification, if available. In addition, when a new ETWS Secondary Notification arrives at the eNB, it replaces an old ETWS Secondary Notification, if available. Hence, when any new warning message is received for an ETWS Primary Notification or an ETWS Secondary Notification, the UE should discard any previously buffered segment of the old warning message for the ETWS Notification.

Unlike ETWS Notifications, the CMAS supports concurrent broadcasts of multiple CMAS Notifications from a single cell, because there is a requirement that the CMAS should support concurrent transmissions of different warning messages at any one time. This means that if segmentation is applied to multiple CMAS Notifications, different sets of segments will be broadcast concurrently for different warning messages from a cell. Thus, a UE capable of receiving CMAS Notifications should be able to assemble the segments in parallel for different warning messages. In other words, while the UE receives and assembles segments for one CMAS Notification, the UE should be able to receive and assemble segments for another CMAS Notification.

It should be noted that when the broadcast of one warning message follows the broadcast of another warning message for two different CMAS Notifications, all the segments of one warning message should be broadcast before segments of another warning message start to be broadcast. Segments of one CMAS Notification are not interleaved with segments of another CMAS Notification.

References

1. 3GPP Technical Specification 22.268, "Public Warning System (PWS) requirements (Release 11)", www.3gpp.org, accessed on 14 March 2012.
2. 3GPP Technical Specification 23.041, "Technical realization of Cell Broadcast Service (CBS) (Release 11)", www.3gpp.org, accessed on 14 March 2012.
3. 3GPP Technical Specification 29.168, "Cell Broadcast Centre interfaces with the Evolved Packet Core; Stage 3 (Release 11)", www.3gpp.org, accessed on 14 March 2012.
4. 3GPP Technical Specification 36.413, "S1 Application Protocol (S1AP) (Release 10)", www.3gpp.org, accessed on 14 March 2012.
5. 3GPP Technical Specification 36.331, "Radio Resource Control (RRC); Protocol Specification (Release 10)", www.3gpp.org, accessed on 14 March 2012.

11

Multimedia Broadcast/Multicast Service (MBMS)

The MBMS service is a point-to-multipoint service provided by 3GPP systems including E-UTRAN, UTRAN, and GERAN. The MBMS provides user data from a single network entity to multiple UEs. MBMS services are transmitted on the same carrier frequencies used by mobile operators to transmit normal services such as voice calls and the Internet. Operators can provide various types of services as MBMS services, for example, TV broadcasting, message distribution, and file downloading. When it comes to TV broadcasting, the MBMS offers an opportunity for mobile operators to be able to provide TV broadcasting services to users without buying frequency bands licensed for TV broadcasting.

The MBMS is supported from Release 9 in E-UTRAN. 3GPP started standard work on MBMS over E-UTRAN for Release 8, but it was finally specified in Release 9 due to lack of time during the discussions for Release 8. The MBMS in E-UTRAN is also known as the eMBMS (Evolved MBMS).

Only basic operations of the MBMS were specified in Release 9, compared to what are specified in UTRAN. For example, no uplink message related to MBMS is defined in Release 9. MBMS features have been enhanced in Release 10 and Release 11. Enhancements to the MBMS after Release 9 are described in Section 11.6.

11.1 MBMS Services

Within 3GPP systems, operators can provide MBMS user services, which are defined as MBMS services provided to end-users. The MBMS user service supports various media types such as text, images, video, speech, and audio. End-users can be provided with streaming services, file downloading services, TV services, and so on, through MBMS user services.

To provide MBMS user services, 3GPP systems use MBMS bearer services, which are defined as services provided over IP multicast in the PS domain of 3GPP systems. One MBMS user service is transported over one or more MBMS bearer services on the network. For example, end-users may be provided with one MBMS user service, such as a TV service, over two MBMS bearer services, one for audio and one for video.

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

The MBMS bearer service offers two modes in 3GPP systems [1]:

- broadcast mode;
- multicast mode.

For an MBMS bearer service in multicast mode, only member UEs which have joined the MBMS bearer service can receive the service. Thus, when the UE is interested in an MBMS bearer service, the UE should join the MBMS bearer service to become a member of a multicast group, which is a group of UEs interested in the MBMS bearer service. For a UE that has joined, UE-specific information related to a particular MBMS bearer service is established on the network. Then, while the UE performs handover, the UE-specific information is transferred from one node to another node on the network side. Joining and the management of UE-specific information related to the MBMS bearer service introduce non-trivial complexity for the network in terms of providing MBMS bearer services in multicast mode.

For an MBMS bearer service in broadcast mode, UEs can receive the MBMS bearer service without joining. The broadcast mode is less complex than the multicast mode because some of the signaling flows required in the multicast mode are not necessary in the broadcast mode. However, in the broadcast mode, the network may not know where UEs interested in the MBMS bearer services are located. Thus, the network may broadcast MBMS bearer services blindly, wasting radio and network resources. The network may be able to avoid broadcasting certain MBMS bearer services, for example in a rural area due to lack of audience. However, in such a case, a certain UE that suddenly comes into that area cannot receive MBMS bearer services of interest if they are not broadcast in that area. As a result, the broadcast mode may degrade user experience from time to time.

The GERAN and the UTRAN support both the broadcast mode and the multicast mode. However, the E-UTRAN provides MBMS bearer services in the broadcast mode only. The main reason why the multicast mode is not supported in the E-UTRAN is to simplify MBMS operation.

The following subsections explain how the EPC and the E-UTRAN provide MBMS services. For simplicity, the term “MBMS service” is used to mean an MBMS bearer service in the following subsections. When the term “MBMS service” is intended to mean an MBMS user service, it is unambiguously written out as “MBMS user service”.

11.2 Architecture and Functions for MBMS

The following functional entities are introduced to the EPC and E-UTRAN for support of the MBMS:

- the BM-SC (Broadcast-Multicast Service Center);
- the MCE (Multi-cell/multicast Coordination Entity);
- the MBMS Gateway.

The BM-SC serves as a source of MBMS transmissions in 3GPP systems. The BM-SC is a functional entity for provisioning and delivery of an MBMS user service by initiating transmission of an MBMS bearer service related to the MBMS user service. The BM-SC starts or stops transmission of a session for the MBMS bearer service by using “MBMS session control signaling”, such as MBMS Session Start and MBMS Session Stop.

In addition, the BM-SC can initiate a service announcement. A service announcement informs UEs about forthcoming MBMS user services. Service announcements can be realized by SMS, PUSH, IETF protocols such as HTTP, and so on [2].

The MCE participates in the MBMS session control signaling and performs admission control and radio resource allocation. When a new session of an MBMS bearer service starts, the MCE makes a decision on whether or not to establish a radio bearer for the session depending on, for example, available radio resources and the priority of the MBMS bearer service. The MCE also controls the MBMS Counting procedure from Release 10 (see Section 11.6.1).

In addition, the MCE determines radio configurations of MBMS bearer services and signals them to eNBs, and the eNB signals the radio configurations sent by the MCE to UEs via RRC messages. The MCE should set the same configurations across all eNBs broadcasting the same MBMS service, in terms of configurations of the RLC layer, the MAC layer, and the Physical layer.

The MBMS Gateway exists between the BM-SC and the eNBs for forwarding MBMS user data to eNBs, and exists between the MME and eNBs for MBMS session control. The MBMS Gateway uses the IP Multicast protocol for forwarding MBMS user data to eNBs that have joined an MBMS service. The MBMS Gateway serves as the source of IP multicast distribution toward eNBs, and allocates an IP multicast address to an eNB that joins an MBMS service.

The additional network entities for MBMS introduced the following new reference points, as shown in Figure 11.1:

- M1: the reference point between the MBMS Gateway and the eNB for the user plane which is used to forward MBMS user data;

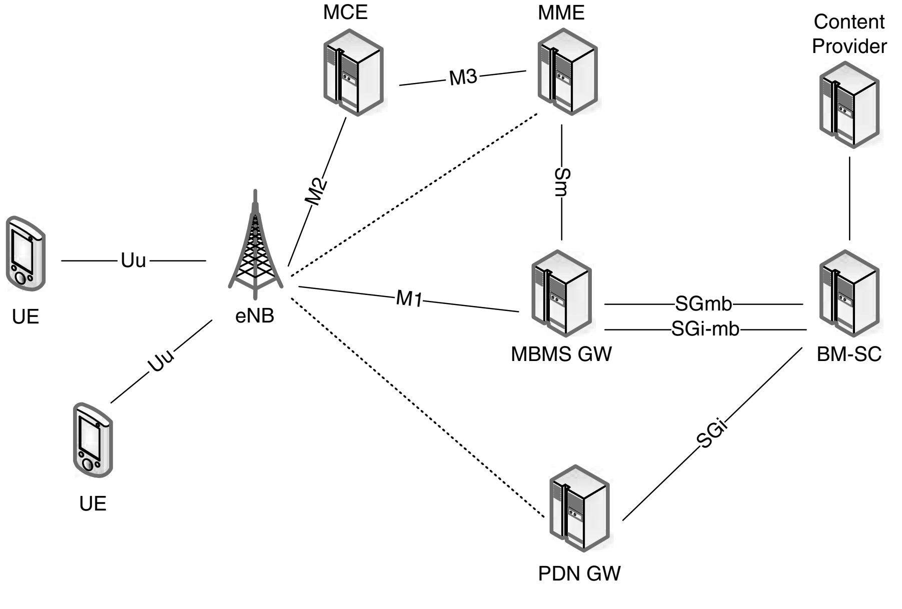

Figure 11.1 MBMS reference architecture for E-UTRAN. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

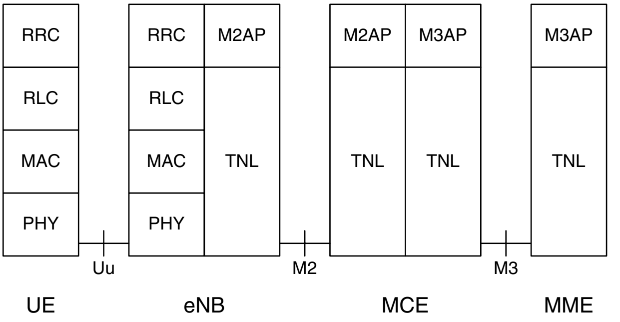

Figure 11.2 Protocol stack for the MBMS control plane. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

- M2: the reference point between the MCE and the eNB for the control plane;
- M3: the reference point between the MME and the MCE for the control plane;
- Sm: the reference point between the MME and the MBMS Gateway for the control plane;
- SGi-mb: the reference point between the BM-SC and the MBMS Gateway for the user plane which is used to forward MBMS user data;
- SGmb: the reference point between the BM-SC and the MBMS Gateway for the control plane;
- SGi: the reference point between the BM-SC and the PDN Gateway for the user plane which is used to realize MBMS user services over EPS bearers.

The protocol stack for the MBMS control plane is shown in Figure 11.2. The M2 Application Protocol (M2AP) is used to control MBMS-related signaling flows between the eNB and the MCE. The M3 Application Protocol (M3AP) is used to control MBMS-related signaling flows between the MCE and the MME. The M2AP and the M3AP are normally used for communication of MBMS session control signaling such as Session Start and Session Stop.

The protocol stack for the MBMS user plane is shown in Figure 11.3. The IP multicast protocol is used over the M1 interface for forwarding MBMS user data from the MBMS

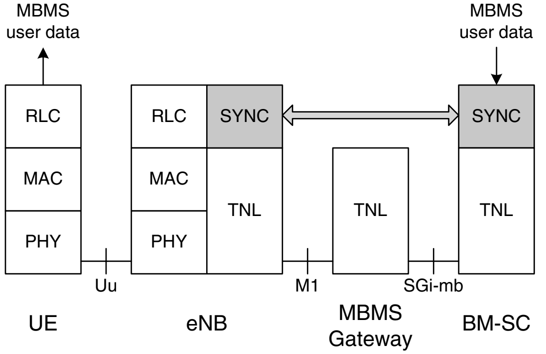

Figure 11.3 Protocol stack for the MBMS user plane. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

Gateway to the eNBs. In addition, the MBMS synchronization (SYNC) protocol is used to convey MBMS user data from the BM-SC to the eNBs for coordinating transmissions of the same MBMS user data from different eNBs.

For the SYNC protocol, which is specified in [3], the BM-SC includes synchronization-related information such as time stamps in SYNC PDUs conveying MBMS user data. When the BM-SC sets a time stamp, the BM-SC should take into account transmission delay from the BM-SC to the eNB, processing delay in the eNB, and so on. The eNBs use time stamps included in SYNC PDUs for synchronized transmissions from multiple cells on the radio interface. When the SYNC protocol conveys more MBMS user data than the amount of data that the eNB can transmit on the radio interface, the eNB is allowed to drop some of the MBMS user data.

11.3 MBSFN Transmissions

The MBMS can be provided in a point-to-multipoint manner over the radio interface. A point-to-multipoint transmission can be realized by an MBMS over Single Frequency Network (MBSFN) transmission, which is a simulcast transmission technique where identical waveforms are transmitted simultaneously from multiple cells. All cells involved in MBSFN transmissions should be time synchronized and deliver the same contents to UEs at the same time.

MBSFN transmissions from multiple cells are combined in the UE such that it is seen as a single transmission after combining on the UE side. The combining of MBSFN transmissions provides UEs with a diversity gain, and the diversity gain is especially important for MBMS services at cell boundaries where UEs' reception performance may be degraded due to weak signal strength. This benefit of combining MBSFN transmissions exists only when multiple cells are synchronized for transmissions of the same content. However, the E-UTRAN does not preclude the case where MBSFN transmission is provided only from a single cell, in which case there is no diversity gain from MBSFN transmissions. Operators can provide such an MBSFN transmission only from a single cell when MBMS services are provided to UEs only in the area of the single cell.

MBMS services are broadcast on one or more carriers where non-MBMS services such as unicast services are provided. Within the carriers, MBMS services and non-MBMS services are provided in different subframes. The E-UTRAN does not support the "MBMS-dedicated carrier", which provides only MBMS services and excludes non-MBMS services.

Unlike non-MBMS transmissions, MBMS services should be transmitted only via a specific type of subframe, called an "MBSFN subframe" [4]. MBSFN transmissions can take place only in these MBSFN subframes. All cells participating in MBSFN transmissions should configure MBSFN subframes to provide MBMS services. Configuration of MBSFN subframes is cell specific. However, for a single MBMS service, all cells should configure the same MBSFN subframes to ensure synchronized transmission.

Note that MBSFN subframes can be used not only for MBMS transmissions but also for non-MBMS transmissions such as relay backhaul transmissions and unicast transmissions. Up to Release 9, UEs consider unicast transmissions as not being assigned to MBSFN subframes, but some UEs may be able to receive unicast transmissions in MBSFN subframes from Release 10, as specified in [5].

Synchronous transmissions of MBMS services from multiple cells are performed within an MBSFN area consisting of one or more cells, as shown in Figure 11.4. Cells belonging to

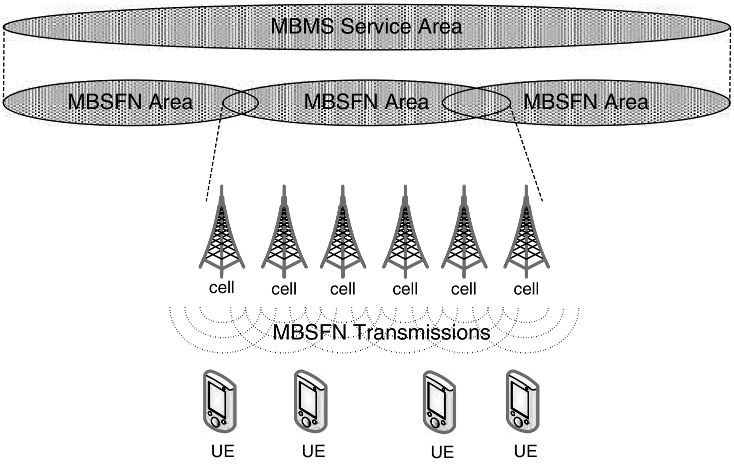

Figure 11.4 MBSFN transmissions over an MBSFN area. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

the same MBSFN area should be synchronized to participate in MBSFN transmissions of MBMS control information and MBMS services. Different MBSFN areas can overlap, such that a single cell can belong to multiple MBSFN areas, up to a maximum of eight.

Different cells can have different combinations of MBSFN areas. For example, one cell might cover MBSFN area #1 and MBSFN area #2, but another cell might cover MBSFN area #2 and MBSFN area #3. Each cell informs UEs of the MBSFN subframe configuration of the cell via system information.

On top of MBSFN areas, the network defines an MBMS service area where a certain MBMS service is broadcast. One MBMS service area consists of one or more cells and corresponds to one or more MBSFN areas. The BM-SC informs network entities about the MBMS service area corresponding to the session. The network entities can use the MBMS service area to determine the eNBs to which the corresponding MBMS service should be distributed on the network side.

11.4 Radio Protocols for MBMS

11.4.1 Layers 1 and 2 for MBMS

The E-UTRAN provides two point-to-multipoint logical channels for the MBMS on the radio interface:

- the Multicast Control Channel (MCCH);
- the Multicast Traffic Channel (MTCH).

The MCCH exists in the control plane of the radio interface between the UE and the eNB. This logical channel is used to broadcast MBMS-related RRC messages corresponding to an

MBSFN area. The RRC layer of the eNB creates RRC messages based on MBMS control information received from the MCE.

The MCCH is mapped to a transport channel called the Multicast Channel (MCH), which is mapped to a physical channel called the Physical Multicast Channel (PMCH). When the MCCH is mapped to the MCH, only one MCCH exists for one MCH.

The MTCH exists in the user plane of the radio interface between the UE and the eNB. This logical channel is used to broadcast MBMS user data related to an MBMS service. The eNB receives MBMS user data from the BM-SC via the MBMS Gateway by using the SYNC protocol.

The MTCH is also mapped to the MCH, and the MCH is mapped to the PMCH. There is one-to-one correspondence between an MBMS service and an MTCH. More than one MTCH can be mapped to one MCH in order to obtain a multiplexing gain in MBMS data transmission.

There is one-to-one correspondence between an MBSFN area and an MCCH. Thus, the E-UTRAN provides different MCCHs for different MBSFN areas. As shown in Figure 11.5, when multiple MBSFN areas overlap, one cell should provide multiple MCCHs. Different MCCHs should be broadcast in different MBSFN subframes.

A cell supporting MBMS uses the PDCCH to notify UEs about a change in "MCCH information"; that is, MBMS control information delivered via one or more MCCHs. For MCCH information change notifications, the MBMS Radio Network Temporary Identifier (M-RNTI) is carried on the PDCCH. Even though multiple MBSFN areas can overlap at a cell, only a single M-RNTI on the PDCCH is used at a cell in the E-UTRAN. The PDCCH carrying the M-RNTI also indicates the MBSFN area(s) for which the change of MCCH information is relevant.

The intention of using the M-RNTI on the PDCCH is to enable a UE to save battery consumption by monitoring the PDCCH only, before receiving user data of a session for an MBMS service of interest. Once the UE starts to receive user data for the session, the UE

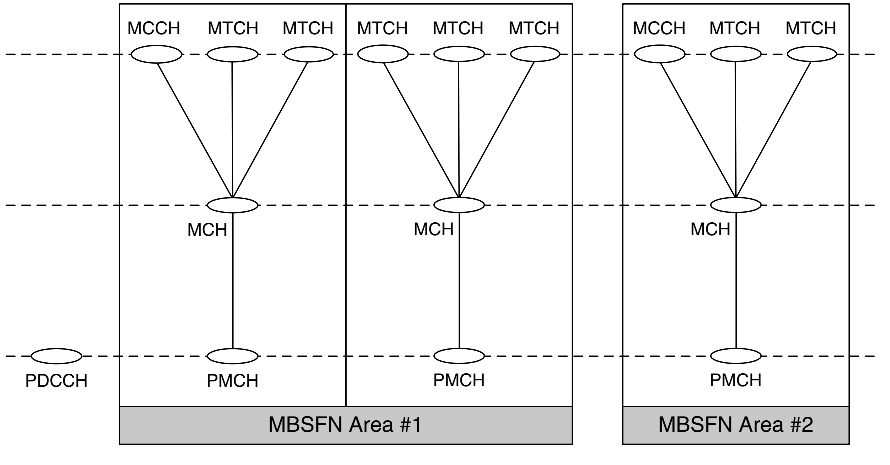

Figure 11.5 Example of MBMS channel mapping for multiple MBSFN areas at a cell. Reproduced by permission of 3GPP, © 2011. 3GPP^{TM} TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

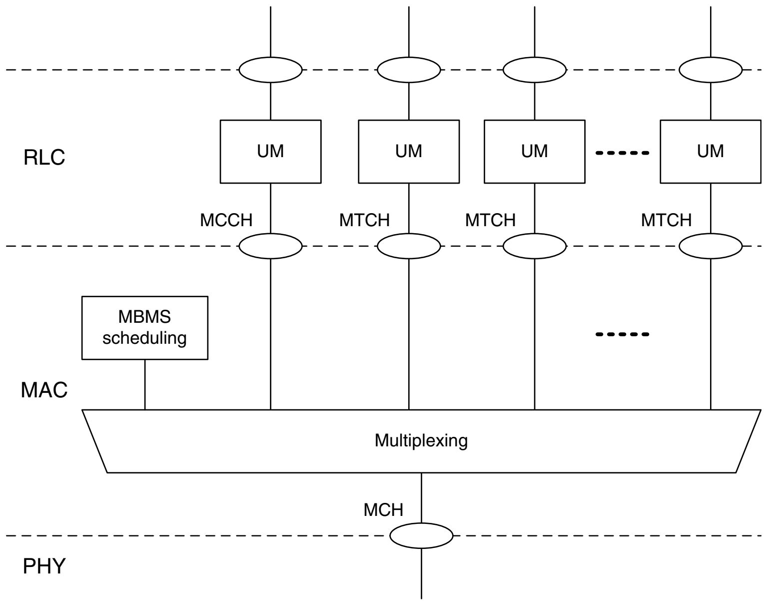

Figure 11.6 Layer 2 model for one MBSFN area on the eNB side. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

continuously receives not only user data for the session on an MTCH, but also MCCH information on an MCCH related to the session. Thus, the eNB indicates the M-RNTI on the PDCCH only when MCCH information changes due to a session start or a counting request.

As shown in Figure 11.6, the PDCP layer is not applicable for either the MCCH or the MTCH. Therefore, header compression and AS security are not supported for either MCCH information or MBMS user data.

The RLC layer configures the Unacknowledged Mode (UM) for both the MCCH and the MTCH. Re-ordering is not applicable for RLC entities configured for MCCH and MTCH because HARQ is not used for data delivery through the MCCH and MTCH.

In the MAC layer, the eNB can multiplex one or more MTCHs together with one MCCH. The LCID field in the MAC PDU indicates whether the corresponding part is for the MCCH or MTCH. For the MCCH and the MTCH, the MAC layer in the eNB delivers at most one transport block (i.e., one MAC PDU) to the Physical layer for one TTI.

The MCH Scheduling Information MAC Control Element is used to inform UEs when MBMS service user data are scheduled for each MTCH. The MCH Scheduling Information MAC Control Element delivered over one MCH includes scheduling information for the MTCH that is delivered over the MCH.

11.4.2 Layer 3 for MBMS

The RRC layer broadcasts MCCH information on an MCCH to inform UEs about MBMS control information. The MCCH information is broadcast periodically for each MCCH in

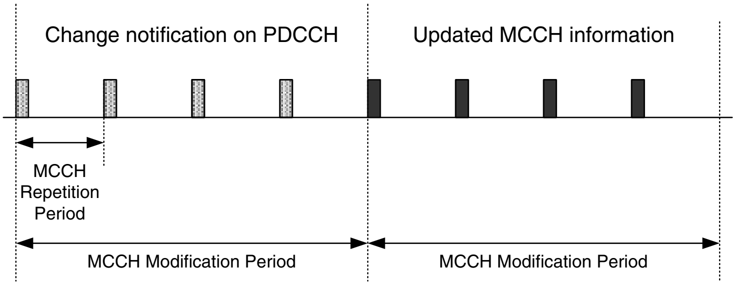

Figure 11.7 Update of MCCH information. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

each MBSFN area. For the periodical broadcast of MCCH information, the RRC layer uses the concepts of the MCCH repetition period and the MCCH modification period, as illustrated in Figure 11.7.

For reliable delivery of the MCCH information, the RRC layer in the eNB repeatedly broadcasts identical content of the MCCH information every MCCH repetition period. The eNB can start transmitting updated MCCH information from the boundary of the MCCH modification periods, like system information on the BCCH.

When the MCCH information needs to be changed, the eNB notifies UEs about the change in MCCH information by indicating the M-RNTI on the PDCCH in one MCCH modification period before the updated MCCH information starts to be delivered. The PDCCH indicating the M-RNTI also informs UEs about one or more MBSFN areas for which the change in MCCH information is relevant, by using eight bits. Different bits correspond to different MBSFN areas. Each bit indicates the change in MCCH information for the corresponding MBSFN area. The eNB sets the bit to notify UEs about the start of a session or an ongoing counting procedure. The PDCCH indicating the M-RNTI is transmitted periodically based on the shortest MCCH modification period, and it is common to all MBSFN areas.

The MCE informs the eNBs about configurations of MCCHs such as the MCCH repetition period, the MCCH modification period, the MBSFN subframes allocated for MCCHs, and the modulation and coding scheme on the MBSFN subframes. The MCE can configure multiple MCCHs, one for each MBSFN area, in the eNB contributing to MBSFN transmissions, by using M2AP signaling procedures such as the M2 Setup procedure.

The RRC layer broadcasts configurations of MCCHs via system information. The SIB for MBMS services – SIB13 – is used to inform UEs about configurations of MCCHs. SIB13 is broadcast over the DL-SCH from a single cell, like normal SIB types. Thus, no MBSFN transmission is applied to SIB13. SIB13 also contains the configuration of notifications on the PDCCH.

In addition to SIB13, the RRC layer provides MCCH information based on M2AP signaling received from the MCE. The following downlink RRC messages can be used to give MCCH information:

- the *MBSFNAreaConfiguration* message;
- the *MBMSCountingRequest* message.

The *MBSFNAreaConfiguration* message contains most of the MCCH information except counting-related information. UEs need to receive the *MBSFNAreaConfiguration* message to receive one or more MBMS services from a single MBSFN area. Different *MBSFNArea-Configuration* messages are broadcast for different MBSFN areas.

The *MBMSCountingRequest* message used to initiate the MBMS Counting procedure contains counting-related information. Different *MBMSCountingRequest* messages should be broadcast for different MBSFN areas. It should be noted that the MBMS Counting procedure is supported from Release 10.

When the UE supports the MBMS, the UE will receive SIB13 from a cell belonging to an MBSFN area to configure an MCCH on an MBSFN area. The UE configures the MCCH on the MBSFN area when, for instance, the UE becomes interested in MBMS services from an MBSFN area, or a UE interested in MBMS services enters an MBSFN area.

Multiple neighboring cells may contribute to this MBSFN transmission with the same configuration of MBSFN subframes. Thus, while the UE is moving across cells within the MBSFN area, the UE can continue to receive the MCCH without reconfiguration, unless the eNB reconfigures the MCCH.

It should be noted that when the UE is capable of receiving MBMS services and is interested in MBMS services, the UE is allowed to receive MCCH information and MBMS services in both RRC_IDLE and RRC_CONNECTED.

11.5 MBMS Procedures

In general, networks provide an MBMS service to one or more UEs in accordance with the following sequence:

- When a session of the MBMS service starts, the eNB joins the MBMS service and then transmits MBMS data to one or more UEs.
- When the session starts, the UE receives MCCH information related to the MBMS service and then establishes a radio bearer related to the MBMS service called an MBMS Point-to-Multipoint Radio Bearer (MRB).
- While the eNB is transmitting the session to one or more UEs, the session may be updated, for example, with an update of the service area.
- When the session of the MBMS service stops, the eNB leaves the MBMS service and stops transmitting MBMS data to one or more UEs.
- When the session of the MBMS service stops, the UE receives updated MCCH information and then releases the MRB related to the MBMS service.

11.5.1 MBMS Session Start

For an MBMS service, the BM-SC initiates the MBMS Session Start procedure by sending a Session Start Request message, as illustrated in Figure 11.8. The MBMS Session Start procedure triggers one or more eNBs to join the concerned MBMS service and initiate transmission of MBMS user data to the UEs.

Before initiating the MBMS Session Start procedure, the BM-SC provides MBMS service announcements to UEs supporting MBMS, as specified in [2]. The MBMS service announcements cause UEs interested in MBMS services to be ready to receive MBMS services. Thus,

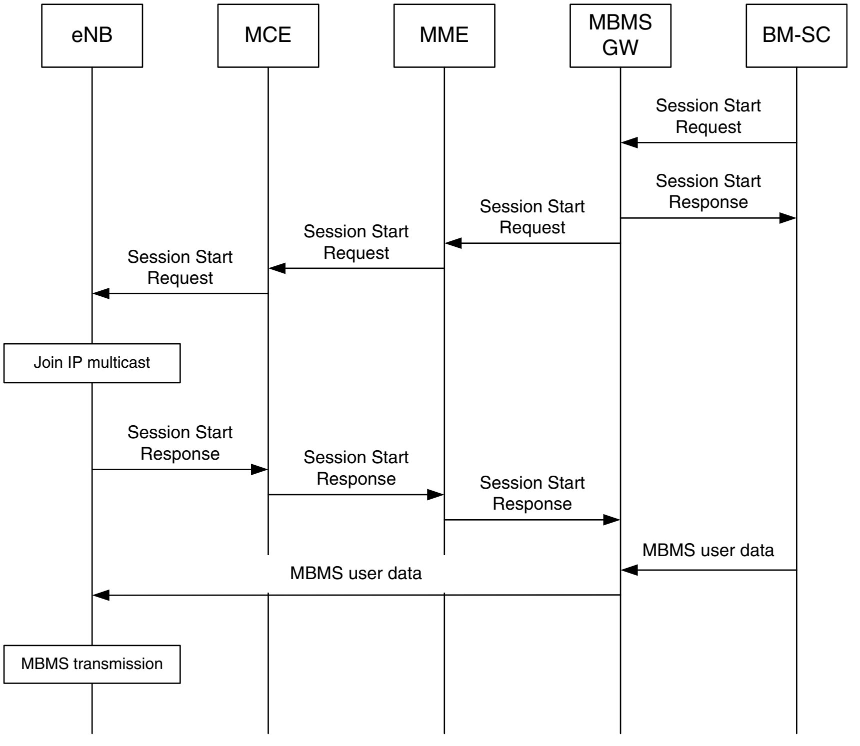

Figure 11.8 MBMS Session Start. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

when a session starts on the radio interface, interested UEs are ready to receive MBMS user data over the MTCH.

When initiating transmission of an MBMS session for the MBMS bearer service, the BM-SC assigns a Temporary Mobile Group Identity (TMGI) for identification of the MBMS bearer service and a session ID to distinguish each MBMS session. The BM-SC also determines session attributes such as QoS information and an MBMS service area where the MBMS session will be transmitted, and then includes them in the MBMS Session Start message.

When the session starts, the IP multicast address allocated by the MBMS Gateway is delivered to the appropriate eNBs via the Session Start Request message. After receiving the Session Start Request message, the eNBs should join the service to enable reception of MBMS user data from the MBMS Gateway. Then, the eNBs establish radio resources such as the configuration of MBSFN subframes in preparation for transmission of MBMS user data to UEs.

The BM-SC starts to send MBMS user data to the eNBs via the MBMS Gateway after the MBMS Session Start procedure. The BM-SC can apply forward error correction to MBMS user data. The MBMS Gateway uses the IP multicast protocol to forward the received MBMS user data to the eNBs that joined the service.

The BM-SC should make sure that the MCE and the eNB are ready for MBSFN transmission of this session by waiting for some time before sending MBMS user data. To be ready for MBSFN transmission, the eNBs need to configure the necessary radio resources and channels, and broadcast updated MCCH information reflecting this session, under the control of the MCE. After receiving MBMS user data from the MBMS Gateway, the eNBs broadcast MBMS user data on an MTCH configured for this session. Then, UEs can receive MBMS user data on the MTCH if they are interested in this session.

11.5.2 MCCH Information Acquisition and MRB Configuration

When a session of the MBMS service starts, MBSFN transmissions are coordinated across the eNBs under the control of the MCE. For this coordination, the MCE sends an MBMS Scheduling Information message to the eNBs, as illustrated in Figure 11.9. As specified

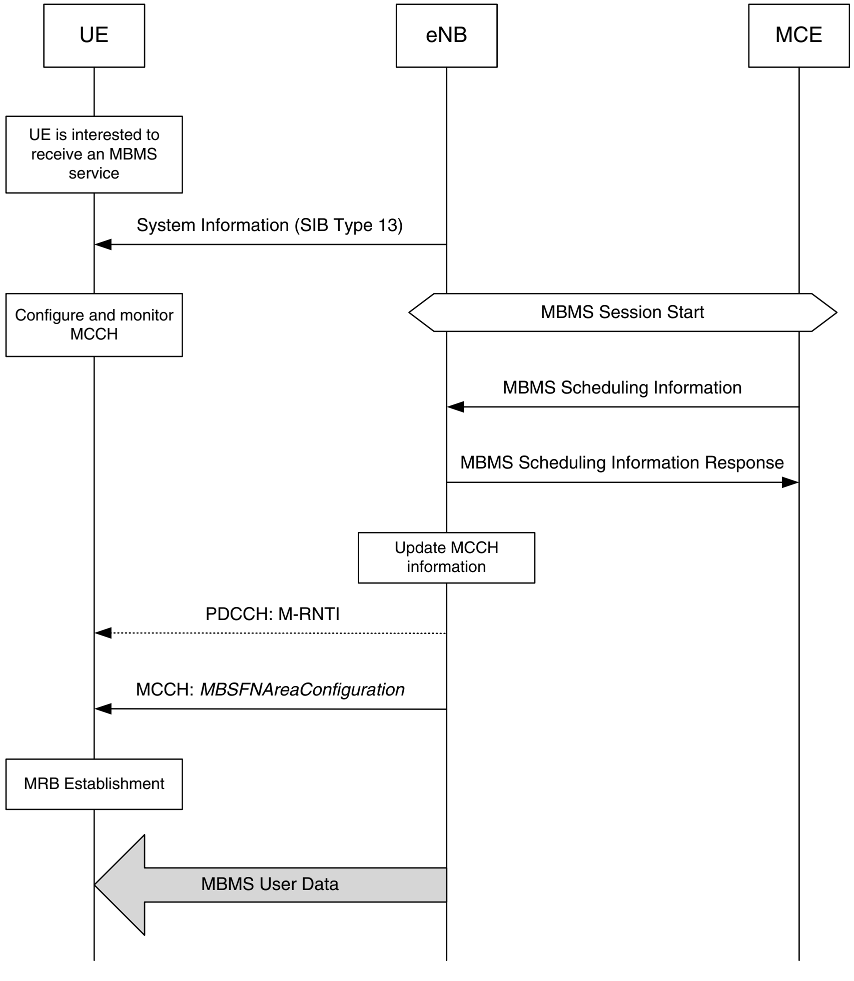

Figure 11.9 An example of MCCH information acquisition and MBMS user data reception. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

in [6], this message includes updated MCCH information such as configuration of an MRB corresponding to the session and the MCCH update time indicating when the eNB should apply the updated MCCH information.

When the MBMS Scheduling Information message is received for the session start, the eNBs indicate the M-RNTI on the PDCCH in order to pass information to UEs in an MCCH modification period. Then, the eNBs update the MCCH information broadcast on the MCCH in the next MCCH modification period, according to the MCCH update time indicated by the MCE.

If a UE supporting the MBMS is interested in the MBMS service, it keeps monitoring the PDCCH using the configuration contained in SIB13. When the PDCCH indicates a change in MCCH information in the MBSFN area concerning the MBMS service, the UE receives the MCCH to acquire the updated MCCH information for the MBSFN area.

Then, the UE establishes an MRB corresponding to the session with the updated MCCH information. When the MRB is established, the MTCH and the PMCH corresponding to the MRB are also configured.

After MRB establishment, the UE may receive MBMS user data from the MTCH. While the UE is moving across cells within the MBSFN area, the UE can continue to receive MBMS user data via the established MRB.

11.5.3 MBMS Session Update

While an MBMS session is ongoing, the BM-SC may modify session attributes such as the MBMS service area. For this modification, an MBMS Session Update Request message is sent to the MCE and the eNBs, as illustrated in Figure 11.10.

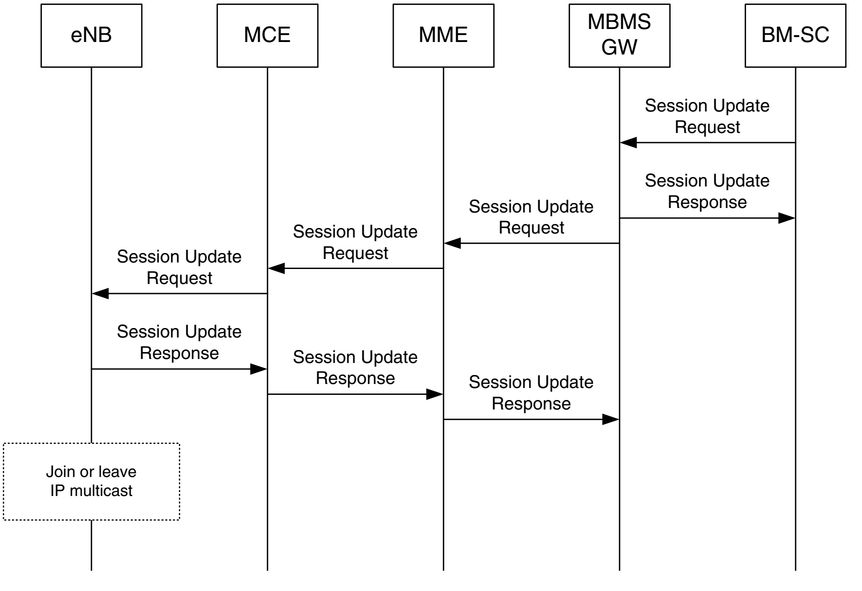

Figure 11.10 MBMS Session Update. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

When the Session Update Request message is received, an eNB that previously joined the MBMS service may leave the service, or a new eNB may join the service for the ongoing session, owing to modification of the MBMS service area for the ongoing session.

If an eNB leaves the service, the eNB releases the configurations of radio resources and channels. This will trigger UEs to release the concerned MRB. If a new eNB joins the service, the eNB applies the configurations of radio resources and establishes channels for this session. This will trigger UEs to establish the relevant MRB.

11.5.4 MBMS Session Stop

When no more MBMS user data are supposed to be transmitted for a particular session of the MBMS service, the BM-SC considers the session terminated. In this case, an MBMS Session Stop Request message is sent to the MCE and the eNBs, as illustrated in Figure 11.11.

When the Session Stop Request message is received, an eNB that has already joined for this session leaves the service in order to disable reception of MBMS user data from the MBMS Gateway. When the eNB leaves the service, it releases the configurations of radio

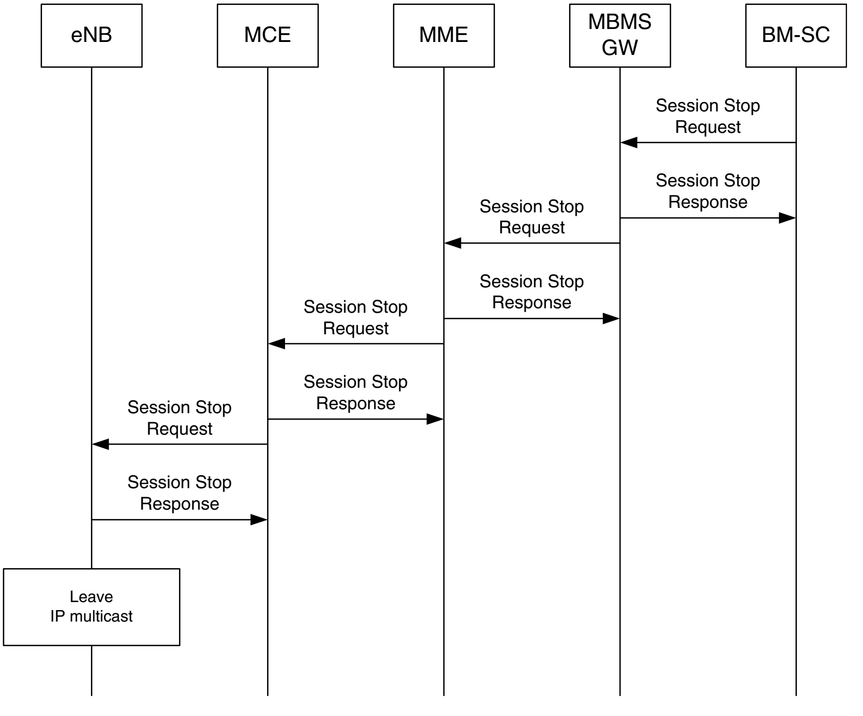

Figure 11.11 MBMS Session Stop. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

resources and channels for this session. This Session Stop triggers the UEs to release the concerned MRB.

11.5.5 MRB Release

When the session stops, the MCE sends an MBMS Scheduling Information message to the eNBs contributing to this MBSFN transmission, as illustrated in Figure 11.12. When the MBMS Scheduling Information message is received, the eNB updates the MCCH information for the concerned MBSFN area and stops MBSFN transmission for the MRB corresponding to the session.

While a session of the MBMS service is ongoing, the interested UE periodically receives the MCCH information from the MBSFN area concerning the MBMS service. The UE can recognize the Session Stop from the updated MCCH information. The UE considers that the session related to this MRB stops when there is no channel configuration for the established

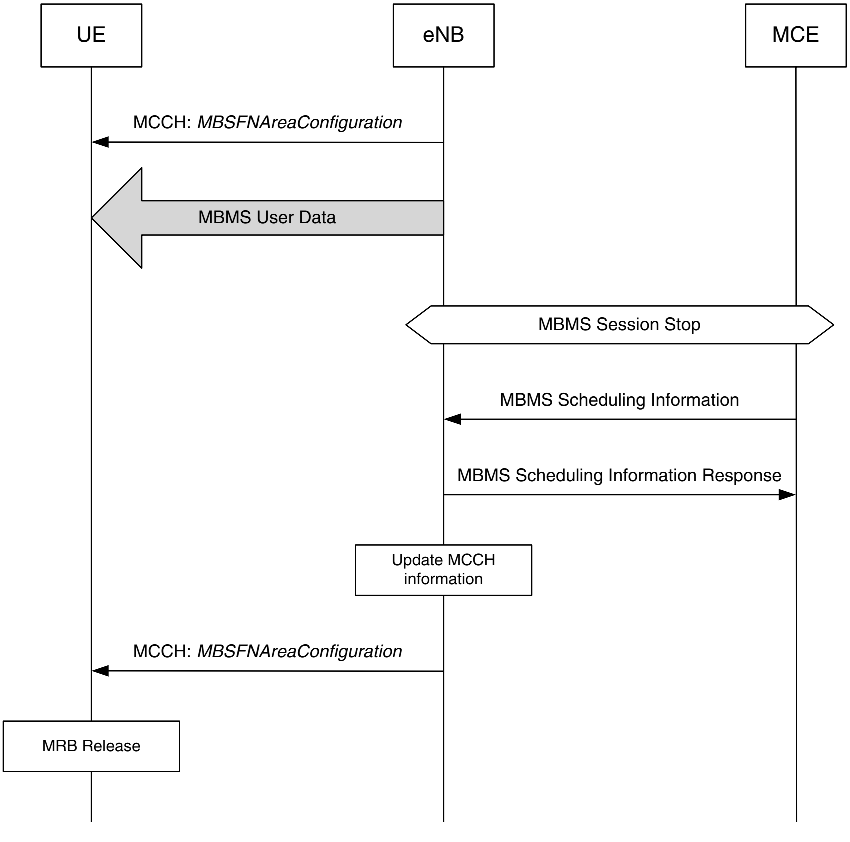

Figure 11.12 An example of MRB release. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

MRB in the MCCH information. When the session stops, the UE releases the established MRB for the session.

In addition to the session stop, the UE can also release the established MRB to stop receiving the session, for example:

- when the UE leaves the MBSFN area concerning the session;
- when the UE is no longer interested in the session because a user revokes interest;
- when the UE is not capable of receiving the session of the MBMS service any more, for example, due to reception of another MBMS/non-MBMS service that has a higher priority than the current MBMS service.

11.6 MBMS Enhancements in Releases 10 and 11

11.6.1 MBMS Counting

In Release 10, 3GPP enhanced the MBMS with the MBMS Counting procedure. The MBMS Counting procedure allows the MCE to count the number of UEs that are interested in an MBMS service within a whole MBSFN area. Depending on the results of the MBMS Counting procedure, the MCE may suspend an MBMS service or resume a suspended MBMS service in a whole MBSFN area.

In Release 9, the E-UTRAN does not know whether a particular UE is receiving MBMS services or not, even if the UE is in RRC_CONNECTED. In addition, the E-UTRAN does not know how many UEs are receiving a particular MBMS service. Thus, in a worst case scenario, the E-UTRAN may transmit an MBMS service in an MBSFN area with no UEs receiving the MBMS service for some time.

The MCE requests initiation of the MBMS Counting procedure by sending an MBMS Service Counting Request message to the eNB, as illustrated in Figure 11.13. This message includes an MBSFN area ID and a list of TMGIs corresponding to MBMS services for counting. This message also includes the MCCH update time, indicating when the MCCH information should be updated.

When the MBMS Service Counting Request message is received, the RRC layer of the eNB initiates the MBMS Counting procedure on the radio interface. The MBMS Counting procedure in the RRC layer consists of two RRC messages: the *MBMSCountingRequest* message in the downlink and the *MBMSCountingResponse* message in the uplink. The *MBMSCountingRequest* message is broadcast on an MCCH, and the *MBMSCountingResponse* message is transmitted on a DCCH via SRB1.

In the MBSFN area indicated by the MCE, the eNB broadcasts the *MBMSCountingRequest* message including the list of TMGIs in the MCCH modification period corresponding to the MCCH update time received from the MCE. The PDCCH indicating the M-RNTI will also be updated due to the presence of the *MBMSCountingRequest* message in the MCCH information in the MBSFN area. The MCE can signal an MBMS Service Counting Request message to several eNBs belonging to the same MBSFN area for the same MBMS service.

A UE that is interested in MBMS services or receiving MBMS services in the MBSFN area will receive the *MBMSCountingRequest* message from the eNB. If the MBMS service that the UE is interested in or is receiving corresponds to one of the listed TMGIs in the *MBMSCountingRequest* message, the UE responds to the *MBMSCountingRequest* message by sending an *MBMSCountingResponse* message. Since the *MBMSCountingRequest*

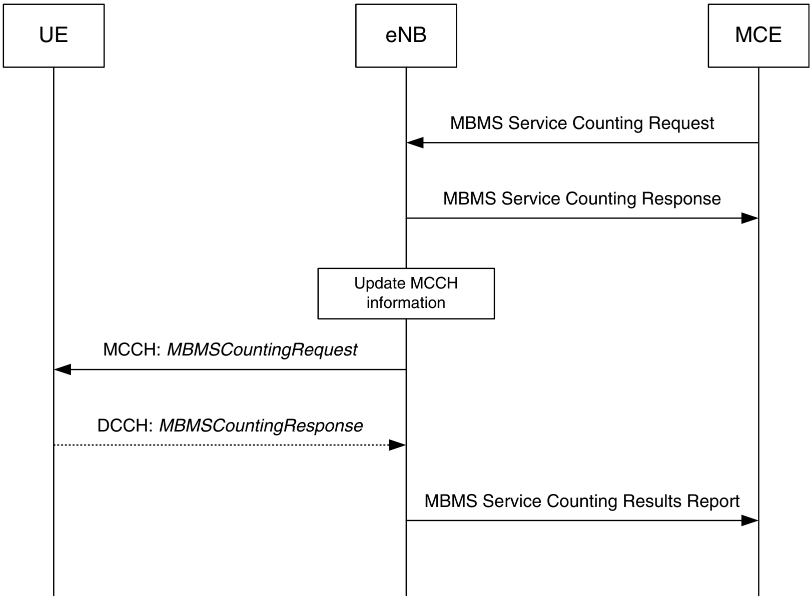

Figure 11.13 The MBMS Counting procedure. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

message is broadcast in an MCCH modification period, the eNB may receive multiple *MBMSCountingResponse* messages from multiple UEs responding to the same *MBMSCountingRequest* message.

The *MBMSCountingResponse* message does not include the TMGI itself, but indicates the index of the concerned TMGI among the list of TMGIs in the *MBMSCountingRequest* message for radio efficiency. Thus, when more than one *MBMSCountingRequest* message is transmitted for multiple MBSFN areas, the eNB may be confused by multiple *MBMSCountingResponse* messages from different UEs that may concern different MBSFN areas. Thus, when more than one MBSFN area overlaps at a cell, the UE should indicate the index of the concerned MBSFN area among the list of MBSFN areas broadcast in SIB13.

It should be noted that when the UE receives the *MBMSCountingRequest* message, the UE should be in RRC_CONNECTED to transmit the *MBMSCountingResponse* message via SRB1. Unlike MBMS counting in the UTRAN, UEs in RRC_IDLE do not perform the RRC Connection Establishment due to the MBMS Counting procedure in the E-UTRAN. Thus, the E-UTRAN will count only the number of RRC_CONNECTED UEs interested in a particular MBMS service by using the MBMS Counting procedure. However, the E-UTRAN may be able to estimate a ratio of the number of UEs in RRC_IDLE to the number of UEs in RRC_CONNECTED. Based on this ratio, the E-UTRAN may be able to estimate the number of UEs interested in the particular MBMS service for both RRC_IDLE and RRC_CONNECTED.

After sending the *MBMSCountingRequest* message, the eNB may receive one or more *MBMSCountingResponse* messages from one or more UEs in RRC_CONNECTED. The eNB collects the responses from the UEs and then reports the number of RRC_CONNECTED UEs that are receiving or are interested in the MBMS services via an MBMS Service Counting Results Report message to the MCE. The eNB includes the number of RRC_CONNECTED UEs for each listed TMGI in the message. The eNB can report this even in cases where no counting responses are received for a particular MBMS service.

The MCE may receive more than one MBMS Service Counting Results Report message from more than one eNB belonging to the MBSFN area. Based on the report, the MCE may decide to suspend transmission of the MBMS service in the MBSFN area, or resume transmission of a suspended MBMS service in the MBSFN area. In any case, the MCE can suspend or resume transmission of an MBMS service in the MBSFN area by sending an MBMS Scheduling Information message to the eNBs concerning the MBSFN area. It should be noted that resumption and suspension of sessions are no different from starting and stopping a session on the radio interface. In other words, UE behaviors for resumption and suspension are the same as those for Session Start and Session Stop.

11.6.2 MBMS Service Continuity

Until 3GPP starts to discuss mobility enhancements supporting MBMS service continuity for Release 11, it has been assumed that MBMS features do not affect mobility procedures in E-UTRA. Thus, some UEs that are receiving or interested in MBMS services may be unable to receive MBMS services due to cell reselection in RRC_IDLE or handover in RRC_CONNECTED.

For example, when a UE in RRC_IDLE is receiving an MBMS service at one carrier frequency, the UE in RRC_IDLE may reselect to another carrier frequency where no MBMS is supported, for example, due to cell reselection priorities. In addition, when a UE in RRC_CONNECTED is receiving an MBMS service at one cell, the eNB may move the UE to another cell where no MBMS is supported via the handover procedure, because the eNB does not know whether or not the UE is receiving MBMS services. Accordingly, some UEs may not continue to receive MBMS services.

It would be possible for UEs supporting MBMS to avoid such interruption of MBMS reception with dual receivers. For instance, if a UE had dual receivers, the UE could use one receiver to receive unicast services such as voice calls and an Internet service from a serving cell selected based on normal mobility procedures, and the other receiver to receive MBMS services from a cell supporting MBMS. However, one drawback of dual receivers is increased complexity in terminals supporting the MBMS. Thus, if possible, it is desirable for the UE to receive both unicast services and MBMS services on the same carrier frequency by using a single receiver.

In Release 11, 3GPP further enhanced the MBMS with support for MBMS service continuity. For this purpose, the network can provide assistance information to inform UEs about mapping information between carrier frequencies and MBMS services, and transmission timing of MBMS services.

By using the assistance information, when the UE is interested in a particular MBMS service, the UE in RRC_IDLE can autonomously set the carrier frequency carrying the MBMS service to the highest cell reselection priority for the scheduled time. As a result, it is likely that the UE will reselect to a cell on the carrier frequency carrying the MBMS service.

It should be noted that UEs in RRC_IDLE are allowed to set autonomously the highest cell reselection priority for MBMS not only from Release 11 but also from Release 9.

However, UEs in Releases 9 and 10 are not provided with the assistance information of Release 11. Hence, those UEs can do so only when they have knowledge of mapping information between carrier frequencies and MBMS services, for example, via application layers.

In Release 11, while the UE is in RRC_CONNECTED, the UE can inform the serving cell about carrier frequencies where MBMS services of interest are scheduled to be transmitted. For this purpose, the RRC layer introduces a new uplink message called the *MBMSInterestIndication* message.

When a UE is receiving or is interested in MBMS services, a UE in RRC_CONNECTED can send an *MBMSInterestIndication* message including one or more carrier frequencies where the MBMS services are scheduled. Then, when the eNB receives the *MBMSInterestIndication* message from the UE, it is likely that the eNB will move the UE to a cell on the carrier frequency carrying the MBMS service by initiating handover.

These MBMS enhancements in Release 11 will improve the MBMS user experience by reducing interruption of MBMS reception in both RRC_IDLE and RRC_CONNECTED.

References

1. 3GPP Technical Specification 23.246, "Multimedia Broadcast/Multicast Service (MBMS); Architecture and functional description (Release 10)", www.3gpp.org, accessed on 14 March 2012.
2. 3GPP Technical Specification 26.346, "Multimedia Broadcast/Multicast Service (MBMS); Protocols and codecs (Release 10)", www.3gpp.org, accessed on 14 March 2012.
3. 3GPP Technical Specification 25.446, "MBMS synchronisation protocol (SYNC) (Release 10)", www.3gpp.org, accessed on 14 March 2012.
4. 3GPP Technical Specification 36.211, "Physical Channels and Modulation (Release 10)", www.3gpp.org, accessed on 14 March 2012.
5. 3GPP Technical Specification 36.331, "Radio Resource Control (RRC); Protocol Specification (Release 10)", www.3gpp.org, accessed on 14 March 2012.
6. 3GPP Technical Specification 36.443, "M2 Application Protocol (M2AP) (Release 10)", www.3gpp.org, accessed on 14 March 2012.

12

Carrier Aggregation (CA)

At the World Radiocommunication Conference 2007 (WRC07), it was strongly argued that additional spectrum would be required for future mobile systems considering the explosive increase in mobile traffic expected over the next 15 years. The WRC07 defined several new bands for IMT-Advanced:

- 450 MHz band;
- UHF band (698–960 MHz);
- 2.3 GHz band;
- C-band (3400–4200 MHz).

In March 2008, the Radiocommunication sector of the International Telecommunications Union (ITU-R) commenced developing ITU-R Recommendations for the terrestrial components of the International Mobile Telecommunications-Advanced (IMT-Advanced) radio interface(s). The key features of IMT-Advanced in the recommendations are:

- a high degree of functional commonality worldwide while retaining the flexibility to support a wide range of services and applications in a cost-efficient manner;
- compatibility of services within IMT-Advanced and with fixed networks;
- capability of interworking with other radio access systems;
- high quality mobile services;
- user equipment suitable for worldwide use;
- user-friendly applications, services, and equipment;
- worldwide roaming capability; and
- enhanced peak data rates to support advanced services and applications (100 Mbps for high mobility and 1 Gbps for low mobility were set as research targets).

Given these features, the ITU-R invited proposals for candidate radio interface technologies for IMT-Advanced. In order to provide the 3GPP proposals for IMT-Advanced to the ITU-R, 3GPP defined LTE-Advanced in LTE Release 10 with the following general requirements:

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

Table 12.1 Performance requirements for IMT-Advanced and LTE-Advanced

	IMT-Advanced	LTE-Advanced
Peak data rate	100 Mbps for high mobility 1 Gbps for low mobility	1 Gbps in DL 500 Mbps in UL
Bandwidth	40 MHz	Up to 100 MHz
User plane latency	10 ms	Improved compared to LTE Releases 8/9
Control plane latency	100 ms	50 ms
Peak spectrum efficiency	15 bps/Hz in DL 6.75 bps/Hz in UL	30 bps/Hz in DL 15 bps/Hz in UL
VoIP capacity	Up to 200 UEs per 5 MHz	Improved compared to LTE Releases 8/9

- LTE-Advanced shall be an evolved version of LTE;
- LTE-Advanced shall meet or exceed IMT-Advanced requirements within the ITU-R time plane; and
- LTE-Advanced shall meet operator requirements.

Considering the above, 3GPP set performance requirements for LTE-Advanced in June 2008. Table 12.1 compares the requirements between IMT-Advanced and LTE-Advanced.

In addition to the performance requirements, LTE-Advanced is also required to meet spectrum requirements such as spectrum flexibility and spectrum compatibility. Spectrum flexibility means that LTE-Advanced is able to operate on non-contiguous spectrum, and spectrum compatibility means that LTE-Advanced is able to operate with LTE on the same spectrum.

In order to meet the spectrum requirements as well as the performance requirements, LTE-Advanced extends LTE Releases 8/9 with support for Carrier Aggregation (CA), where two or more Component Carriers (CCs) are aggregated to support wider transmission bandwidth up to 100 MHz, as shown in Figure 12.1. An LTE-Advanced UE can aggregate any of the component carriers in a non-contiguous manner depending on the capabilities.

Figures 12.2 and 12.3 show the Layer 2 structures for CA in the downlink and the uplink, respectively. Compared to the Layer 2 structure of Releases 8/9, a key difference to support CA is that one HARQ entity is used per CC. That is, a UE should be able to process multiple transport blocks in a TTI, one for each component carrier. Other than this, the Layer 2 protocols are kept Release 8/9 compliant. It is important to note that CA operation is invisible to

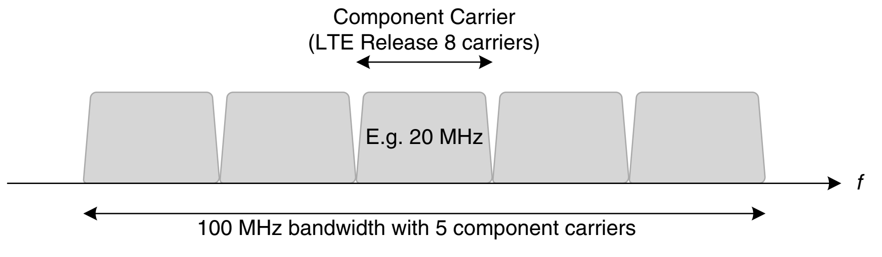

Figure 12.1 Concept of Carrier Aggregation

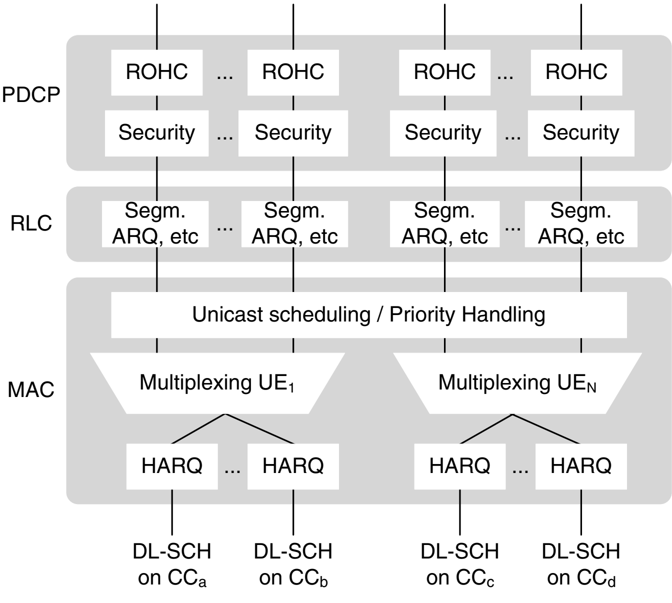

Figure 12.2 Layer 2 structure for DL with CA. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

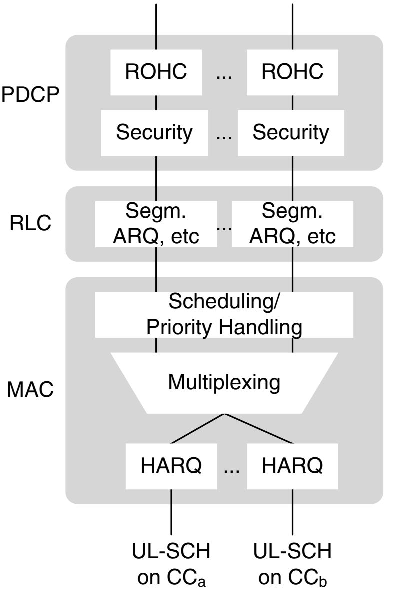

Figure 12.3 Layer 2 structure for UL with CA. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

the RLC and PDCP layers; that is, the RLC and PDCP layers are not impacted by the CA operation.

The impact of CA operation is restricted to some of the MAC procedures such as handling of multiple PDSCHs and PUSCHs. Each HARQ entity handles the data stream of each component carrier. Since the HARQ operation is performed for each component carrier, the associated control signaling, for example, PDCCH, PHICH, PCFICH, and SRS, is also defined for each component carrier. Multiplexing or de-multiplexing of multiple data streams is performed above the HARQ entity in the MAC layer.

12.1 Spectrum and Deployment Scenarios

12.1.1 Spectrum Scenarios

In CA, three different spectrum scenarios are considered, as illustrated in Figure 12.4.

Intra-band contiguous CA is used when multiple CCs belonging to the same band are allocated in a contiguous manner. Although this may be a less likely scenario considering today's frequency allocations, it will be likely to occur if new frequency bands like 3.5 GHz are allocated in the future.

Non-contiguous CA is used when multiple CCs belonging to the same band are allocated in a non-contiguous manner or when multiple CCs belonging to different bands are

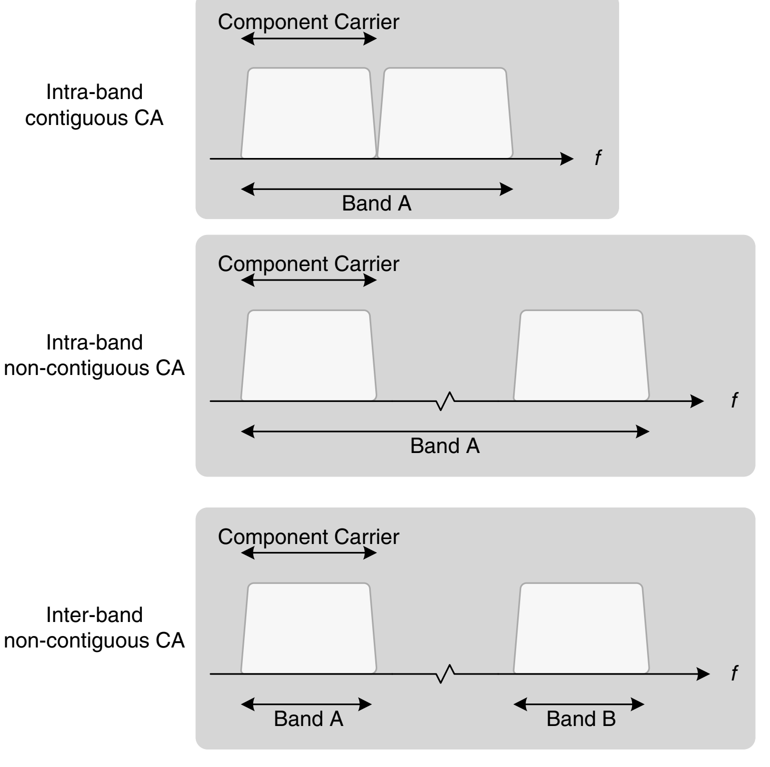

Figure 12.4 CA spectrum scenarios

allocated. As a result of the existing spectrum allocation policy, and because of the fact that the spectrum resource in the low frequency band (<4 GHz) is scarce, it is difficult to allocate contiguous 100 MHz bandwidth for a mobile network. Therefore, the non-contiguous CA technique provides a practical approach to enable an operator to utilize its current spectrum resources fully, including unused scattered frequency bands and the frequency bands already allocated for legacy mobiles. Non-contiguous CA can be classified further into intra-band non-contiguous CA and inter-band non-contiguous CA.

Intra-band non-contiguous CA is used when multiple CCs belonging to the same band are allocated in a non-contiguous manner. This scenario is expected in countries where a single band is allocated and the middle carriers are loaded with other users or used by shared networks.

Inter-band non-contiguous CA is used when multiple CCs belonging to different bands are allocated. With this type of aggregation, mobility robustness can potentially be improved by exploiting different radio propagation characteristics of different bands [1].

In Release 10, for the UL, 3GPP has decided to support only intra-band CA because inter-band CA requires additional functionalities such as maintenance of multiple uplink timing. On the other hand, for the DL, both intra-band and inter-band CA are supported.

12.1.2 Deployment Scenarios

CA deployment scenarios depend largely on the operator's needs. For efficient deployment, operators take into consideration various factors, such as spectrum allocation, urban/suburban/rural areas, antenna direction/location, hot spots, presence of frequency-selective repeaters, and so on. Some of the possible deployment scenarios are illustrated in Figure 12.5 [2]. It should be noted that CA is applicable only when multiple cells have overlapped coverage and belong to the same eNB.

Scenario 1 is considered the most typical scenario in Release 10. In this scenario, the antennas of the cells are collocated, and the cells are overlaid with little frequency separation. Therefore, the overlaid cells provide almost the same coverage. The simplicity of this scenario enables simple CA operation, which facilitates standardization work for CA.

Scenario 2 is where the antennas of the cells are collocated, but the cells are overlaid with large frequency separation, which leads to different cell coverage. Due to different pathloss, cell coverage with a higher frequency is less than that with a lower frequency. This scenario is also considered important as operators can, in practice, be allocated with spectrum across different bands.

Scenario 3 is where the antennas of cells are collocated, but the antenna directions are different in order to fill the coverage hole at the cell boundary. Though the antenna configuration is different, it is assumed that there are no new requirements compared to scenarios 1 and 2.

In scenario 4, the antennas of cells are not collocated because Remote Radio Head (RRH) cells are used in addition to normal eNB cells. As an RRH is a low-power node compared to the normal eNB, the cell coverage of RRH cells is typically less than that of eNB cells. Therefore, RRH cells are usually placed in hot spots to enhance throughput, thereby serving as a cost-efficient deployment option to the operator. Note that although the antennas of cells are not collocated, the CA technique can still be applied because eNB scheduling is performed in one place; that is, both eNB cells and RRH cells are under the control of the eNB. However, since the antenna location of the eNB cells is different from that of the RRH cells,

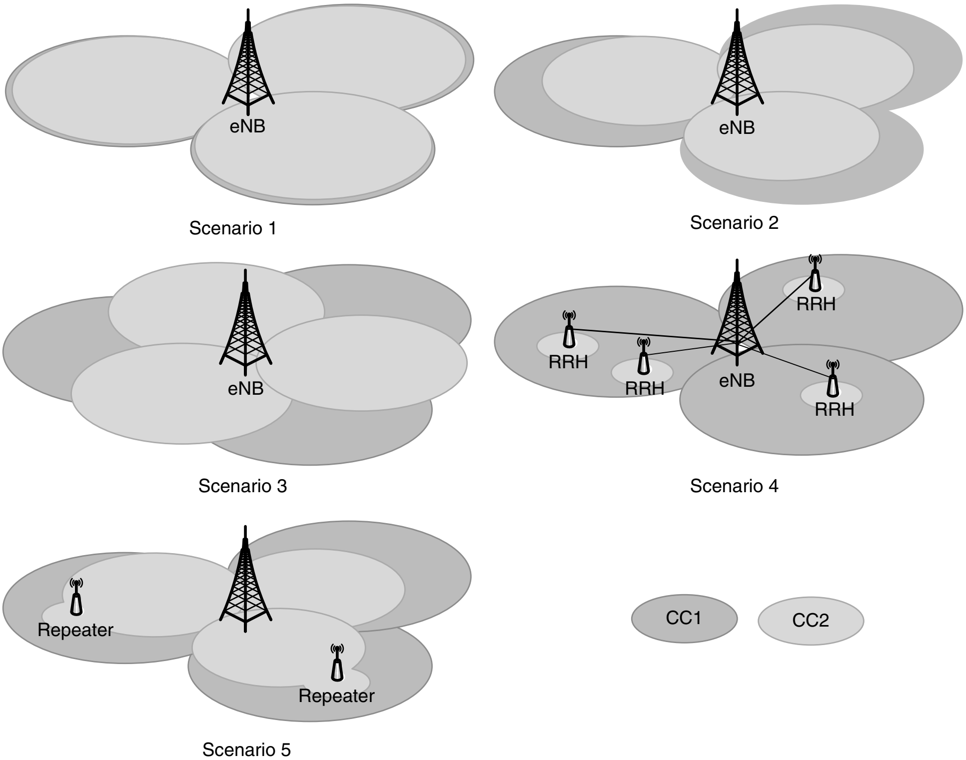

Figure 12.5 CA deployment scenarios. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

the propagation delays of eNB cells and RRH cells are also different, and thus the UE is required to maintain respective uplink timing for eNB cells and RRH cells.

Scenario 5 is where a part of the cells is amplified by frequency-selective repeaters. A repeater may be deployed by operators to extend cell coverage not only in rural areas but also in urban areas. For urban areas, a repeater is one of the deployment choices to ensure cell coverage indoors and underground, where penetration loss is large. When operators provide services over multiple bands, repeaters are not always deployed for all of the bands. The bands for which a repeater should be deployed are determined by the operator's deployment policy, taking into account link budget, deployment cost, and so on. Similar to scenario 4, different cells experience different propagation delays due to the presence of repeaters. Therefore, this scenario also requires the maintenance of multiple uplink timing.

As mentioned above, maintaining multiple uplink timing is necessary for scenarios 4 and 5. Since the complexity of this function is not trivial, 3GPP decided not to consider these scenarios in Release 10 for the uplink direction. This means that, in Release 10, the UE is required to maintain only single uplink timing. For the downlink direction, however, all scenarios from 1 to 5 are supported because downlink CA operation does not affect uplink timing. Support for multiple uplink timing is being considered seriously in Release 11.

12.2 Cell Management

12.2.1 PCell and SCell

The most outstanding feature in CA is that multiple serving cells are configured for the RRC_CONNECTED UE. The configured multiple serving cells are aggregated together to serve the UE. Here, an interesting question can be raised: are the functions of the aggregated serving cells equal?

In Releases 8/9, the Physical Cell Identity (PCI) of the serving cell is used to derive the security key (see Section 3.5). Only one PCI is used for key derivation. In CA, however, multiple PCIs are available, so one of the PCIs should be selected. Otherwise, the security structure needs to be changed.

In order to maintain the legacy security structure, 3GPP decided that one special serving cell would be used for security key derivation. This special serving cell is called the Primary Cell (PCell), and it plays a similar role to a Release 8/9 serving cell. The RRC_IDLE UE makes an RRC connection to a cell, from which the cell becomes the PCell of the UE. In RRC_CONNECTED, the UE operates on the PCell, performing all required procedures such as security, measurement, mobility, and so on. Without handover, the PCell never changes for the UE.

In CA, the eNB can configure additional serving cells to the RRC_CONNECTED UE depending on the amount of traffic. These additional serving cells are called Secondary Cells (SCells), and they provide additional radio resource to the UE. Since the configuration of SCells depends on the amount of traffic, an SCell can be configured only with a DL CC. This is the major difference from the PCell, where both DL CC and UL CC are always configured together.

Due to the asymmetric characteristic of CCs, the terminology "Carrier Aggregation" is not deemed correct, in that one PCell and one or more SCells are aggregated in this technology. Although the exact terminology is "Cell Aggregation", the familiarity of "Carrier Aggregation" makes this terminology be kept used. Throughout this section, "CC" is used when a DL CC or UL CC is mentioned separately, and "PCell" or "SCell" is used otherwise.

Since an SCell serves only as additional radio resource, most of the UE procedures are not applied to an SCell. For example, radio link monitoring, RA procedure, and Semi-Persistent Scheduling are not applied. Only data transfer related procedures are performed for an SCell.

As per CA in the Release 10 specification, the following functions are performed only through the PCell:

- **Random Access (RA) procedure:** The two purposes of the RA procedure are requesting uplink radio resource and aligning uplink transmission timing. Since the same transmission buffer and the same transmission timing are used for all UL CCs, there is no need to perform the RA procedure on an SCell.
- **Radio Link Monitoring (RLM):** RLM for an SCell is not needed because it can be done through the CQI/SRS report for the SCell. As a result, RLM is performed only for the PCell, and a Radio Link Failure (RLF) is declared only when the channel quality of the PCell is poor.

- **Handover procedure:** While the handover procedure is always required for a change of PCell, a change of SCell can be performed without the handover procedure.
- **PUCCH transmission:** The PUCCH transmission carrying the HARQ ACK/NACK, the CSI (CQI/PMI/RI/PTI) reports, and the Scheduling Request (SR) is performed only on the PCell.

12.2.2 Signaling of Configuration Information

When the UE completes the RRC Connection Establishment procedure successfully, only one serving cell is configured, which is the PCell. After that, if there is a need to add, modify, or remove SCells for the UE, this can be achieved through the RRC Connection Reconfiguration procedure. While an *RRCConnectionReconfiguration* message including *sCellToAddModList* information is used for the addition and modification of an SCell, an *RRCConnectionReconfiguration* message including *sCellToReleaseList* information is used for the removal of an SCell. An *RRCConnectionReconfiguration* message for SCell management may include the following parameters:

- *cellIdentification;*
- *sCellIndex;*
- *radioResourceConfigCommonSCell;*
- *radioResourceConfigDedicatedSCell.*

The *cellIdentification* contains the PCI of the SCell and the Absolute Radio Frequency Channel Number (ARFCN) of the SCell.

The *sCellIndex* is used to identify each SCell within the configured SCells for the UE. In the PHY layer, when cross-carrier scheduling is used, it indicates the target SCell to which the resource allocation information over the PDCCH applies. In the MAC layer, it indicates the SCell's corresponding bit position in the bitmap of the Extended Power Headroom MAC CE and the SCell Activation and Deactivation MAC CE. In the RRC layer, it identifies the SCell to be removed or modified. Because this 3-bit parameter is shorter than the 9-bit PCI, it can reduce the amount of signaling overhead.

The *radioResourceConfigCommonSCell* contains essential information for the UE to operate properly on the SCell. The essential information is the common information that applies to all UEs operating on the SCell, and includes parameters such as the PHY layer configuration parameters. Because the essential information corresponds to system information, the UE is not required to read the system information directly from the SCell. Furthermore, when the system information of the SCell changes, the updated information is also delivered to the UE through the *RRCConnectionReconfiguration* message. As a result, while the UE should acquire the essential information of the PCell directly by receiving the system information of the PCell, the UE can skip reading the system information of the SCell.

The *radioResourceConfigDedicatedSCell* contains UE-specific configuration information for operation on the SCell.

12.2.3 Linkages and References

The *ul-CarrierFreq* parameter included in SIB2 indicates the UL CC associated with the DL CC. Because this mapping information is delivered through SIB2, this association between the DL CC and the UL CC is called the *SIB2 linkage*. When CA is not used for the UE, there is only one DL CC and only one UL CC. Therefore, in this case, the SIB2 linkage is the only possible linkage between the UL CC and the DL CC.

On the other hand, when CA is used for each UE, there are multiple UL CCs and DL CCs. Furthermore, the configured number of DL CCs can be different from the configured number of UL CCs. Therefore, in addition to the SIB2 linkage option, there are multiple choices for a possible linkage between the configured DL CCs and the configured UL CCs. As a result, a UE-specific linkage can be configured such that different DL CCs for the same UL CC can be configured for different UEs and vice versa.

However, only the SIB2 linkage is used for the following two reasons.

First of all, a UE-specific linkage leads to inefficient use of the radio resource. During the contention-based RA procedure, when the RA preamble transmission is detected on the specific UL CC, the eNB has to transmit the RA response. For the SIB2 linkage, because there is a unique DL CC associated with the UL CC, the eNB knows on which DL CC the RA response has to be transmitted. However, if a UE-specific linkage is used for each UE, there may be multiple DL CCs associated with the UL CC. In this case, because the eNB does not know which UE has made the RA preamble transmission on the UL CC, the eNB has to transmit the RA response on all possible DL CCs. This leads to a waste of radio resources.

Secondly, with a UE-specific linkage, the RA procedure gets more complex without any benefit. During the non-contention-based RA procedure, when the eNB detects the RA preamble transmission on a specific UL CC, it can uniquely identify the UE that has made the RA preamble transmission. After that, based on the linkage configuration of the UE, the eNB can transmit the RA response on the DL CC associated with the UL CC. However, this means that the UE has to monitor a different DL CC than the SIB2-linked DL CC, depending on whether the performed RA procedure is non-contention-based or contention-based. Thus, a UE-specific linkage incurs unnecessary complexity.

Because the UE adjusts its UL transmission power based on the DL pathloss estimation, the UE should decide which DL CC should be used in estimating the DL pathloss for the transmission on the specific UL CC. As a simple rule, for each UL CC, the DL CC linked through the SIB2 linkage is used for this pathloss estimation. However, to support HetNet deployment where the pathloss estimation for some SCells is not reliable, use of the PCell as a reference for the pathloss estimation instead of the DL CC linked by the SIB2 linkage is allowed.

The uplink transmission timing of the UE is adjusted in reference to the downlink reception timing. In CA, the PCell DL CC is used as the downlink timing reference for all UL CCs. The reasoning behind this is that only intra-band CA is supported in Release 10. In the intra-band CA scenario, the UL propagation delay is almost the same for all UL CCs and the DL propagation delay is almost the same for all DL CCs. As a result, timing management can be undertaken using a single UL CC/DL CC pair. Because time alignment through the RA procedure is always supported on the PCell, the PCell is chosen as the reference for the timing.

12.2.4 Cross-Carrier Scheduling

When CA is not used, the UE is configured with only one DL CC and only one UL CC. As a result, the resource allocation information through the PDCCH of the DL CC indicates the allocation of radio resources for the DL CC and the associated UL CC. This is shown on the left side of Figure 12.6.

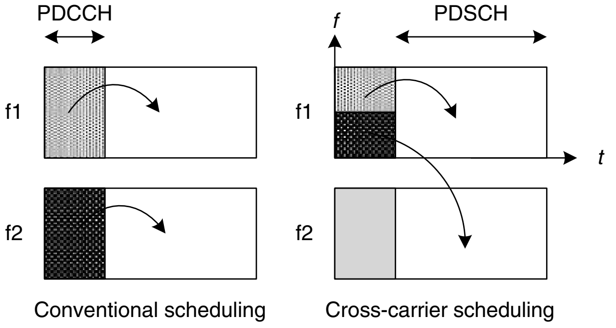

Figure 12.6 Conventional scheduling and cross-carrier scheduling. Reproduced by permission of 3GPP, © 2010. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

However, when CA is used for the UE, there are multiple UL CCs and DL CCs that the eNB can choose for the radio resource allocation. Based on this fact, cross-carrier scheduling is introduced for CA, as illustrated on the right side of Figure 12.6. This cross-carrier scheduling allows the resource allocation information through the PDCCH of one DL CC to indicate the radio resources on the other CCs. A new field called the Carrier Indicator Field (CIF) delivered through the PDCCH indicates the target CC where the allocated radio resource exists.

There are three benefits to cross-carrier scheduling.

Firstly, cross-carrier scheduling is effective in combating an adverse environment such as a Heterogeneous Network (HetNet), where neighboring cell interference is significant. In this case, if the PDCCH signals are not reliable in an interfered cell, radio resource allocation for the interfered cell may not be possible. This problem can be resolved by cross-carrier scheduling because the eNB can configure cells other than the interfered cell for PDCCH transmission.

Secondly, cross-carrier scheduling improves the scheduling flexibility of the eNB. Without cross-carrier scheduling, the resource allocation information can be delivered through only one DL CC. However, with cross-carrier scheduling, the eNB has multiple DL CC options for the transmission of the resource allocation information. For example, when a DL CC with enough DL-SCH resource experiences a shortage of PDCCH resource, the PDCCH resource of other DL CCs can be used to transmit the resource allocation information for the DL CC.

Lastly, cross-carrier scheduling reduces the burden of the UE's PDCCH blind decoding. When cross-carrier scheduling is used, the number of monitored DL CCs can be made smaller than the total number of configured DL CCs. Because PDCCH blind decoding requires a lot of UE processing power, a reduction in the amount of PDCCH blind decoding directly relieves the requirement on the UE processing performance.

The PCell is the cell where all elementary procedures for connection management are performed. Thus, cross-carrier scheduling is not used for the PCell; that is, the PDCCH of an SCell cannot transport the resource allocation information for the PCell.

For each cell, the configuration of the cross-carrier scheduling is done by using the parameter *CrossCarrierSchedulingConfig* included in the *RRCConnectionReconfiguration* message. This parameter first indicates whether another cell is transmitting the resource allocation information for this cell. If so, this parameter further indicates which cell is transmitting the resource allocation information for this cell. If not, this parameter further indicates whether the PDCCH of this cell can deliver the resource allocation information for another cell.

12.2.5 Extended Measurements

The measurement framework in Releases 8/9 is built upon the assumption that a UE supports only a single serving cell. In Release 10, the measurement framework has been extended such that it is applicable to UEs supporting multiple serving cells.

12.2.5.1 Extended Definition of Measurement Types

The UE needs to perform three types of measurement: (1) serving cell measurements, (2) intra-frequency neighboring cell measurements, (3) inter-frequency neighboring cell measurements. Each measurement type is extended for a UE configured with multiple serving cells, as follows:

- **Serving cell measurements:** This refers to the measurements of all configured serving cells. As examples of serving cell measurements (see Figure 12.7), the UE needs to perform measurements of the PCell and SCell.
- **Intra-frequency neighboring cell measurements:** This refers to measurements of neighboring cell(s) on the same frequency as the serving cells configured to the UE. The UE can perform intra-frequency neighboring cell measurements without using any measurement gap. As an example (see Figure 12.7), all the cells on frequencies 1 and 2 except for the PCell and SCell are subject to intra-frequency neighboring cell measurements.

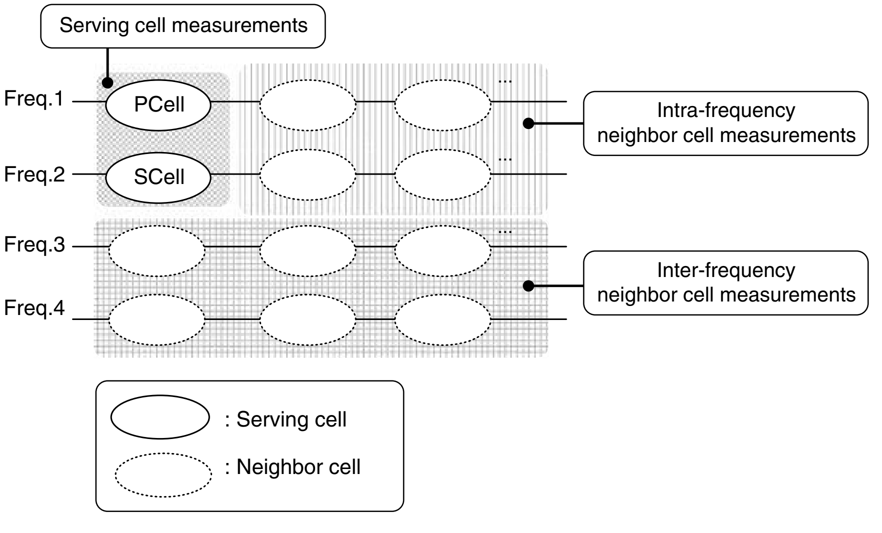

Figure 12.7 Extended definitions of each measurement type

- **Inter-frequency neighboring cell measurements:** This refers to the measurements of neighboring cell(s) on a frequency different from any frequency of serving cells configured to the UE. The UE may need measurement gaps, depending on the UE capability, to perform inter-frequency neighboring cell measurements. As an example (see Figure 12.7), all the cells on frequencies 3 and 4 are subject to inter-frequency neighboring cell measurements.

12.2.5.2 Generalization of Existing Measurements

The measurement performance requirements are different depending on whether the measured frequency is a primary or a secondary frequency and whether the SCell on the measured frequency, if configured, is activated or deactivated. The details of measurement performance requirements can be found in [3].

In Releases 8/9, the UE is required to perform serving cell measurements. This requirement is extended in Release 10 such that the UE needs to perform measurements of all configured serving cells. In Releases 8/9, the measured results of a serving cell are always included in all *MeasurementReport* messages, in order to assist the serving cell to perform the proper mobility procedure. This behavior is also extended in Release 10 such that the UE has to include the measured results of all serving cells configured to the UE; that is, the available measurement results of the PCell and SCells are always included in all *MeasurementReport* messages.

Several measurement reporting events are defined for Release 8/9 UEs. The extension of these events for Release 10 is made simply with the clarification that the "serving cell" in the existing events in Releases 8/9 refers to the PCell in Release 10. The affected events are event A1, A2, A3, and A5 for E-UTRAN measurements, and B2 for inter-RAT measurements. Other events such as A4 or B1 are not affected by this clarification, as these events are irrelevant to serving cell measurements.

The UE measurement behaviors for energy-efficient measurement are often referred to as the "s-measure," and the s-measure mechanism in Release 10 is, in principle, the same as that defined in Releases 8/9. The s-measure mechanism in Release 8 means that the UE is allowed to omit neighboring cell measurements if the measured RSRP of its serving cell is higher than the s-measure threshold. The only clarification made for the Release 10 s-measure is that the serving cell in the s-measure-related UE behaviors refers to the PCell. That is, if the measured RSRP of the PCell is higher than the signaled *s-Measure* parameter, the UE can skip neighboring cell measurements (intra-/inter-frequency neighboring cell measurements). On the other hand, if the measured RSRP of the PCell is not higher than the *s-Measure* parameter or if the *s-Measure* parameter is not signaled, the UE needs to perform all the required neighboring cell measurements.

3GPP discussed the need for s-measure enhancement to facilitate the addition or replacement of SCells even in cases where the RSRP of the PCell is sufficiently good. Since the existing s-measure mechanism allows the UE to skip neighboring cell measurements in such cases, the UE may not detect a better SCell candidate. However, the conclusion of the discussion was that s-measure enhancement is not essential as the employment of existing mechanisms such as DRX, SCell deactivation, and careful measurement configuration can avoid significant battery consumption for the UE configured with multiple serving cells.

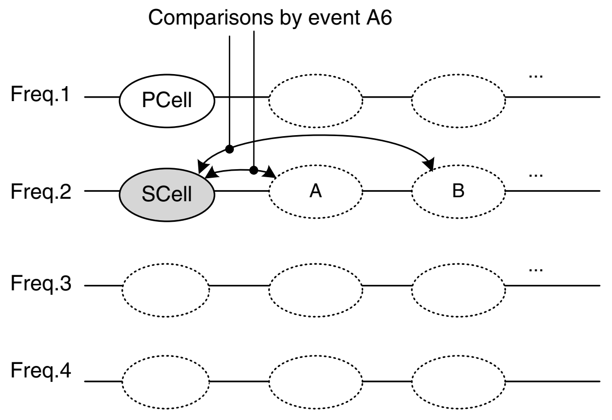

Figure 12.8 Measurement reporting evaluation with event A6

12.2.5.3 New Measurement Reporting Event (A6)

A new event, A6, is introduced in Release 10 to support the detection and reporting of a better cell or cells than the SCell on the same frequency. Figure 12.8 shows an example of measurement evaluations enabled by event A6 configured on the secondary serving frequency, where the SCell is compared with neighboring cell A and neighboring cell B on the same frequency.

The motivation behind new event A6 is to assist replacement of an SCell with a better cell, if available, on the concerned frequency. In general, the efficiency of spectrum usage is maximized if the eNB configures the UE with the best cell, in terms of radio condition, as a serving cell on the concerned frequency. The configuration of the best cell as a serving cell on the frequency becomes possible if the UE can detect and report a better cell or cells including the best cell on the concerned frequency. For the detection and reporting of better cells than the PCell on the primary frequency, the existing event A3 can be used, where the measured result of the PCell is evaluated against the measured results of the neighboring cells on the primary frequency. However, there are no existing reporting events applicable to detection and reporting of better cells than the SCell on the concerned secondary frequency. The new event A6 works for this purpose.

The introduction of event A6 requires some additional handling in terms of the measurement identity swapping that was introduced in Release 8. Measurement identity swapping provides continuation of previously configured measurements on the serving frequency after handover or re-establishment. Figure 12.9 shows an example where additional handling is required against an invalid configuration introduced by measurement identity swapping.

In Figure 12.9, there is an invalid association between the event A6 and the MeasObj 1 via MeasId 2 after measurement identity swapping, because the event A6 is only applied to the secondary frequency. Such an invalid configuration takes place if no serving cell remains on the source primary frequency as a result of handover/re-establishment to a cell on a frequency other than the source primary frequency. To avoid the invalid configuration remaining in the UE, the UE needs autonomously to release the invalid measurement identity for

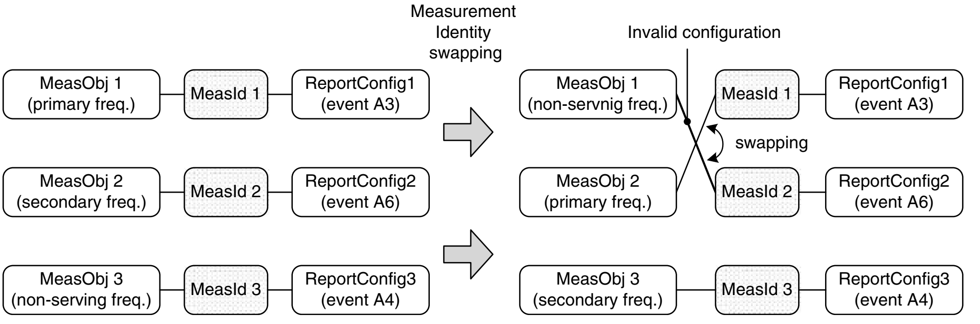

Figure 12.9 Invalid measurement configuration after measurement identity swapping

which a measurement event becomes linked inappropriately to a measurement object. In the example of Figure 12.9, the UE has to release MeasId 2.

12.2.6 SCell Management

To utilize the available spectrum resources efficiently, a network needs to manage the configuration/reconfiguration of SCells towards UEs carefully. The management of SCells includes SCell addition/removal/replacement, and SCell activation/deactivation. The measurement reporting events can be used to assist in the management of SCells, as follows:

- **SCell Addition:** If event A4 is configured towards the non-serving frequency, it can trigger a measurement report when a new SCell candidate is detected on that frequency.
- **SCell Removal:** If event A2 is configured towards the configured serving frequency, it can trigger a measurement report when the measured level of the SCell on the frequency becomes severely degraded and thus may be better deactivated or even removed from the configured SCell(s).
- **SCell Activation:** If event A1 is configured towards the configured serving frequency, it can trigger a measurement report when the measured level of the SCell on the frequency becomes good enough to use.
- **SCell Deactivation:** Event A2 can be used in a similar manner to SCell Removal.
- **SCell Replacement:** For the replacement of an SCell with another cell (new SCell candidate) on the same frequency, event A6 can be used to detect the SCell candidates with measured levels higher than the current SCell and to trigger a measurement report.

Note that the SCell management described above is just an example, and it is up to the operational policy of the network.

12.2.7 Mobility with Carrier Aggregation

For a handover, the general principle is that the radio configuration to be used at the target cell is decided by the target cell and then signaled to the UE via a handover command in the source cell. This principle also applies for carrier aggregation, and the handover command can include the configuration of the SCells to be used at the target cell.

Upon receiving the handover command including the configuration of the SCell(s), the UE configures the SCell(s) according to the received configuration but deactivates the SCell(s) during the handover. The UE keeps the SCell(s) deactivated until the UE receives an SCell activation command from the target cell after the completion of the handover.

The selection of the PCell to be used at the target cell in Release10 is no different from the selection of a serving cell to be used at the target cell in Releases 8/9. The selection of the SCell(s) for use in the target cell, however, did not exist in Releases 8/9. To assist in the selection of the SCell(s) at the target cell, additional signaling has been introduced to provide information from the UE to the target cell via the source cell.

To make this assistance information available at the source cell, the reporting of the best neighboring cell on each serving frequency has been introduced. For the reporting of the best neighboring cell, the serving cell provides the UE with a measurement reporting configuration including the indication, *reportAddNeighMeas.* If the measurement reporting condition is met, the UE indicates, in the associated measurement report, the best neighboring cell of each serving frequency for all configured serving frequencies. The measured values of RSRP and RSRQ of the best cell are also included in the measurement report.

Based on the measurement results received from the UE, the source cell constructs a list of SCell candidates. An entry on the list consists of a physical cell identifier, the carrier frequency, and RSRP and RSRQ values. It is assumed that the source cell makes sensible decisions in constructing the list. The source cell then forwards the list of SCell candidates to the target cell during handover preparation. The target cell utilizes the received information on the SCell candidates for configuration of SCells to be used by the UE after handover.

12.3 Extended MAC Functions

12.3.1 SCell Activation and Deactivation

As already explained, SCells are added and removed by means of RRC signaling. However, in order to utilize a DL CC for PDCCH/PDSCH reception and a UL CC for PUSCH transmission, the SCell needs not only to be added but also activated. This means that there is a separate activation step to make the SCell usable. The motivation behind having a separate activation step is to reduce signaling overhead incurred by the frequent addition and removal of SCells. Here, it should be noted that activation and deactivation are applied only to SCells. The PCell is never deactivated in order to maintain a reliable connection between the UE and the eNB.

The activation and deactivation of an SCell are realized by MAC signaling rather than RRC signaling. By using MAC signaling, the eNB can quickly change the activation/deactivation status of an SCell in accordance with data activity, which improves UE battery consumption with less overhead. Given that the deactivated SCell state has been introduced for power-saving purposes, the activity of RF and baseband parts of the UE should be set to a minimum in this state. Therefore, it should be enough for the UE to perform only RSRP/RSRQ measurements for a deactivated SCell, without performing CSI measurements that consume considerable amounts of UE power. Power saving can be realized by not performing CSI measurements for deactivated SCells.

Not only is measurement on the DL CC disabled, but transmission on the UL CC is also disabled for a deactivated SCell. Since the eNB does not schedule PUSCH on a deactivated SCell, it is useless to transmit SRS on it. Rather, SRS transmission on a deactivated SCell

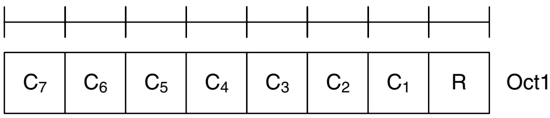

Figure 12.10 Activation/Deactivation MAC CE format. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

causes inter-cell interference, wasting UE battery. Moreover, since pathloss estimation from the DL CC of a deactivated SCell is not reliable, power control on the UL CC would not work accurately. For this reason, activation or deactivation of an SCell forces both DL CC and UL CC to be enabled or disabled at the same time.

As mentioned above, MAC signaling is employed for activation and deactivation of an SCell. A new MAC CE called the Activation/Deactivation MAC CE has been introduced for this purpose. This MAC CE is identified by the MAC PDU subheader with LCID. It has a fixed size and consists of a single octet containing seven C_i-fields and one R-field, as illustrated in Figure 12.10.

The C_i field indicates the activation/deactivation status of the SCell with *SCellIndex i*. That is, the C_i field "1" indicates that the SCell with *SCellIndex i* shall be activated, and the C_i field "0" indicates that the SCell with *SCellIndex i* shall be deactivated. The R field is a reserved bit. Since the bitmap is included in the MAC CE format, multiple SCells can be activated and/or deactivated by a single MAC CE.

For deactivation, a second method can be used in addition to the explicit deactivation command by the Activation/Deactivation MAC CE – the SCell can be deactivated implicitly by a timer called *sCellDeactivationTimer*. The UE relies on the deactivation timer if the explicit deactivation command is lost or is not provided by the eNB.

The *sCellDeactivationTimer* is configured for each SCell, though the value is common for all SCells. The timer is started when the explicit command activating the SCell is received, and restarted when a PDCCH indicating an uplink grant or downlink assignment is received on the SCell. Upon the expiry of the timer, the UE deactivates the SCell. Therefore, the SCell is considered activated only when its *sCellDeactivationTimer* is running.

When the eNB activates the SCell, it cannot immediately use the SCell for data transmission. This is because the UE needs some time to decode the activation command and activate the receiver chain. The required time period is called the *activation period*, and its length has been chosen to be 8 ms to give a reasonable tradeoff between performance and complexity. Due to the activation period, there is a timing difference between the transmission of an activation command and actual use of the SCell. For example, when the UE receives an Activation/Deactivation MAC CE activating the SCell in subframe n, normal SCell operation starts in subframe $n + 8$.

The same principle applies for deactivation. To provide the UE with enough time to decode the deactivation command and deactivate the receiver chain, an 8 ms deactivation period is also defined. However, the UE is allowed to deactivate the SCell earlier than 8 ms in order to save UE battery. Therefore, when the UE receives an Activation/Deactivation

MAC CE deactivating the SCell in subframe n, normal SCell operation is finished in a subframe no later than $n+8$. To maintain consistent UE behavior, the same deactivation period is applied for SCell deactivation by the *sCellDeactivationTimer*.

Though early deactivation is allowed, CSI reporting should continue up to subframe $n+8$. The reason is as follows. In the case where CSI reporting and PUSCH transmission coincide in the same subframe, the CSI is multiplexed in PUSCH. In this case, the length of the CSI varies depending on the number of activated SCells. If the CSI reporting is allowed to stop before subframe $n+8$, the eNB cannot predict the exact timing of CSI reporting termination, and thus the eNB has to perform blind decoding of PUSCH during the whole deactivation period. To remove this unnecessary eNB complexity, CSI reporting is forced to continue for the whole deactivation period even though other SCell operations can stop in a subframe earlier than $n+8$.

When an SCell is activated, normal SCell operation is performed. Normal SCell operation includes the following behaviors:

- SRS transmissions on the SCell;
- CSI reporting for the SCell;
- PDCCH monitoring on the SCell;
- PDCCH monitoring for the SCell on another cell (this is applicable only when cross-carrier scheduling is configured for the SCell).

12.3.2 Power Headroom Reporting (PHR)

The Power Headroom (PH) indicates the difference between the nominal maximum output power and the estimated PUSCH transmit power. This information is useful for the eNB scheduler to perform uplink scheduling for the UE with optimal transmission power. In Releases 8/9, a single PH per UE is sufficient because the UE has only a single UL CC. However, in CA where multiple UL CCs are used, each CC may experience different channel conditions, and thus the PUSCH transmit power may be different among UL CCs. Therefore, the PH should be reported per CC, and the power control should also be performed per CC. This is the reason why per-CC PH is introduced in CA. The per-CC PH indicates the difference between the maximum CC output power and the estimated PUSCH transmit power on the CC.

Moreover, in Release 10, where CA is introduced, the MAC layer is enhanced so that the UE is configured to transmit PUSCH and PUCCH in the same subframe. The support for simultaneous PUSCH and PUCCH transmission depends on UE capability. Then, the transmit power in the UE needs to be distributed between two channels. This means that, even if a single UL CC is configured, the eNB cannot know the PH of the CC unless the estimated PUCCH transmit power is reported. Therefore, another type of PH is introduced to indicate the difference between the maximum CC output power and the estimated PUSCH + PUCCH transmit power on the CC. This type of PH is only applicable for the PCell because PUCCH transmission is allowed only on the PCell.

Consequently, two types of per-CC PH are defined in CA – one considers only PUSCH transmit power, and the other considers both PUSCH and PUCCH transmit power. The former is called *Type 1 PH* and the latter is called *Type 2 PH*. The calculation of each type of PH is performed with the formula below, where $P_{CMAX,c}$ is the maximum output power of serving cell c. As explained previously, Type 2 PH is only applicable for the PCell while Type 1 PH is applicable for both the PCell and SCells.

- Type 1 PH = $P_{CMAX,c}$ − PUSCH power.
- Type 2 PH = $P_{CMAX,c}$ − (PUSCH power + PUCCH power).

In cases where PUSCH or PUCCH are not transmitted, Type 1 or Type 2 PH is calculated based on the reference format that is known to both eNB and UE. This type of PH is called *virtual PH*, meaning that the PH is calculated assuming virtual PUSCH or PUCCH transmission. The virtual PH is also considered useful for the eNB's uplink scheduling, and thus is reported.

The introduction of virtual PH brings another issue to the eNB, in that it has to identify the type of PH of each serving cell. Without type identification, the eNB cannot know whether the reported PH is based on actual or virtual transmission, which will lead to incorrect estimation of the corresponding cell's UL CC power situation. One way to solve this problem is for the eNB to take the uplink scheduling into consideration for PH type identification. However, this increases eNB complexity considerably, and, furthermore, sometimes it does not give an accurate result due to PDCCH detection errors. Therefore, it is necessary to indicate the type of PH – normal or virtual – explicitly.

For both Type 1 and Type 2 PHs, $P_{CMAX,c}$ should be reported to the eNB; this was not the case for Release 8/9 PH. This is because per-CC PH may not provide sufficient information on the UE power status to the eNB. For example, there is a case where the reported PHs indicate a good enough power headroom for each UL CC while the total UE power exceeds the maximum. Therefore, the UE should report the $P_{CMAX,c}$ together with the associated PH. One exception to this is virtual PH, which is calculated based on the reference format. Since the $P_{CMAX,c}$ is already known by the eNB, it is pointless to report the $P_{CMAX,c}$ to the eNB. Thus, $P_{CMAX,c}$ is not reported if the reported PH type is virtual. However, for normal PH, the $P_{CMAX,c}$ is always reported together with the associated PH.

Transmission of PHR is controlled by timers and parameters; that is, the *prohibitPHR-Timer*, *periodicPHR-Timer*, and *dl-PathlossChange*, similar to Releases 8/9. These are configured per UE, and the same values apply to all serving cells. The *prohibitPHR-Timer* controls the frequency of PHR transmission, and the *periodicPHR-Timer* controls the periodic transmission of PHR. The *dl-PathlossChange* controls event-based PHR transmission; that is, PHR is triggered when a change in the measured DL pathloss exceeds *dl-PathlossChange* from the last transmitted PHR.

In Release 10, PHR is also triggered when a change in required power backoff exceeds *dl-PathlossChange* from the last transmitted PHR. This trigger has been introduced in order to indicate that the UE autonomously reduces transmission power to meet the Specific Absorption Rate (SAR) requirements when the UE transmits on different radio interfaces. Since the PH is calculated after applying autonomous power backoff, it should be indicated that the PH calculation is done with reduced transmission power.

Another new PHR trigger introduced in CA is SCell activation. Since the power situation of an SCell is different before and after activation, it is useful to provide the updated power situation to the eNB in a timely manner. Therefore, PHR is triggered on the activation of an SCell. However, PHR is not triggered on deactivation because the SCell will not be used until it is activated again.

In CA, since multiple serving cells are configured and the power situation of each serving cell is monitored independently, PHR can be triggered at different times for each serving cell. In such a case, the question is which serving cell's PH should be reported when a PHR is triggered by one of the serving cells.

When the *dl-PathlossChange* threshold is satisfied for a serving cell, it is likely that the pathloss or power backoff of other serving cells will also have changed to quite some extent. As a result, it would be better for the eNB to know the power situation of those serving cells together. Actually, it is always beneficial for the eNB to know the power situation of all activated serving cells for future uplink scheduling. Hence, when PHR is triggered by any of the activated serving cells, the PHs of all activated serving cells are reported together at the same time.

For this purpose, new PHR, namely Extended PHR, has been introduced in Release 10. Extended PHR includes PHs of all activated serving cells regardless of whether the PH is normal or virtual. Note that for the PCell, both Type 1 and Type 2 PHs are included, and for an SCell, only Type 1 is included. Note also that Type 2 PH is included only if simultaneous PUSCH and PUCCH transmission is configured.

Since Extended PHR includes PHs of multiple serving cells, identification of the serving cell to which a PH belongs is required. Therefore, an identifier field is included in Extended PHR. Thanks to the identifier field, Extended PHR is self-decodable, and therefore it can be transmitted on any of the activated serving cells. This is different from Releases 8/9, where PHR is transmitted only on the corresponding serving cell. Transmitting Extended PHR on any activated serving cell gives great flexibility to the eNB because the eNB does not have to allocate a UL grant for each serving cell whose PH is reported.

Extended PHR is transported through a new MAC CE called an Extended PH MAC CE. A new LCID is allocated to indicate the Extended PH MAC CE. The format of the Extended PH MAC CE is shown in Figure 12.11.

The C_i field indicates the presence of a PH field of an SCell with *SCellIndex i*. The value "1" indicates that the PH of the SCell *i* is reported, and the value "0" indicates that the PH of the SCell *i* is not reported. Since the presence of a PH field depends on the value of the C_i field, the length of an Extended PH MAC CE varies. Thus, the L field is required in the MAC subheader to indicate the length of the MAC CE (see Section 6.11).

The V field indicates the type of reported PH, whether it is normal or virtual. The value "1" indicates that the PH is virtual, and the value "0" indicates that the PH is normal. In addition, the V field indicates whether the corresponding $P_{CMAX,c}$ field is present or not. Since $P_{CMAX,c}$ is not needed for virtual PH, the $P_{CMAX,c}$ field is present only when the value of the V field is set to "0".

The P field indicates whether the UE applies power backoff due to power management. The value "1" indicates that the corresponding $P_{CMAX,c}$ field would have had a different value if no power backoff due to power management had been applied. With this field, the eNB can know that the UE reduces its transmission power autonomously.

The PH field indicates the power headroom of the corresponding serving cell, and the $P_{CMAX,c}$ field indicates the $P_{CMAX,c}$ used to calculate the PH. For the PCell, the Type 2 PH and corresponding $P_{CMAX,c}$ fields are present only when simultaneous PUSCH and PUCCH transmission is configured.

The R field is the reserved bit.

C_7	C_6	C_5	C_4	C_3	C_2	C_1	R
P	V	PH (Type 2, PCell)					
R	R	$P_{CMAX,c}$ 1					
P	V	PH (Type 1, PCell)					
R	R	$P_{CMAX,c}$ 2					
P	V	PH (Type 1, SCell 1)					
R	R	$P_{CMAX,c}$ 3					
...							
P	V	PH (Type 1, SCell n)					
R	R	$P_{CMAX,c}$ m					

Figure 12.11 Extended Power Headroom MAC CE. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

12.3.3 Logical Channel Prioritization (LCP)

The UE performs the LCP procedure to allocate transmitting resource to each of the multiplexed logical channels (see Section 6.6). By using logical channel priority and the Prioritized Bit Rate (PBR), the UE can allocate the proper amount of transmitting resource to each radio bearer. The transmitting resource is calculated from the UL grant allocated to the UL CC. In Releases 8/9, at most one UL grant is available in a TTI because only one UL CC is configured.

In CA, the LCP procedure needs to be changed in order to process multiple UL grants in a TTI. Since each CC can be provided with a UL grant, the number of UL grants that should be processed in a TTI can be up to the number of activated UL CCs. An issue here is how the UE processes multiple UL grants in a TTI; that is, either independently or jointly.

Independent processing means that the UE processes UL grants one by one in order of receipt. Even within the same subframe, the actual time of receipt could be different for each serving cell. Therefore, it is possible for the UE to identify the order of receipt of UL grants and apply the LCP procedure to each received UL grant. Since multiple UL grants are processed serially, this is sometimes called *serial processing*. The only important aspect in this scheme is that multiple LCP procedures are performed in a TTI. It should be noted that even if multiple LCP procedures are performed in a TTI, the PBR is enforced only once in a TTI by definition of the PBR.

Joint processing means that the UE first sums up the UL grants received in a TTI and then applies the LCP procedure to the total UL grant. The transmitting resource of each logical channel is calculated based on the total UL grant. In this scheme, the LCP procedure is performed only once in a TTI.

In joint processing, RLC PDUs included in each MAC PDU are different depending on the processing order of UL grants. This means that the QoS of the radio bearer will be impacted by the UL grant processing order. If radio conditions of UL CCs are different, it is better to process the UL grant of a better quality UL CC first and a poor quality UL CC last in order to ensure that high priority data are transmitted on the better quality UL CC.

However, at least in Release 10, it is reasonable to assume that all activated UL CCs experience similar channel quality. Considering that only intra-band scenarios are supported in Release 10, the characteristics of UL CCs will not be much different. Moreover, since the eNB is aware of the UE transmit power (by PHR) and the channel quality (by CQI), it will select appropriate UL grants to provide UL CCs with similar quality. Given that the channel quality is similar among UL CCs, the processing order is not such an important factor to consider in joint processing. Therefore, it is left to UE implementation to decide which UL grant the UE processes first in joint processing.

Comparing the two schemes, independent processing is simpler from the implementation point of view, because it is a simple extension of the existing LCP procedure. It is enough to apply the LCP procedure multiple times in a TTI. However, from a performance point of view, joint processing is better, in that it results in less segmentation in the RLC layer. Although there is a tradeoff between implementation complexity and segmentation performance, the difference is not so significant, and thus both schemes are allowed in CA. Therefore, it is left to UE implementation whether to take independent processing or joint processing for multiple UL grants received in a TTI.

12.3.4 Buffer Status Report (BSR)

The BSR procedure in CA is not much different from Releases 8/9. Three types of BSR are still used depending on the triggering condition – that is, Regular BSR, Periodic BSR, and Padding BSR – and at most one BSR is included in a MAC PDU (see Section 6.5.1). The only difference is the number of BSRs transmitted in a TTI. In Releases 8/9, at most one BSR is transmitted in a TTI because only one serving cell is configured. However, in CA, multiple BSRs can be transmitted in a TTI because multiple serving cells are configured (one MAC PDU per UL CC).

One restriction in transmitting multiple BSRs in a TTI is that only one Regular or Periodic BSR is allowed in a TTI. The reason for this restriction is that a Regular or Periodic BSR has higher priority over user data in Logical Channel Prioritization, therefore, including a Regular or a Periodic BSR in all MAC PDUs would consume much of the transmitting resource that could be allocated for user data.

It is, however, acceptable to transmit multiple Padding BSRs in a TTI. The inclusion of a Padding BSR in a MAC PDU is decided for each MAC PDU using the same rule as in Releases 8/9; that is, include one when the space remaining after the inclusion of MAC SDUs and the necessary MAC CEs can accommodate a Padding BSR.

When multiple BSRs are transmitted in a TTI, all BSRs reflect the same Buffer Size. This is because the Buffer Size of the BSR is determined after all MAC PDUs are built for the

TTI. One may think that multiple transmission of the same BSR would waste radio resource. However, since a Regular or Periodic BSR is transmitted only once and a Padding BSR utilizes only padding bits, there is no extra cost for transmitting multiple BSRs with the same Buffer Size in the same TTI. Rather, it can improve the BSR reliability.

An important aspect in CA BSR is that the Buffer Size can indicate a wider range than a Release 8/9 BSR. More specifically, the Buffer Size in Releases 8/9 can indicate up to 150 000 bytes while the Buffer Size in CA can indicate up to 3 000 000 bytes, which is 20 times larger than the Release 8/9 Buffer Size. Since the data rate is increased significantly in Release 10 with the aid of CA and MIMO techniques, the Buffer Size level has also been changed to cope with such a high data rate. This is called the Extended Buffer Size level as compared to the legacy Buffer Size level. For a single UE, whether to use the Extended Buffer Size level or the legacy Buffer Size level is configured by the eNB.

12.3.5 Discontinuous Reception (DRX)

In order to extend the DRX operation to CA, two different alternatives were considered: UE-specific DRX and Cell-specific DRX. In UE-specific DRX, a common DRX operation is applied to all configured cells. Since the Active Time is the same for all cells, the UE monitors the PDCCH of all DL CCs in the same subframe. DRX-related timers and parameters are configured per UE for the UE-specific DRX operation. In Cell-specific DRX, however, each cell performs its own DRX operation. Since the Active Time is configured independently, one cell's activity does not affect another cell's DRX operation. DRX-related timers and parameters need to be configured per cell in this case.

Between the two alternatives, 3GPP decided to use UE-specific DRX for CA. Although Cell-specific DRX provides a better power saving, UE-specific DRX was chosen because it is simple and provides much commonality with Release 8/9 DRX. Moreover, the power saving can be improved by SCell deactivation, which is another reason for choosing UE-specific DRX.

12.3.6 Semi-Persistent Scheduling (SPS)

Semi-Persistent Scheduling is used to alleviate the load on the PDCCH when a large number of simultaneous VoIP users need to be scheduled (see Section 6.4). For a single UE, only one pattern of SPS resource is configured because the UE typically uses only one VoIP service at a time.

In CA, the SPS resource configuration is limited to the PCell. If the SPS resource were configured to an SCell, reconfiguration would be needed each time the SCell was activated/deactivated or added/removed. Since the PCell is never deactivated and not removed unless handover is performed, the SPS resource on the PCell can be utilized stably without further complications.

In a certain subframe, the SPS resource can be overridden by the dynamic resource when dynamic scheduling is performed for a subframe. Since the SPS resource is only configured to the PCell, it is obvious that the dynamic resource of the PCell can override the SPS resource. However, whether the dynamic resource of an SCell could override the SPS resource in the PCell was the subject of debate.

Allowing the SCell dynamic resource to override the PCell SPS resource would give flexibility to the eNB in using the PDCCH. However, it would complicate the UE implementation, in that the PDCCH reception on the SCell would need to be considered in the SPS operation. Moreover, the probability of PDCCH error would be multiplied by the number of serving cells, which would make it difficult for the eNB to predict the UE's actual transmission. For this reason, it was decided that only the PCell dynamic resource should be able to override the SPS resource.

References

1. Iwamura, M., Etemad, K., Fong, M.-H. *et al.* (2010) Carrier aggregation framework in 3GPP LTE-advanced. *IEEE Communications Magazine*, **48**, 60–67.
2. 3GPP Technical Specification 36.300, "E-UTRA and E-UTRAN Overall Description Stage 2 (Release 10)", www.3gpp.org, accessed on 14 March 2012.
3. 3GPP Technical Specification 36.133, "Requirements for support of radio resource management (Release 10)", www.3gpp.org, accessed on 14 March 2012.

13

Relay

Relay is the technology that relays radio transmissions between the eNB and the UE in a wireless manner. It is regarded as one of the key technologies introduced in LTE Release 10, and is realized by an additional network node called the Relay Node (RN) located between the eNB and the UE. By using the Relay Node, operators can improve cell coverage and throughput in various types of locations with relatively low CAPital EXpenditure (CAPEX) and OPerating EXpenditure (OPEX).

13.1 Deployment Scenarios

Deploying an eNB in every location is not cost effective and sometimes not easy, in that it requires fixed and wired backhaul. During the study of relay technology, various scenarios were identified where relay deployment was deemed much more useful than normal eNB deployment [1]. The identified scenarios are summarized below:

- **Coverage extension:** An RN can be deployed at a cell edge to extend the coverage of the eNB. A typical deployment would be at a cell edge in a rural area where the population is sparse.
- **Dead spot mitigation:** An RN can be deployed in a dead spot to remove a coverage hole from the cell. A coverage hole is usually caused by physical obstructions such as buildings, tunnels, and so on.
- **Throughput enhancement:** An RN can be deployed in a hot spot or indoor area to increase the throughput in that particular area.
- **Temporary coverage:** An RN can be deployed in an area hosting special events (e.g., sports games, music concerts) or experiencing natural disasters (e.g., an earthquake, a tsunami) to provide temporary coverage.
- **Group mobility:** An RN can be deployed in a moving vehicle such as a bus, train, and so on. Unlike the other scenarios, the RN in this scenario is subject to mobility.

Although relay deployment is deemed useful in various scenarios, the scenario of coverage extension in rural areas has been prioritized in LTE Release 10, focusing on the operator-deployed, stationary-type, single-hop RN. The simplification of the deployment scenario

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

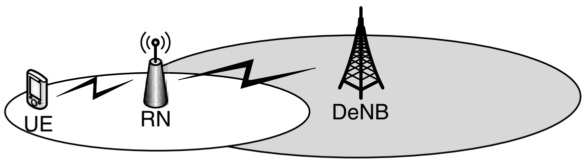

Figure 13.1 Relay Node deployment scenario – coverage extension in a rural area

reduced significant standardization efforts, and thus enabled the RN to be available in the timeframe of Release 10. The prioritized scenario of coverage extension is illustrated in Figure 13.1.

13.2 Network Architecture for the Relay Node

The overall network architecture supporting the RN is shown in Figure 13.2. Compared to the conventional network architecture, a new network node called the RN has been introduced between the UE and the eNB. Due to the introduction of the RN, a new radio interface called the Un interface has also been introduced between the RN and the eNB [2].

As the Un interface is considered not much different from the Uu interface, most of the Uu functions are applied equally to Un radio protocols. However, some Uu functions have been modified for RN-specific operation, as explained in Section 13.4. To support two different radio interfaces, an RN is equipped with two wireless transmitter/receiver pairs, one for the Uu interface and one for the Un interface.

The eNB serving the RN is called the Donor eNB (DeNB). The difference between the DeNB and the normal eNB is that the DeNB supports both Uu and Un radio protocols while the normal eNB supports only Uu radio protocols. Since the DeNB is equipped with both Uu

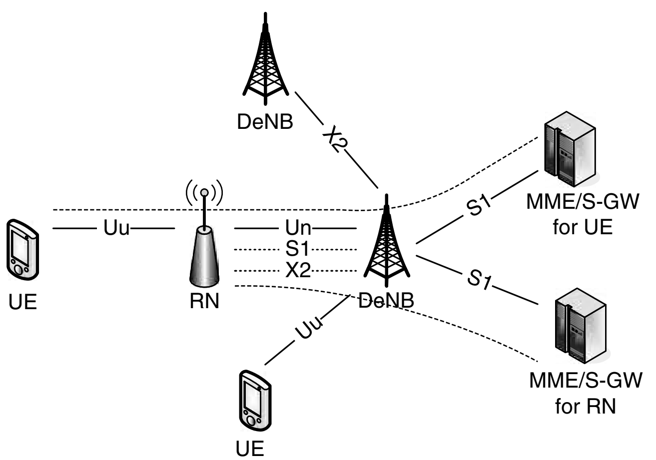

Figure 13.2 Network architecture supporting a Relay Node

and Un radio protocols, it is able to serve both the RN and the UE simultaneously. The number of RNs that can be served by a DeNB is not limited.

The RN is also equipped with both Uu and Un radio protocols. Before connecting to a DeNB, the RN acts as a UE and uses Uu radio protocols initially to connect to a DeNB. Once connected, the RN can act as an eNB and uses Un radio protocols to communicate with the DeNB and Uu radio protocols to communicate with UEs under its coverage. The detailed procedure for the RN to become an eNB is explained in Section 13.5.

Also, from the network interface point of view, the RN has both UE and eNB aspects. As a UE, the RN is connected to its own MME and S-GW through the DeNB. In this case, the S1 and X2 interfaces terminate in the DeNB. However, when the RN becomes an eNB, it is connected additionally to the UE's MME and S-GW through the DeNB. The S1 and X2 interfaces for the serving UE terminate in the RN in this case. This means that the Un radio protocols should support transfer of S1 and X2 messages as well as GPRS Tunneling Protocol-User plane (GTP-U) data, which is the major difference from Uu radio protocols.

The DeNB can provide S1 and X2 proxy functionality between the RN and other network nodes (i.e., MMEs, S-GWs, and eNBs). In other words, the DeNB plays the role of MME (for the S1 control plane), S-GW (for the S1 user plane), or eNB (for X2) to the RN, and consequently the DeNB appears as corresponding network nodes to the RN. The S1 and X2 proxy functionality includes passing UE-dedicated S1 and X2 signaling messages as well as GTP data packets between the S1 and X2 interfaces associated with the RN and other network nodes.

The control plane and user plane protocol architectures supporting the RN are shown in Figures 13.3 and 13.4, respectively. Note that both control plane messages (S1AP/X2AP) and user plane data (GTP-U) are transmitted as IP packets in the Un interface.

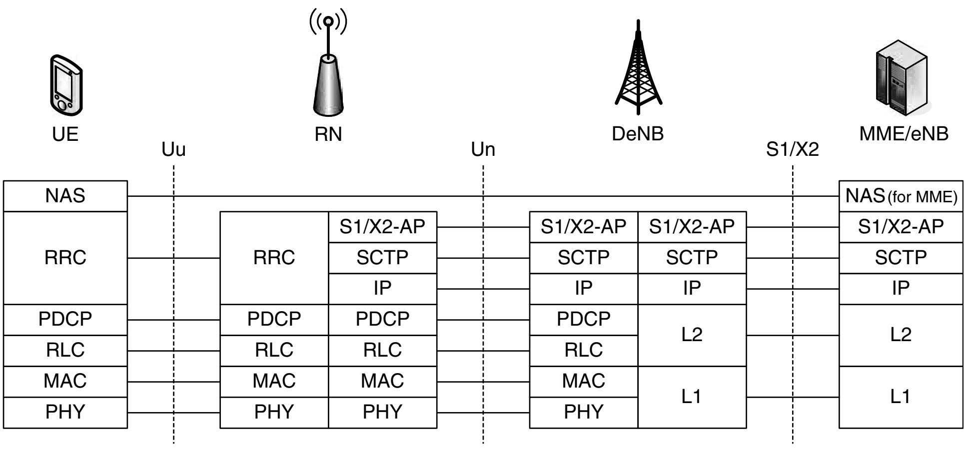

Figure 13.3 Control plane protocol architecture supporting a Relay Node. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

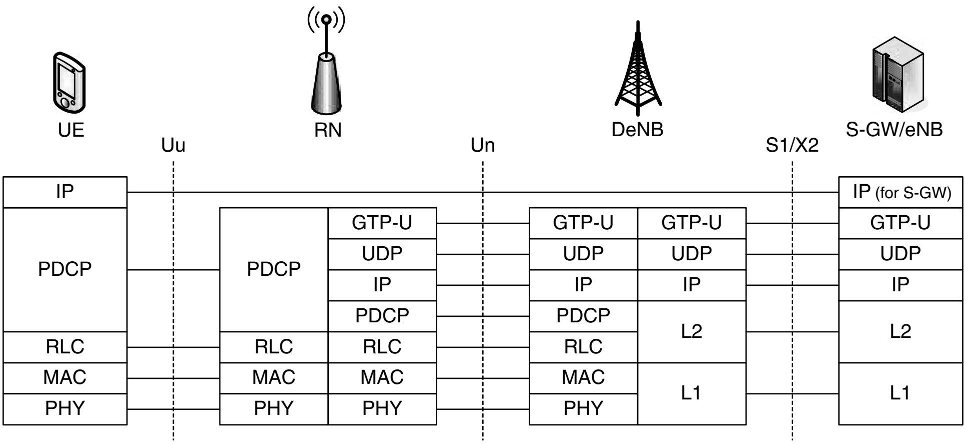

Figure 13.4 User plane protocol architecture supporting a Relay Node. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

13.3 Types of Relay Node

During the study of relay technology, various types of RN were discussed. Variation among the types is based mainly on two aspects: which layer performs relaying and whether the Un link shares a frequency with the Uu link. Each aspect is explained in the following subsections.

13.3.1 Layer Performing Relaying

In terms of the layer performing relaying, three types of RN were considered: Layer 1 RN, Layer 2 RN, and Layer 3 RN. These types of RN are shown in Figures 13.5–13.7, respectively.

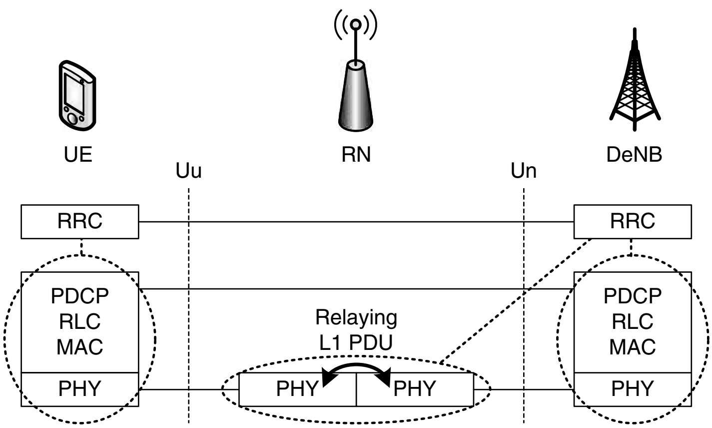

Figure 13.5 Protocol architecture for a Layer 1 RN

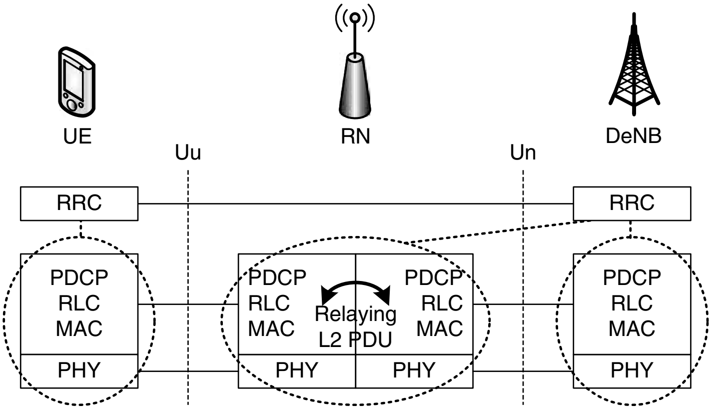

Figure 13.6 Protocol architecture for a Layer 2 RN

A Layer 1 RN, sometimes called a repeater, relays RF signals in the Physical layer, only amplifying the RF signals. Since no further data processing is performed, the processing delay is shorter than for other types of RN, but it cannot improve the Signal to Interference and Noise Ratio (SINR) because the interference and noise are amplified together with the desired signal. The Physical layer functions of this RN are controlled by the RRC layer of the DeNB.

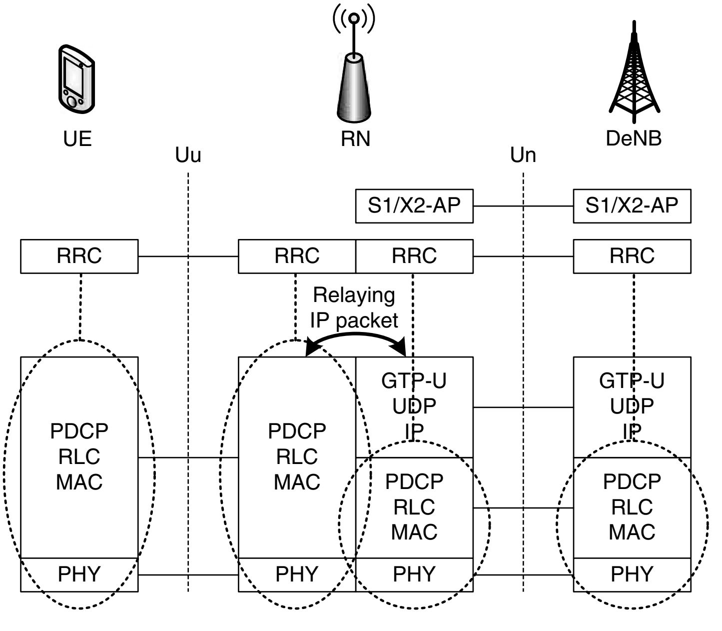

Figure 13.7 Protocol architecture for a Layer 3 RN

A Layer 2 RN is designed to overcome the drawback of a Layer 1 RN by decoding and encoding of the RF signals in layer 2. Although the processing delay is increased slightly compared to a Layer 1 RN, interference and noise can be eliminated by the decoding and encoding processes in layer 2. Depending on the actual sub-layer performing relaying, a Layer 2 RN can be classified further into a MAC RN, an RLC RN, or a PDCP RN. Since there is no RRC layer in the RN, the RRC layer in the DeNB should control all layer 1 and layer 2 functions in the RN; for example, data formatting, retransmission, and scheduling in the RN need to be controlled closely by the DeNB. This is the main drawback of a Layer 2 RN, because the coordination of two network nodes is not an easy task. Moreover, additional layer 2 functionalities are required to support close coordination between the RN and the DeNB. For this reason, a Layer 2 RN is not preferred, even though it has superior performance.

A Layer 3 RN is similar to a Layer 2 RN in that interference and noise can be eliminated by layer 2 processing. However, a Layer 3 RN has full layer 3 capability; that is, it is equipped with an RRC layer as well as layer 1 and layer 2 radio protocols. The existing Uu radio protocols apply equally to the Layer 3 RN except for a small modification to the Un side. Relaying is performed for IP packets above the PDCP layer, and the RRC layer controls all layer 1 and layer 2 radio protocols. Thanks to the RRC layer, a Layer 3 RN can control its cells with their own cell identities (PCIs), and, as a result, it appears to the UE to be a normal eNB. The drawback of a Layer 3 RN is the increased processing delay caused by layer 2 protocols where the IP packets are reassembled and formatted again.

Among the three RN types, the Layer 3 RN has been adopted for LTE Release 10 for the following reasons:

- **Standard impact:** The existing Uu radio protocols are applied to both Uu and Un sides, with only a small modification to the Un side. This eases the implementation of RNs in the Release 10 timeframe.
- **Backward compatibility:** For the Uu side, a Layer 3 RN operates as a normal eNB with the existing Uu radio protocols. Thus, legacy UEs (e.g., Release 8 UEs) can connect to the Layer 3 RN without any special mechanism.
- **Extensibility:** Once a Layer 3 RN is connected to the DeNB, it can operate independently from the DeNB; this means that a multi-hop RN could be supported easily by a Layer 3 RN. Although a multi-hop RN is not supported in Release 10, it is thought to be a useful feature for future releases.

13.3.2 Frequency Separation of Uu and Un Links

The RN can be categorized into in-band and out-band RNs depending on the relationship between the operating frequencies used in Uu and Un links [3].

In an in-band RN, the Un link is operating on the same frequency as the Uu link. Since an additional frequency is not required, it is economical for operators to deploy this type of RN. However, two radio links sharing the same frequency causes self-interference problems; that is, transmission to one link causes interference to another link. Thus, to prevent self-interference problems in an in-band RN, a special resource separation method such as time-domain separation should be used in Uu and Un links. Time-domain separation ensures that the in-band RN does not transmit and receive simultaneously on two radio links.

If the Un link is operating on a different frequency to the Uu link, this is called an out-band RN. Since the Uu and Un links are isolated by frequency, the out-band RN can transmit and receive simultaneously on both links. No additional mechanism is required for separation of Uu and Un links. Although the operation is much simpler than for an in-band RN, it is not an economical choice to allocate a separate frequency to the Un interface.

In LTE Release 10, it was decided to support both in-band and out-band RNs. In other words, an RN should be capable of operating in one of the in-band and out-band modes. Operators can choose the operating mode depending on their frequency planning.

It has already been explained that the relay type supported in terms of the relaying layer is the Layer 3 RN. Therefore, the supported relay types in LTE Release 10 are an in-band Layer 3 RN and an out-band Layer 3 RN. In the 3GPP standard, the former is called a Type 1 RN, and the latter is called a Type 1a RN.

13.4 Relay Node-Specific Operation

The RN supports most of the Uu functions as a UE or an eNB. The connection to the DeNB is controlled by the Uu functions of the UE, and the connection to the UEs under the RN is controlled by the Uu functions of the eNB. The Uu functions are applied to the RN without any modification.

The Un functions are used in the Un interface when the RN operates as an eNB. Although the Uu functions are also well suited for Un operation, more functions have been introduced to support efficient RN operation.

This section explains the new functions specific to RN operation.

13.4.1 Bearer Mapping

A UE can support up to eight Data Radio Bearers (DRBs) in the Uu interface (see Section 3.6). Likewise, the RN can support up to eight DRBs in the Un interface. However, the RN should support more than eight DRBs in the Uu interface because the RN serves multiple UEs and each UE can be configured with up to eight DRBs.

Due to the limited number of Un DRBs, the RN performs *N*-to-1 mapping between Uu and Un DRBs for the uplink direction. Bearer mapping is performed based on the QoS Class Identifier (QCI, see Section 1.5.2) so that multiple Uu DRBs having the same or similar QoS are mapped to a Un DRB. For the downlink direction, the same *N*-to-1 mapping is performed by the DeNB.

The RN obtains the actual mapping rule between Uu and Un DRBs by two pieces of information: the QCI-to-DiffServ Code Point (DSCP) and the DSCP-to-Un DRB. The QCI-to-DSCP is provided by the Operation, Administration, and Maintenance (OAM), indicating which QoS level of Uu DRBs should be mapped to a DSCP value. The DSCP field, which is included in the GTP-U header, is associated with a Un DRB, where the association between the two is signaled by the DeNB by means of a Traffic Flow Template (TFT). In this way, the RN can map the Uu DRB to a Un DRB unambiguously. Note that when this mapping information is received depends on implementation.

An example of bearer mapping in the RN is shown in Figure 13.8. In the figure, all VoIP bearers of UEs are mapped to UnRB1, Web-browsing bearers to UnRB2, and streaming bearers to UnRB3.

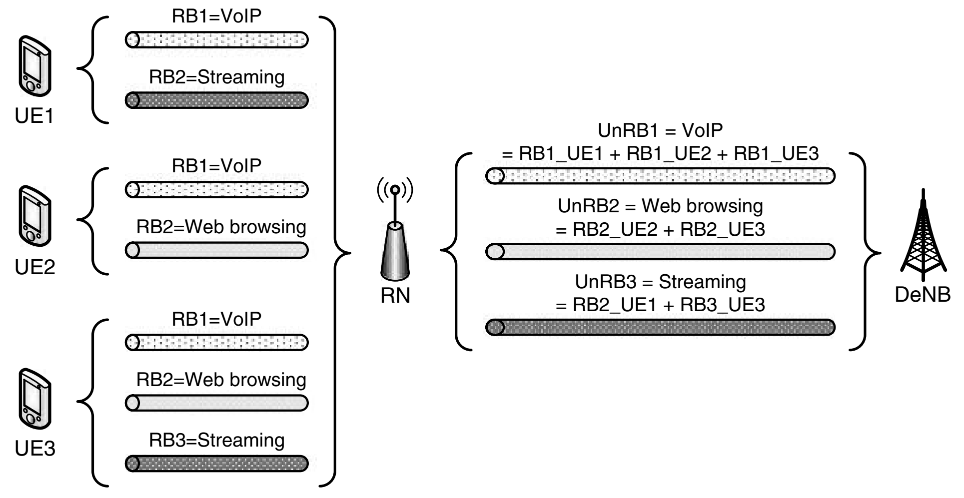

Figure 13.8 Bearer mapping per QCI in an RN

N-to-1 mapping means that multiple Uu DRBs are multiplexed into a single Un DRB. Since only a Layer 3 RN is supported in Release 10, the RN performs multiplexing above the PDCP layer for IP packets. The identification of multiplexed Uu DRBs in a Un DRB is achieved by the GTP-U Tunneling Endpoint ID (TEID) included in the GTP header. Figure 13.9 shows IP packet multiplexing performed in the RN.

13.4.2 Integrity Protection for a UnDRB

In the LTE Release 8 PDCP protocol, integrity protection is not applied to DRBs because applying integrity protection would result in an additional 4-byte overhead in a PDCP PDU (see Section 4.3). It was deemed not worth increasing the overhead in user plane data

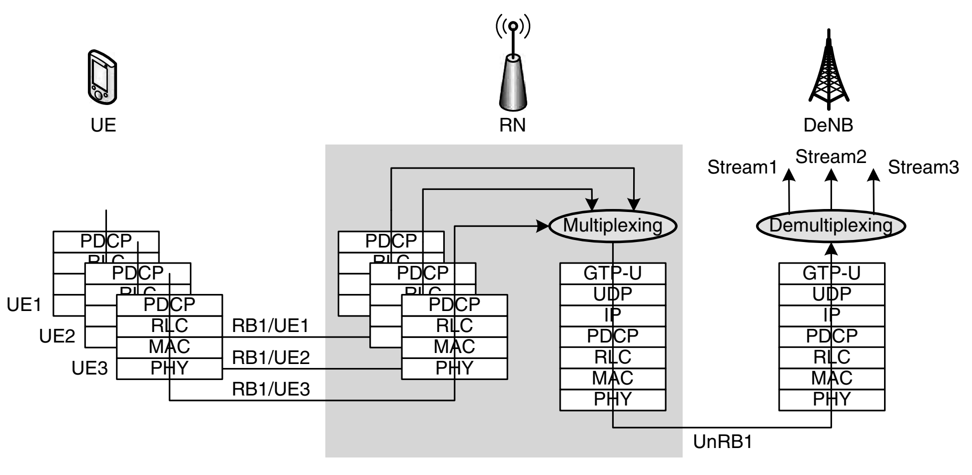

Figure 13.9 IP packet multiplexing for uplink transmission

considering the importance of the contents of the packet. Thus, in Release 8, integrity protection was applied only to SRBs.

In the relay environment, however, the situation is different because control plane messages such as S1AP and X2AP are transmitted on DRBs in the Un interface, as shown in Figure 13.3. Therefore, it is required that integrity protection be applied to DRBs in addition to SRBs in the Un interface.

To support integrity protection on DRBs in the Un interface, the RRC and PDCP of Un protocols have been changed slightly from those of Uu protocols. The changes are explained below.

First, to make integrity protection configurable for each DRB, an indicator called *rn-IntegrityProtection* is included in the PDCP configuration message, *PDCP-Config*. This indicator shows whether integrity protection should be applied to the DRB or not. The reason for having configurability is that integrity protection is required for S1AP/X2AP messages but is still not necessary for user plane data. The use of the integrity protection function is optional for DRBs carrying user plane data.

Secondly, to perform integrity protection on DRBs, a separate AS security key called K_{UPint} is derived from the K_{eNB} similar to other security keys, K_{RRCint}, K_{UPenc}, and K_{RRCenc} (see Section 3.5). In Release 8, where integrity protection is applied only to SRBs, the integrity key is derived only for SRBs; that is, K_{RRCint}. However, for the RN, two integrity keys (K_{UPint}, K_{RRCint}) are derived in addition to two ciphering keys (K_{UPenc}, K_{RRCenc}). The management of K_{UPint} is the same as for other security keys.

Thirdly, a new PDCP PDU format is introduced to support integrity protection for the user plane PDCP Data PDU. In Release 8, the MAC-I field is included in the PDCP Data PDU for SRBs, but is not included in the PDCP Data PDU for DRBs (see Section 4.7). Thus, to include the MAC-I field in the PDCP Data PDU for DRBs, a new PDCP Data PDU format has been defined, as shown in Figure 13.10. This format is the same as the PDCP Data PDU format for DRBs using 12-bit PDCP SN except that a 4-byte MAC-I field is attached at the end of the PDU. Note that only a 12-bit PDCP SN is used in this case. The RN shall use this format in the Un interface when integrity protection is configured for the Un DRB.

Lastly, handling of integrity verification failure of DRBs is different from that of SRBs. In Release 8, when integrity verification failure occurs for SRBs, the PDCP indicates the failure to the RRC to trigger the RRC Connection Re-establishment procedure (see Section 4.3). For an RN, the same applies to SRBs, but different behavior applies to DRBs.

When an integrity verification failure occurs for Un DRBs, the RN just discards the integrity-failed PDU without triggering the RRC Connection Re-establishment procedure. This is because, if the DRB is an AM DRB, the COUNT value is not reset by the RRC Connection Re-establishment procedure in order to support lossless handover (see Section 4.6). Since the integrity failure comes mostly from corrupted COUNT values, the integrity problem may not be fixed if the COUNT value is not reset. Thus, triggering the RRC Connection

D/C (1bit)	PDCP SN (12bits)	Data	MAC-I (4bytes)

Figure 13.10 PDCP Data PDU format for Un DRBs supporting integrity protection. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

Re-establishment procedure in this case is not deemed to be useful. Rather than implementing a special mechanism, the 3GPP decided to leave the handling of DRB integrity failure to RN implementation. Since the probability of integrity failure is, in any case, very low, leaving this up to implementation should not degrade the RN performance.

13.4.3 RN Subframe Configuration

A Layer 3 RN can operate in two different modes: in-band and out-band. Out-band operation is straightforward because the Uu and Un links are isolated by frequency. Additional frequency is only required for out-band operation. However, in-band operation is quite complex, in that the Uu and Un links need to be isolated in the time domain.

Time-domain isolation for the downlink direction is realized by creating "gaps" in the RN-to-UE transmission. During the "gaps", the RN does not transmit to UEs but receives from the DeNB. The "gap" is a special subframe reserved for DeNB-to-RN transmission. The downlink operation of an in-band RN is shown in Figure 13.11.

The "gap" information is signaled from the RN to UEs by means of the MBSFN subframe configuration (see Section 11.3). The MBSFN subframe was originally designed for MBMS transmission, but it can be used for RN purposes, in which case it is called the RN subframe. With the knowledge of RN subframe configuration, the UEs attached to the RN can avoid any attempt to receive data from the RN in the RN subframes; that is, the UEs should not listen to the RN at all in the RN subframes.

To ensure backward compatibility to the UEs, the RN should transmit some control information such as the PDCCH, PCFICH, and/or PHICH in every subframe. Even in RN subframes, the RN should transmit the control information. The transmission of control information takes place at the start of each subframe.

When the RN transmits the control information to the Uu link, the RN cannot receive control information from the DeNB. Without the control information, the RN cannot receive data from the DeNB. Therefore, a new physical control channel called the Relay-PDCCH (R-PDCCH) has been introduced. The R-PDCCH is transmitted at the part of an RN subframe where the RN is not transmitting control information to the Uu link.

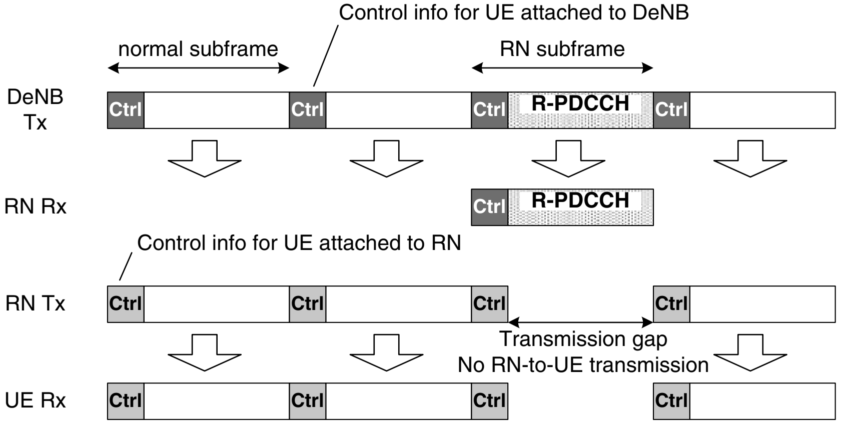

Figure 13.11 Downlink operation of an in-band RN

For the uplink direction, such an RN subframe configuration is not needed because the RN has full control over UEs' uplink transmission. When the RN wants to transmit to the DeNB, it can block UE transmission by not allocating uplink grant to the UEs. For this reason, the RN subframe is sometimes called the downlink subframe.

When the RN establishes an RRC connection to a DeNB, it indicates the need for RN subframe configuration in the *RRCConnectionSetupComplete* message. If RN subframe configuration is requested, the DeNB provides the information on RN subframe configuration using the *RNReconfiguration* message. Upon receipt of the RN subframe configuration, the RN applies the configuration immediately. Therefore, it is possible for the subframe configurations on Uu and Un to be misaligned temporarily until the subframe configuration on Uu is changed according to the RN subframe configuration.

Once configured, the RN subframe configuration is maintained until released. The RN releases the RN subframe configuration when the RN initiates an RRC Connection Re-establishment procedure or goes into RRC_IDLE.

13.4.4 Update of System Information

In Release 8, the UE acquires updated system information through the System Information Acquisition procedure (see Section 3.2). This procedure can be initiated by various triggers, such as receipt of a *Paging* message, expiry of the maximum validity duration (i.e., 3 hours), return from out-of-coverage, and so on. When the procedure is initiated, the UE reads the system information broadcast in the predefined subframes and gets updated system information. The updated system information replaces all stored system information.

The same mechanism is applied to an out-band RN for acquiring updated system information. Since an out-band RN uses a different frequency in the Un interface, it can receive updated system information without any subframe restriction.

For an in-band RN, however, it may not always be possible to receive updated system information, because an in-band RN monitors only Un subframes complying with RN subframe configuration. Therefore, dedicated RRC signaling is used to provide updated system information to an in-band RN.

Upon a change in any system information relevant to the RN, the DeNB provides the updated system information to an in-band RN using the dedicated RRC message, *RNReconfiguration*. The difference between system information acquired by the acquisition procedure and dedicated signaling is that the dedicated system information contained in the *RNReconfiguration* message does not replace all stored system information but only the corresponding stored system information. The dedicated system information takes precedence over any corresponding system information acquired by the acquisition procedure, and remains valid until overridden; that is, the restriction of maximum validity duration is not applied.

The *RNReconfiguration* message can also be used to configure the RN with different parameter values than those broadcast in system information; this applies to both in-band and out-band RNs.

13.4.5 RN Reconfiguration Procedure

As already explained, the RN Reconfiguration procedure is used to configure/reconfigure the RN subframe configuration and/or to update the system information. This procedure is specific to RN operation and is independent of the RRC Connection Reconfiguration procedure

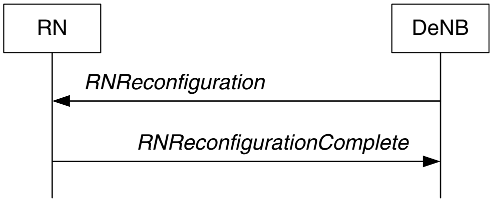

Figure 13.12 RN Reconfiguration procedure. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

used for Uu reconfiguration. The overall RN Reconfiguration procedure is shown in Figure 13.12.

The DeNB initiates this procedure by sending an *RNReconfiguration* message to the RN in RRC_CONNECTED. The DeNB can send this message at any time after security has been activated in the RN. The *RNReconfiguration* message contains only two pieces of information: RN subframe configuration and RN-dedicated system information. The RN-dedicated system information includes System Information Blocks Type 1 and 2.

After applying the configuration provided in the *RNReconfiguration* message, the RN sends an *RNReconfigurationComplete* message to the DeNB as an acknowledgment of the RN reconfiguration. The RN is then ready to serve UEs as an eNB.

13.5 Relay Node Start-Up Procedure

The RN Start-Up procedure is used to configure the RN with the necessary parameters to make it ready to start RN operation. After power-on, the RN performs a two-phase start-up procedure, as shown in Figure 13.13. The reason for having two phases is that at power-on, the RN does not know which cells are allowed to attach. Since not all eNBs can serve RNs (due to operator policy, cell load situations, etc.), the RN needs to identify a cell that can support RN operation. Phase I is used for this purpose; that is, to obtain information on attachable cells. However, if the RN already knows the attachable cells, it can bypass phase I and go directly to phase II. Only after phase II is completed can the RN serve UEs as an eNB.

In phase I, the RN attaches to the MME in the same way as a legacy UE. That is, the RN selects a cell with the best radio quality, makes an RRC connection to the eNB, and performs the UE Attach procedure to the MME. After attaching to the MME, the RN contacts the OAM to download the initial RN configuration including the attachable DeNB cell list. Then, the RN detaches from the MME and goes back to RRC_IDLE.

In RRC_IDLE, the RN selects a cell with the best radio quality among the cells in the attachable DeNB cell list. The RN makes an RRC connection to the DeNB for the selected cell, by which process phase II starts.

In phase II, the RN makes the RRC connection as an RN, which is different from the RRC connection in phase I. The inclusion of an indicator called *rn-SubframeConfigReq* in the *RRCConnectionSetupComplete* message indicates that the RRC connection has been triggered by the RN. Indication of the RN at RRC connection is helpful for the DeNB to search for an RN-capable MME. The indicator also indicates whether the RN requires RN subframe configuration (for an in-band RN) or not (for an out-band RN).

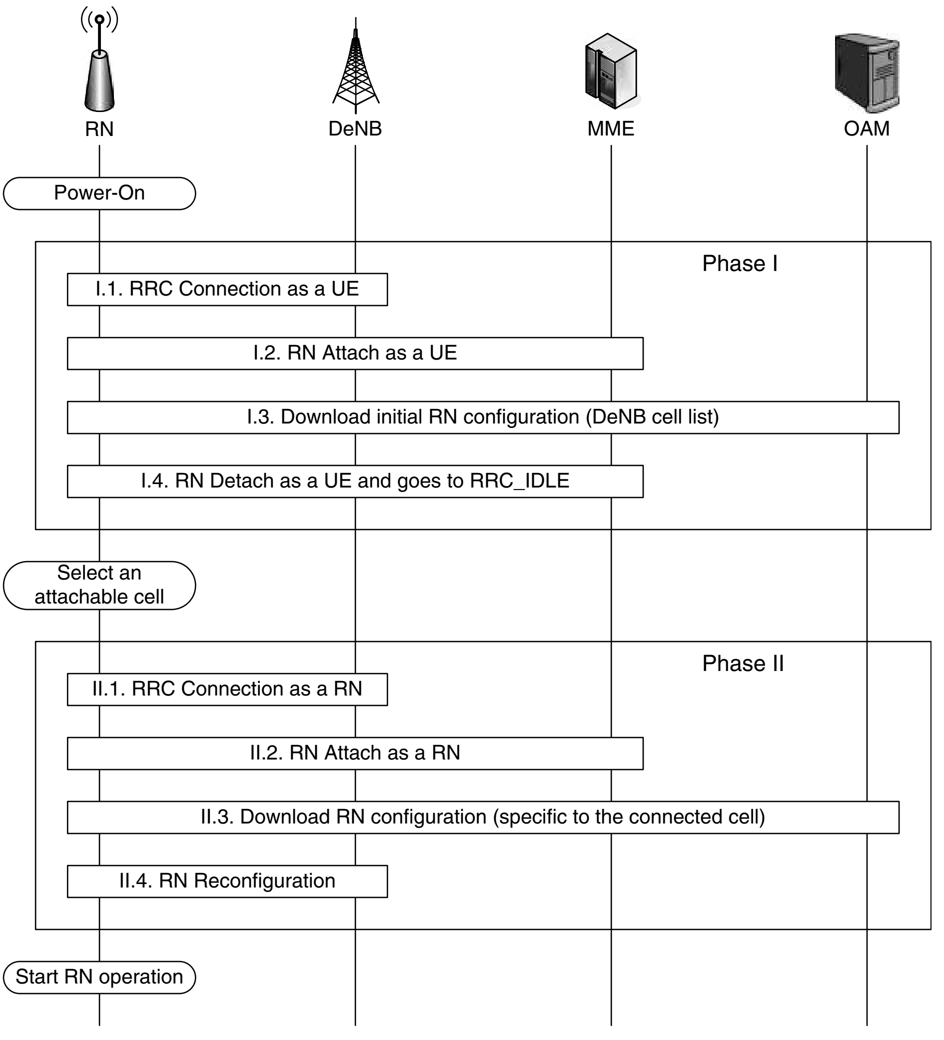

Figure 13.13 RN Start-Up procedure. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

When the DeNB finds an RN-capable MME, the RN performs the RN Attach procedure to it and then contacts the OAM to download further RN configurations specific to the connected cell. After that, the RN sets up S1 and X2 connections with the DeNB, and the DeNB initiates an RN Reconfiguration procedure to provide the RN with RN subframe configuration and RN-dedicated system information. The RN is then ready to start RN operation; that is, it broadcasts system information and prepares for UE access.

13.6 Simplified Operation of Release 10 Relay Node

At the beginning of LTE Release 10, Relay was considered a key feature of LTE-Advanced, and hence many new features were studied and discussed. However, due to the limited time schedule, 3GPP decided to introduce only one simple type of RN in Release 10. As already explained in Section 13.1, the Relay deployment scenario is limited to coverage extension in rural areas, which means that the Release 10 RN does not consider multi-hop operation, mobility between different eNBs, and UE mobility between different RNs. Only an operator-deployed, stationary-type, single-hop RN is considered in Release 10. The simplification of the deployment scenario further leads to simplification of the Un radio protocols. Therefore, it was decided not to include the following features in Release 10 even though they were considered useful.

The first feature is data forwarding at handover. Since neither the RN handover nor the UE handover is supported, it was decided not to consider data forwarding in the Un interface. However, if handover is supported in future releases, data forwarding may be required, in which case the handover interruption time will be increased significantly. At such time, a mechanism to reduce the handover interruption time needs to be considered.

The second feature is downlink flow control in the Un interface. Since the RN has two radio links, it is important to balance the utilization of Uu and Un links. Under-utilization of a radio link leads to radio resource wastage, and over-utilization leads to packet discard. Thus, it is important that the two radio links are utilized properly.

For the uplink direction, radio link utilization is controlled by the RN (for a Uu link) and the DeNB (for a Un link). Thus, no additional mechanism is required to control the transmission in each radio link. However, for the downlink direction, a flow control mechanism may be needed to control the DeNB's downlink transmission depending on the Uu link situation.

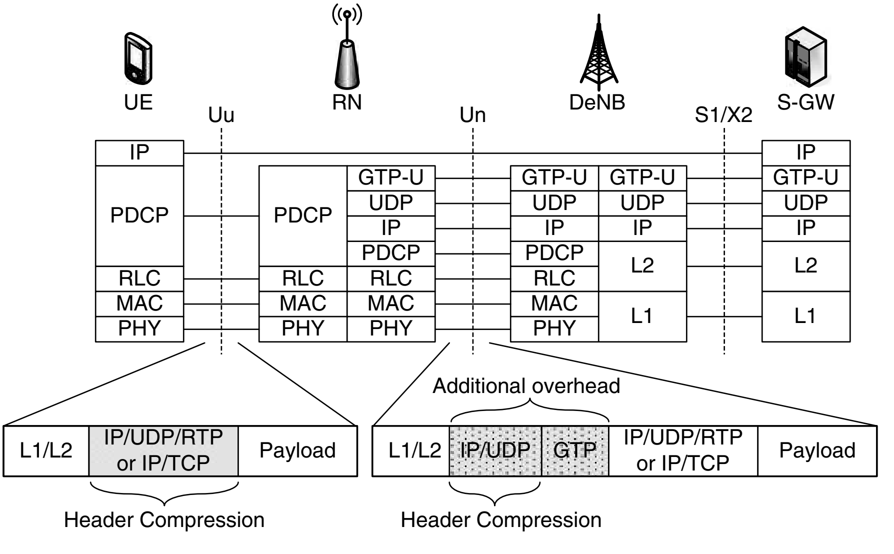

Figure 13.14 Header overhead in the Un interface

Without a flow control mechanism, the RN may discard the data received from the DeNB if the Uu link is congested.

The introduction of a flow control mechanism was considered positively during the course of the Relay discussions, but finally it was decided not to introduce such a mechanism in LTE Release 10. Since the deployment scenario is limited to coverage extension in rural areas, it is generally assumed that a Uu link is not congested. This is the reason why flow control was not considered essential for Release 10.

The final feature is additional header compression in the Un interface. As can be seen in Figure 13.14, the Un interface has additional IP/UDP/GTP layers compared with the Uu interface. Consequently, each Un packet contains additional IP/UDP/GTP headers in addition to IP/UDP/RTP or IP/TCP headers, as shown in Figure 13.14.

The problem with the current header compression mechanism in the PDCP is that only the outer IP headers are compressed (see Section 4.2). That is, the GTP and inner IP/UDP/GTP or IP/TCP headers are transmitted without header compression. This causes a big header overhead in the Un interface, so many mechanisms were discussed to reduce this header overhead.

However, as for flow control, additional header compression was not considered essential for Release 10. Since the deployment scenario is limited to coverage extension in rural areas, it is generally assumed that not many packets are transmitted through the Un interface, and thus the header overhead caused by the Un packets should not degrade the overall system performance.

Not only was the introduction of the above new features avoided, but some of the existing features in Uu radio protocols were also made not applicable to RN operation in Release 10. For example, VoIP-supporting features such as Semi-Persistent Scheduling (see Section 6.4.2) and TTI bundling (see Section 6.8.3), and MBMS transmission (see Section 11.3) are not supported by an RN. These simplifications are also helpful in terms of speedy implementation and deployment of an RN.

References

1. Iwamura, M., Takahashi, H., and Nagata, S. (2010) Relay technology in LTE-Advanced. *NTT DoCoMo Technical Journal*, **12** (2), 29–36.
2. 3GPP Technical Specification 36.300, "E-UTRA and E-UTRAN Overall Description Stage 2 (Release 10)", www.3gpp.org, accessed on 14 March 2012.
3. 3GPP Technical Report 36.912, "Feasibility study for Further Advancements for E-UTRA (LTE-Advanced) (Release 10)", www.3gpp.org, accessed on 14 March 2012.

14

Minimization of Driving Test (MDT)

So far, operators have spent enormous amounts of time and money to optimize network performance by collecting radio measurements and then analyzing them to derive the optimal parameters to apply to networks. The radio measurements are usually collected by test equipment running on cars driven by operators. These driving tests have to be performed extensively to collect radio measurements around the concerned areas. Based on the collected data, optimal parameters are derived carefully, with further impact analysis in the laboratory.

Once the optimal parameters have been selected, the parameters are applied to the networks and another set of driving tests is carried out to evaluate the impact of the parameter tuning. Such calibration may have to be repeated until the expected level of performance is achieved. A change in the radio environment may take place when a cell site is newly installed or even when new buildings are constructed. Deployment of heterogeneous networks, for example, pico/femto cells, drastically complicates the radio environment. To ensure a good quality of service to subscribers under such circumstances, more extensive and frequent driving tests are required, which increases the operational cost. With this in mind, operators have been seeking more efficient solutions.

The Minimization of Driving Test (MDT) is a standardized mechanism introduced in Release 10 to provide operators with network performance optimization tools in a cost-efficient manner. The key to MDT is employing user equipment of normal subscribers that exist any time and anywhere to collect radio measurements, as shown in Figure 14.1. UE involved in MDT is required to store measurement results and report them to the network. The network collects the measurement results from many UEs and utilizes them to fine-tune network parameters.

An overview and an overall description of MDT functionalities for LTE and UMTS are provided in [1]. Note that this section only focuses on MDT functionalities for LTE.

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

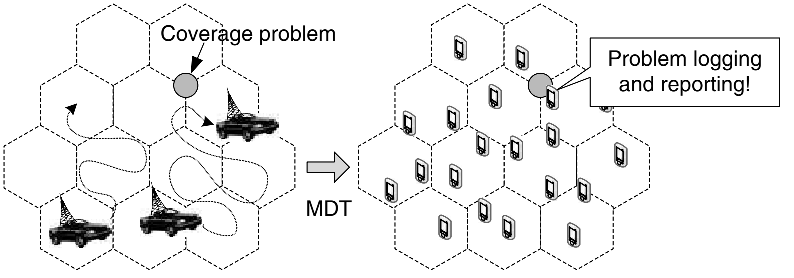

Figure 14.1 Motivation behind Minimization of Driving Test (MDT)

14.1 Architectural Framework

The MDT in Release 10 is built on control plane architecture. The design principle of MDT in Release 10 includes [2]:

- support for real-time and non-real-time measurement reporting;
- acceptable end-user implications, for example, battery consumption and memory requirements should be kept reasonable;
- correlation of the measurement results with time and location, indicating when and where the measurements were obtained, respectively.

Real-Time and Non-Real-Time Measurement Reporting

For the MDT, the UE should provide radio measurement results to the network. Measurement results can be obtained in RRC_IDLE and RRC_CONNECTED.

For measurement results obtained in RRC_IDLE, non-real-time reporting should be used because there is no connection between the UE and the network in RRC_IDLE. To enable non-real-time reporting, the concept of "logging" has been introduced. "Logging" means that the UE stores and accumulates the obtained measurements within its memory, and reports the logged measurements after the UE enters RRC_CONNECTED; that is, non-real-time reporting is used for measurements collected in RRC_IDLE. The procedures associated with measurement and logging in RRC_IDLE and reporting in RRC_CONNECTED are referred to as *Logged MDT*.

For measurement results obtained in RRC_CONNECTED, real-time reporting is already possible by applying the measurement report procedure defined in Section 3.9, where the measurement report is sent immediately upon satisfaction of the reporting criterion. The measurements and reporting for MDT reuse the measurement report procedure as much as possible. The procedures associated with measurement and reporting in RRC_CONNECTED for MDT are referred to as *Immediate MDT*.

UE Battery Consumption and Memory Requirement

The general design principle of the MDT is that the UE's MDT task should not sacrifice UE performance, because the MDT task is purely to assist network optimization efforts rather than providing instantaneous benefit to the UE. Such a design principle means that the UE battery life should not be impacted severely and the interruption of normal data services should be avoided or minimized. To meet the design principle, the complexity and battery consumption in relation to the MDT procedures were considered seriously.

Time and Location Information

To analyze coverage problems using collected MDT data, it is essential for the network to know when and where the measurements were taken by the UE.

Correlation between the time and the measurements for the obtained measurement results can be achieved by the UE and the network, as time information is already available in the UE and the network without additional measurement effort. For Logged MDT, the UE provides time information indicating the moment of measurement. For Immediate MDT, the UE does not provide time information, because the reporting is already in real time.

However, the acquisition of accurate location information at the moment of measurement is not easy because it raises some issues. The acquisition of user location information is strictly under the regulatory regime of the concerned authority, and hence, privacy issues related to user consent need to be considered carefully. Moreover, the frequent acquisition of detailed location information by the UE will entail noticeable battery consumption that is not acceptable to the user.

Taking these issues into account, 3GPP considered several options related to increased availability of detailed location information for the MDT measurement results. The options included forcing the UE to provide accurate location information or equivalent information that could be used by the network to derive the precise location where the measurements were taken, for example, by activating the LTE positioning protocol or Global Navigation Satellite System (GNSS) in the UE. However, the conclusion made for Release 10 was that UEs should not be required to perform additional measurements just to obtain accurate location information for MDT purposes, meaning that only available location information is included in the MDT measurement report. Note that 3GPP plans to study enhancement of this to increase the availability of detailed location information for MDT in Release 11.

Trace Function Related to MDT

The MDT is built upon the framework of the trace function, as defined in [3], with RRC signaling over the Uu interface, as defined in [4,5]. The trace function is one of the operators' tools for Operation, Administration, and Maintenance (OAM). The trace function provides operators with the capability to track and log UE activities such that it can enable the determination of the root cause of malfunctions at the UE. The traced data are collected in a network node, called the Trace Collection Entity (TCE). The operator uses the data collected in the TCE for analysis and evaluation. The trace function used for the MDT includes signaling-based trace functions and management-based trace functions. A signaling-based trace function is used to activate the MDT task for a specific UE, while a management-based trace function is used to activate the MDT task without specifying any particular UE.

From the network point of view, an MDT session consists of collecting user consent to participating in the MDT task, initiating the MDT task with UE selection, and collecting MDT measurements. Before going into the details of an MDT session, we will first examine several use cases for the MDT.

14.1.1 Use Cases

A number of parameters are subject to optimization, such as tuning of transmission power of data/control/pilot channels, antenna location/tilt, mobility parameters, and so on. In the study phase of the MDT, the following use cases were considered, as specified in [2]:

- **Coverage optimization:** Service coverage is the fundamental basis for service provisioning. Since the lack of coverage immediately degrades the quality of service, subscribers immediately notice the problem and express dissatisfaction. Hence, it is very important for the MDT to assist in coverage optimization for operators. Downlink coverage and uplink coverage can be considered separately.
- **Mobility optimization:** Mobility is essential for mobile services. With the MDT, the occurrence of handover failure due to too-early handover, too-late handover, or the selection of an incorrect target cell can be passed on to the network with information that can assist the network in deriving the cause of failure.
- **Capacity optimization:** Distribution of traffic over the service area is important for user QoS and operational management of the network. The MDT can assist the network in detecting a location where traffic engineering, such as installing a new eNB, is required.
- **Parameterization for common channels:** The configuration of the broadcast channel, the paging channel, and the random access channel directly affects network and user performance. The MDT can be used to help the network to detect common channel coverage or to identify performance problems related to common channels, for example, connection setup delay.
- **QoS verification:** Since the end goal of operators is to offer a better QoS to users, the MDT is suitable for providing user QoS verification data, for example, throughput or error rate.

Among these use cases, MDT Release 10 focused on the coverage optimization use case. The QoS verification use case is under the scope of MDT Release 11.

14.1.2 Initiation of the MDT Task with UE Selection

The MDT task involves the collection of measurement data, which include radio measurements, user location, and time information corresponding to the moment of measurement. The nature of collecting measurements may raise serious privacy concerns if a UE is forced to take part in the MDT task without user consent. To avoid privacy concerns, the MDT task should be initiated only for UEs that have provided user consent to participate in the MDT task. In addition to user consent, the MDT functions ensure a sufficient level of anonymization for collected data.

The AS layer does not provide any mechanism for the UE to give/revoke or reject consent to participation in MDT tasks. Instead, the upper layers are responsible for ensuring that the MDT task is initiated for UEs from which user consent has already been granted. User consent is collected by using the operator's customer care process. The user consent information resides in the HSS and is delivered to the MME and possibly to the eNB as part of the user subscription information and the UE context, respectively.

If a signaling-based trace function is used, the MDT task is configured toward a specific UE based on, for example, IMSI, IMEI-SV, and so on. This is called "signaling-based MDT". The selection of UEs for signaling-based MDT can be made by the core network. If a management-based trace function is used, the eNB receives a command from the core network for trace session activation for the MDT task without specifying the UEs for MDT involvement. This is called "management-based MDT". The eNB receiving a management-based trace session activation for the MDT selects the UEs to perform the MDT task. The management-based trace session activation for the MDT can include area restriction for the MDT. For UE selection, the eNB jointly considers the area restriction and UE capability related to the MDT. The eNB also considers user consent information received from the MME. After selecting UE(s) for the MDT task, the eNB configures the selected UE(s) to perform the MDT task via RRC procedures, using the parameters indicated in the trace session activation command.

14.1.3 Collection of MDT Measurement Results

If a UE reports MDT measurement results to the eNB, they are delivered to the Trace Collection Entity (TCE). The eNB receiving the MDT measurement report is responsible for routing the MDT measurement results to the correct TCE.

In the case of Logged MDT, the MDT measurement report sent by the UE includes the TCE ID that is translated into the TCE IP by the eNB for routing purposes. In the case of Immediate MDT, the eNB knows the associated TCE IP for the concerned UE as part of the UE context. The TCE IP is used to identify the TCE to which the MDT measurement results received from the UE should be delivered. The MDT measurement results collected in the TCE are then utilized as raw materials for network performance optimization.

14.2 Logged MDT

Logged MDT is the procedure by which the UE performs logging of measurement results and reporting of the logged measurement results. By the definition of logging, the UE stores and accumulates measurement results within its memory. The logged measurement results are reported afterwards when the reporting condition is met. Figure 14.2 shows the concept of Logged MDT. In this figure, the UE configured with Logged MDT performs logging of

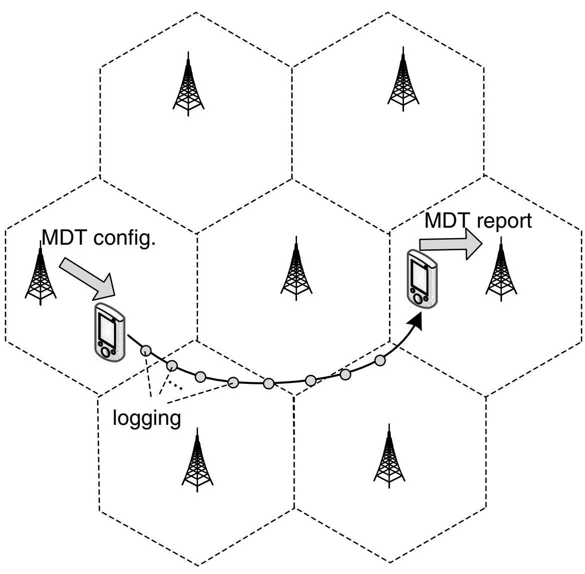

Figure 14.2 Concept of Logged MDT

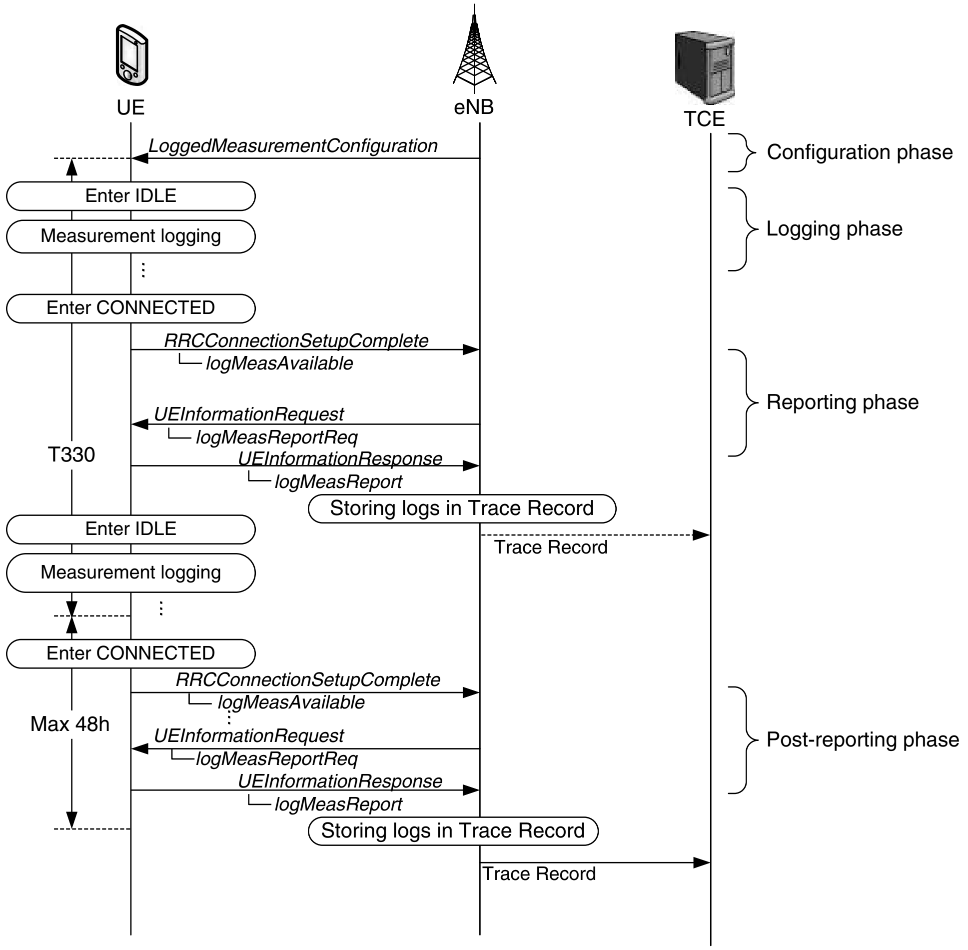

Figure 14.3 Example of Logged MDT signaling flow. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

measurement results across several cells. After the UE enters RRC_CONNECTED, it reports the logged measurement results to the serving cell.

Figure 14.3 illustrates an example of a Logged MDT session. The procedures related to the provisioning of user consent and user selection are not shown in this figure. The UE behaviors for Logged MDT consist of four phases: configuration, logging, reporting, and post-reporting. The details of each phase are described in the following sections.

As shown in Figure 14.3, the MDT task is configured to the UE by an RRC message (configuration phase). If the UE goes into RRC_IDLE, it initiates logging of measurement results while the logging duration timer (T330) is running (logging phase). When the UE enters RRC_CONNECTED with logged measurement results available to report, it may be requested to report the logged measurement results (reporting phase). The logging phase and the reporting phase can occur multiple times prior to expiry of the logging duration timer. The post-reporting phase, as shown by the dotted line, is initiated just in case the UE still has logged measurement results that have not been reported after the expiry of the logging duration timer (post-reporting phase).

The UE informs the network of support for Logged MDT by including *loggedMeasurementsIdle* in the UE capability. The UE also includes *standaloneGNSS-Location* in the UE capability, indicating support for Global Navigation Satellite System (GNSS) positioning.

For a UE supporting Logged MDT, at least 64 Kbytes of memory should be reserved for storing logged measurement results. 64 Kbytes is considered to be the minimum size for a UE to be capable of storing logged measurement results for 3 hours. The duration of 3 hours was determined based on the assumption that a UE will not stay in RRC_IDLE longer than 3 hours due to Tracking Area Update procedures.

Note that when the UE is switched off, measurement configuration and logged measurement results related to the Logged MDT are discarded.

14.2.1 Configuration Phase

A UE initiates the task of Logged MDT upon receipt of a measurement configuration for Logged MDT. The *LoggedMeasurementConfiguration* is used to configure the UE with Logged MDT, as shown in Figure 14.4. The configuration message includes several parameters that define an MDT task for a UE. The parameters are described in Table 14.1.

The parameters for a Logged MDT configuration consist of logging parameters, time information, trace information, and MDT area configuration:

- *Logging Parameters*
  - The *loggingDuration* defines the duration for which the UE is required to perform logging of measurement results. T330 in [4] corresponds to the timer associated with this logging duration. Upon expiry of T330, the UE stops logging and releases the Logged MDT configuration but stores the logged measurement results.
  - The *loggingInterval* defines the periodicity of logging of the measurement results.
- *Reference Time*
  - The *absoluteTimeInfo* is the reference time used by the UE when tagging a time stamp onto the measurement results in the logging phase. The time stamp records the elapsed time since the reference time.
- *Trace Information*
  - The *traceReference* is a Trace Reference, as defined in [3].
  - The *traceRecordingSessionRef* is the Trace Recording Session Reference, as defined in [3].

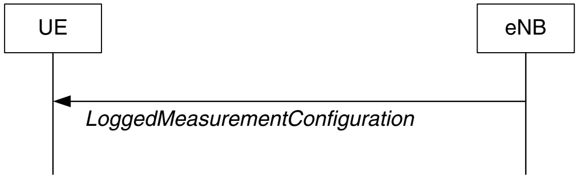

Figure 14.4 Logged MDT configuration

Table 14.1 Parameters in Logged MDT configuration

Parameter Set	Field Name in [4]	Description	Value Range or Data Format	Presence
Logging Parameters	*loggingDuration (associated with T330)*	The duration for which the UE is required to perform logging of measurements	$10 \sim 120$ min	Mandatory
	loggingInterval	Periodicity for logging of measurements	$1280 \sim 61440$ ms	Mandatory
Time	*absoluteTimeInfo*	Reference time to which the UE logs the time difference	YY-MM-DD HH: MM:SS using BCD encoding	Mandatory
Trace Info	*traceReference*	Trace Reference, as defined in [3]	(see left)	Mandatory
	traceRecordingSessionRef	Parameter Trace Recording Session Reference, as defined in [3]	(see left)	Mandatory
	tce-Id	Parameter Trace Collection Entity ID, as defined in [3]	(see left)	Mandatory
Area	*areaConfiguration*	The area for which the UE is requested to perform measurement logging and reporting	List of cells or list of tracking areas	Optional

  - The *tce-Id* is the Trace Collection Entity ID, as defined in [3]. The TCE ID is given to the UE instead of the TCE IP in order to avoid a security risk that could otherwise be encountered. The mapping of TCE IP to TCE ID is performed by the eNB.
- *Area Configuration*
  - The *areaConfiguration* defines the area for which the UE is requested to perform logging and reporting.

14.2.2 Logging Phase

When Logged MDT is configured, the UE needs to perform logging of measurement results. The UE logs measurement results that have been obtained during measurements for cell reselection. The UE does not perform additional measurements other than those measurements.

The UE considers several conditions to decide whether to perform logging. The UE performs logging if and only if all the following conditions are met:

- **Logging duration condition:** The logging duration timer (T330) is running.
- **RAT condition:** The UE is camped on an E-UTRA cell.

- **RRC_IDLE state condition:** The UE is in *CampedNormally* (i.e., no logging in *AnyCellSelection* or *CampedOnAnyCell* state).
- **PLMN condition:** The registered PLMN on the current cell is the same as the registered PLMN of the cell from which the UE received the Logged MDT configuration.
- **Area condition:** The UE is currently camped on a cell belonging to the concerned area. This condition is applicable only when *areaConfiguration* was configured.

If any of the conditions above is not satisfied, logging should be suspended. For example, while the UE is in RRC_CONNECTED or the UE is camping on a non-E-UTRA cell, the UE suspends logging. Note that while logging is suspended, logged measurement results are not discarded until they are reported or the discarding condition is met.

For simplicity, only periodical logging for Logged MDT is supported in Release 10; that is, event-based logging is not supported. The periodicity of logging is defined by the *loggingInterval*.

While a UE performs logging, the logged measurement results are stored in the *LogMeasInfoList*, as defined in [4]. The *LogMeasInfoList* is a list of log entries. Each log entry corresponds to *LogMeasInfo*. The contents of *LogMeasInfo* are listed in Table 14.2.

Each log entry consists of serving cell information, neighboring cell information, time information, and location information. The neighboring cell and location information are included only if available.

Table 14.2 Parameters of each log entry

Parameter Set	Field Name in [4]	Description	Presence
Serving cell identity	*servCellIdentity*	global cell identity of the serving cell	Mandatory
Measured results of serving cell	*rsrpResult*	Measured RSRP of the serving cell	Mandatory
	rsrqResult	Measured RSRQ of the serving cell	Mandatory
Measured results of neighboring cell	*measResultListEUTRA*	Measured results of E-UTRA cells	Optional
	measResultListUTRA	Measured results of UTRA cells	Optional
	measResultListGERAN	Measured results of GERAN cells	Optional
	measResultListCDMA2000	Measured results of CDMA2000 cells	Optional
Time stamp	*relativeTimeStamp*	The moment of logging measurement results, calculated as {current time minus *absoluteTimeStamp*} in seconds	Mandatory
Location information	*locationInfo*	Detailed location information at the moment of logging	Optional

Serving cell information consists of the Evolved Cell Global Identifier (ECGI = PLMN ID + cell ID within the PLMN) and measurement results for the serving cell in terms of RSRP and RSRQ.

The neighboring cell information spans the measurement results of E-UTRA cells, UTRA cells, GERAN cells, and CDMA2000 cells, depending on the UE capabilities. The measurement results of neighboring cells are logged in decreasing order of the ranking criterion used for cell reselection. Within each log entry, a maximum of six intra-frequency neighboring cells and three inter-frequency neighboring cells per frequency can be included for E-UTRAN/UTRAN/CDMA2000 cells, and three GERAN cells per frequency/set of frequencies.

A time stamp is included in each log entry. To reduce the information size, the UE includes relative time with the *absoluteTimeStamp* received in the Logged MDT configuration, calculated as the current time minus the *absoluteTimeStamp*.

Location information indicating the position of the measurements is also included, if available. The location information consists of mandatory *locationCoordinates* and optional *horizontalVelocity* information.

Memory reserved for MDT storage may become full of logged measurement results before the expiry of the logging duration timer. In this case, the UE behaves in the same way as it would upon expiry of the logging duration timer; that is, the UE stops logging, releases the configuration of Logged MDT, and enters the post-reporting phase.

14.2.3 Reporting Phase

Logged measurement results are reported to the network via an RRC message called the *UEInformationResponse*. When a *UEInformationResponse* message contains logged measurement results, it is sent over SRB2, which is normally used for carrying lower priority messages, as the reporting of logged measurement results is not urgent in comparison to other RRC messages that are normally sent over SRB1.

The basic reporting mechanism is "on-demand" reporting; that is, the UE is allowed to report logged measurement results only when the network requests the UE to report. Use of this on-demand mechanism is reasonable because the UE is not aware of whether the current serving cell can decode or is interested in receiving the logged measurement results. Even if the serving cell can decode the logged measurement results, the eNB may want to defer receiving them if there is heavy traffic in the cell.

For the on-demand reporting mechanism to work, the eNB needs to know which UE is subject to MDT reporting; that is, whether or not the UE has logged measurement results to report. For this purpose, a mechanism has been introduced whereby the UE indicates to the eNB the presence of logged measurement results to report. Upon receipt of this indication, the eNB can decide whether to retrieve the logged measurement results from the UE.

14.2.3.1 Availability of Logged Measurement Results

To enable the eNB to decide whether to invoke the log retrieval procedure, a UE with logged measurement results indicates the availability of them to the eNB if all of the following conditions are met:

- **Log availability condition:** The UE has logged measurement results to report.
- **RAT condition:** The UE is camped on an E-UTRA cell.

- **PLMN condition:** The registered PLMN on the current cell is the same as the registered PLMN of the cell from which the UE received the Logged MDT configuration.

Given that all the conditions are satisfied, the UE indicates the availability of logged measurement results. The indication of availability can be made by including *logMeasAvailable* in the following RRC messages:

- *RRCConnectionSetupComplete* message during the RRC Connection Setup procedure;
- *RRCConnectionReestablishmentComplete* message during the RRC Connection Re-establishment procedure;
- *RRCConnectionReconfigurationComplete* message during the handover procedure.

The UE indicates availability of logged measurement results at every RRC state transition, which enables a flexible log retrieval policy from the eNB side.

14.2.3.2 Log Retrieval

Log retrieval over the radio interface is performed by means of the UE Information procedure between the eNB and the UE, as illustrated in Figure 14.5. Upon receipt of a log availability indication, the eNB may request an MDT measurement report from the UE by sending a *UEInformationRequest* message including *logMeasReportReq*, which indicates a request for logged measurement results. The UE responds with a *UEInformationResponse* message including *logMeasReport*, which contains the logged measurement results.

When receiving a request for logged measurement results, the UE checks the following conditions:

- **Log availability condition:** The UE has logged measurement results to report.
- **RAT condition:** The UE is camped on an E-UTRA cell.
- **PLMN condition:** The registered PLMN on the current cell is the same as the registered PLMN of the cell from which the UE received the Logged MDT configuration.

If all of the conditions above are satisfied, the UE constructs a *UEInformationResponse* message by setting the parameters shown in Table 14.3 for the *logMeasReport*. If any of the conditions above is not satisfied, the UE sends a null *UEInformationResponse* (i.e., including nothing in the message) via SRB1. If the *UEInformationResponse* including logged

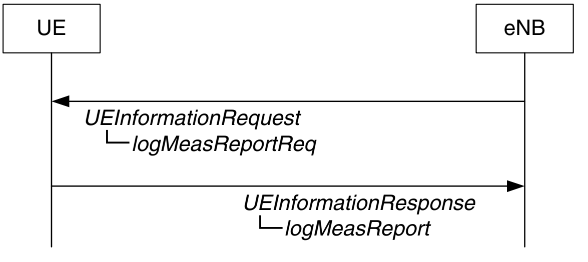

Figure 14.5 UE Information procedure for log retrieval

Table 14.3 Parameters in *UEInformationResponse* for Logged MDT

Parameter Set	Field Name in [4]	Description	Presence
Logs	*logMeasInfoList*	The list of log entries	Mandatory
Time	*absoluteTimeStamp*	Reference time that was received by the eNB at the configuration of Logged MDT	Mandatory
Trace	*traceReference*	Trace Reference, as defined in [3]	Mandatory
	traceRecordingSessionRef	Parameter Trace Recording Session Reference, as defined in [3]	Mandatory
	tce-Id	Parameter Trace Collection Entity ID, as defined in [3]	Mandatory
Segmented log delivery	*logMeasAvailable*	Indication of additional logs stored in the UE	Optional

measurement results is reported successfully, the UE deletes the logged measurement results from its memory.

When constructing the *logMeasInfoList*, the UE places each log entry in chronological order from when the log entry was generated. The size of log entries included in the *logMeasInfoList* should not exceed the maximum size of a PDCP SDU (8188 bytes), which corresponds to the maximum size of an RRC PDU. With this size restriction, it is up to UE implementation to determine how many log entries are included in the *UEInformationResponse*.

14.2.3.3 Multiple Log Retrievals

The logged measurement results stored in the memory for the MDT may exceed the maximum size of a PDCP SDU, because the memory size reserved for storing logged measurement results is 64 Kbytes, which is eight times larger than the maximum size of a PDCP SDU. To handle this case without losing the logged measurement results, it should be possible for a UE to deliver the available logged measurement results in multiple segments via reporting of multiple *UEInformationResponse* messages.

Segmented log delivery via multiple *UEInformationResponse* messages is realized by allowing the UE to indicate log availability (*logMeasAvailable*) in the *UEInformationResponse* if there are still logged measurement results to be retrieved other than those included in the *UEInformationResponse*. The log availability indication in the *UEInformationResponse* makes the eNB aware of the necessity of another log retrieval. Figure 14.6 shows an example of multiple log retrievals where the UE includes the log availability indication in the first *UEInformationResponse*. Since transport of a second *UEInformationResponse* empties the memory reserved for the MDT in the UE in this example, the log availability is not included in the second *UEInformationResponse*.

14.2.3.4 Post-Reporting Phase

It may be that, upon expiry of the logging duration timer, the UE still has logged measurement results that have not yet been retrieved by the network. Expiry of the timer may happen at any time, for example, when log retrieval is not possible due to being in RRC_IDLE,

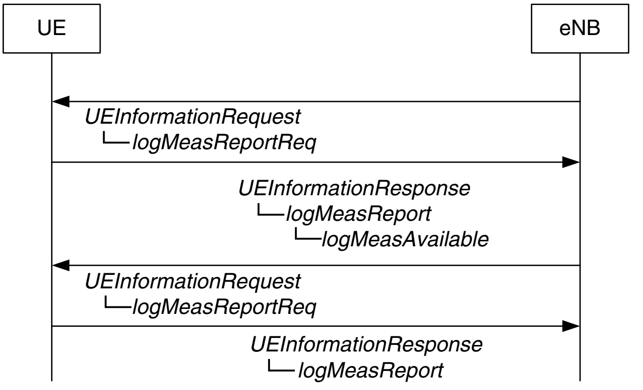

Figure 14.6 Use of multiple UE Information procedures for log retrieval

camping on a non-E-UTRA cell, or even when the network is overloaded with normal user traffic. It would be quite undesirable for expiry of the duration timer always to result in the discarding of logged measurement results that have not been reported to the network.

To provide the network with the chance to retrieve logged measurement results even after the expiry of the logging duration timer, the UE needs to keep the logged measurement results that have not been reported up to a maximum of 48 hours after the expiry of the logging duration timer. This length of time is considered sufficient for both the UE and the network to ensure a successful retrieval of residual logged measurement results. During this post-reporting phase, the UE performs normal log reporting procedures as described above; that is, indicating log availability, if appropriate, and reporting logged measurement results upon a request by the network.

14.3 Immediate MDT

Immediate MDT is the procedure by which the UE performs measurements and reports the measurement results as soon as the reporting condition is met. All behaviors regarding Immediate MDT are applicable for a UE in RRC_CONNECTED.

The general principle of Immediate MDT is that existing RRM measurements and reporting mechanisms, as defined in [4], are reused as much as possible. The only enhancement required to support Release 10 Immediate MDT is that the UE can be configured to tag accurate location information, if available, onto the measurement results reported to the network. Accurate location information tagged onto the measurement results assists the network in correlating the reported measurement results with the location of the measurements.

The UE does not include time information in the measurement report it sends. Instead, when the eNB receives a measurement report for Immediate MDT, the eNB attaches time information to the trace record containing the measurement results and location information reported by the UE. This is possible because Immediate MDT reporting is basically done in real time.

Support for Immediate MDT is mandatory for all UEs, since the procedure for Immediate MDT is based solely on existing RRM measurement procedures.

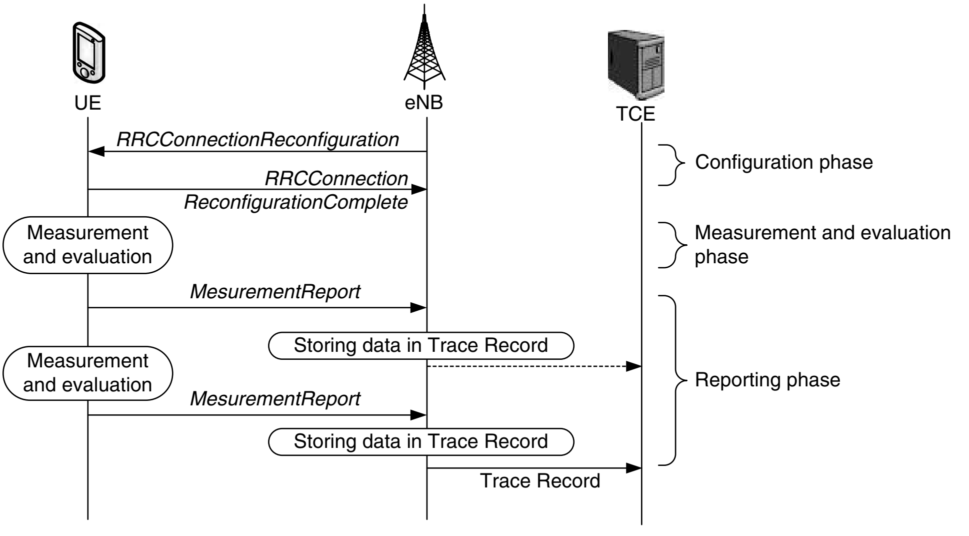

Figure 14.7 Example of Immediate MDT signaling flow

Figure 14.7 shows an example of Immediate MDT procedures, focusing on RRC signaling. Note that the procedures related to the provisioning of user consent and user selection for the MDT task are not shown in this figure. The Immediate MDT procedures consist of three phases: configuration, measurement and evaluation, and reporting. The details of each phase are described in the following sections.

14.3.1 Configuration Phase

Normal RRM measurement configuration is used for Immediate MDT configuration, as shown in Figure 14.8. Note that periodic measurement reporting and event A2-triggered measurement reporting are applicable for reporting conditions of Immediate MDT in Release 10. The only difference to normal RRM measurements is that for Immediate MDT, the eNB may request the UE to include accurate location information, if available, in the RRM measurement report by including *includeLocationInfo* in the reporting configuration of the measurement configuration.

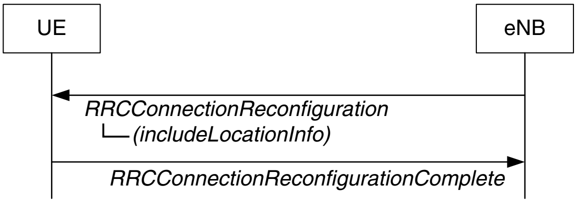

Figure 14.8 Configuration of Immediate MDT

Table 14.4 Parameters in the measurement report for Immediate MDT

Parameter Set	Field Name in [4]	Description	Presence
RRM measurement results	Several fields according to [4]	(see left)	According to [4]
Location Information	*ellipsoid-Point*	Parameter *Ellipsoid-Point* defined in [6]	If *includeLocation-Info* is configured and detailed location information is available
	ellipsoidPointWithAltitude	Parameter *EllipsoidPoint WithAltitude* defined in [6]	
	gnss-TOD-msec	Parameter *Gnss-TOD-msec* defined in [6]	

14.3.2 Measurement and Evaluation Phase

The UE performs measurements according to the measurement configuration for normal RRM measurements; that is, no difference between normal RRM measurements and those for Immediate MDT exists. No additional event for evaluation of reporting conditions other than existing events is defined for Immediate MDT in Release 10.

14.3.3 Reporting Phase

The reporting behaviors are the same as those for normal RRM measurements except that if the network has requested the UE to include accurate location information, the UE includes the accurate location information, if available, in the *MeasurementReport*.

The parameters that can be included in *MeasurementReport* for Immediate MDT are listed in Table 14.4.

14.3.4 MDT Context Transfer

At intra-LTE handover, it is a general principle that the UE context is forwarded from source cell to target cell, and the target cell is in charge of preparing for valid configurations applicable to the UE. However, the handling of Immediate MDT configuration at handover is different depending on how the MDT task was configured for the UE.

If Immediate MDT was configured by a signaling-based trace function, the source eNB forwards the Immediate MDT configuration to the target eNB such that the MDT task of the UE can continue after mobility. However, if the MDT task was configured by a management-based trace function, the Immediate MDT configuration is not forwarded.

The MDT configuration is forwarded during handover even for the case where area restriction was configured and the target eNB does not belong to the concerned area. In such a case, the target eNB is responsible for maintaining a valid configuration for the UE, for example, by releasing part of the configuration that is not applicable in the area.

References

1. 3GPP Technical Specification 37.320, “Universal Terrestrial Radio Access (UTRA) and Evolved Universal Terrestrial Radio Access (E-UTRA); Radio measurement collection for Minimization of Drive Tests (MDT); Overall description; Stage 2 (Release 10)”, www.3gpp.org, accessed on 14 March 2012.
2. 3GPP Technical Report 36.805, “Evolved Universal Terrestrial Radio Access (E-UTRA); Study on minimization of drive-tests in next generation networks (Release 9)”, www.3gpp.org, accessed on 14 March 2012.
3. 3GPP Technical Specification 32.422, “Telecommunication management; Subscriber and equipment trace; Trace control and configuration management (Release 10)”, www.3gpp.org, accessed on 14 March 2012.
4. 3GPP Technical Specification 36.331, “Radio Resource Control (RRC); Protocol Specification (Release 10)”, www.3gpp.org, accessed on 14 March 2012.
5. 3GPP Technical Specification 36.304, “User Equipment (UE) procedures in idle mode (Release 10)”, www.3gpp.org, accessed on 14 March 2012.
6. 3GPP Technical Specification 36.355, “LTE Positioning Protocol (LPP) (Release 10)”, www.3gpp.org, accessed on 14 March 2012.

15

Enhanced Inter-Cell Interference Coordination (eICIC)

On an LTE network, the frequency resources of one cell can be reused by the adjacent neighboring cells to maximize efficiency of frequency spectrum usage, which means that the frequency reuse factor is typically 1. When such a complete frequency reuse operation is utilized, inter-cell interference cannot be avoided.

Figure 15.1 illustrates the inter-cell interference between adjacent cells in an OFDMA network [1], where the resource blocks allocated to the UE in one cell overlap with the resource blocks allocated to another UE in another cell in the time and frequency domains. The overlapping resource blocks generate mutual interference to both UEs.

Such inter-cell interference lowers the Signal to Interference plus Noise Ratio (SINR) associated with these resource blocks. As a result, the ratio of transmission failure increases and then the resource utilization is impacted badly. The impact of inter-cell interference would be more severe if a UE were located at a cell edge where the measured level of the serving cell becomes almost comparable to that of adjacent neighboring cells. Unless a suitable mechanism is applied to manage such inter-cell interference, target QoS cannot be guaranteed.

3GPP introduced the following Inter-Cell Interference Coordination (ICIC) mechanisms for management of such inter-cell interference in LTE:

- frequency domain ICIC in Releases 8/9;
- CA-based ICIC in Release 10;
- time domain ICIC in Release 10.

The following sections describe network deployment scenarios for which inter-cell interference is applicable and examine the ICIC mechanisms supported in LTE.

15.1 Heterogeneous Network Deployment

A homogeneous network means a network that has similar characteristics in terms of the architectural requirements, service coverage, and so on. In a homogeneous network,

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

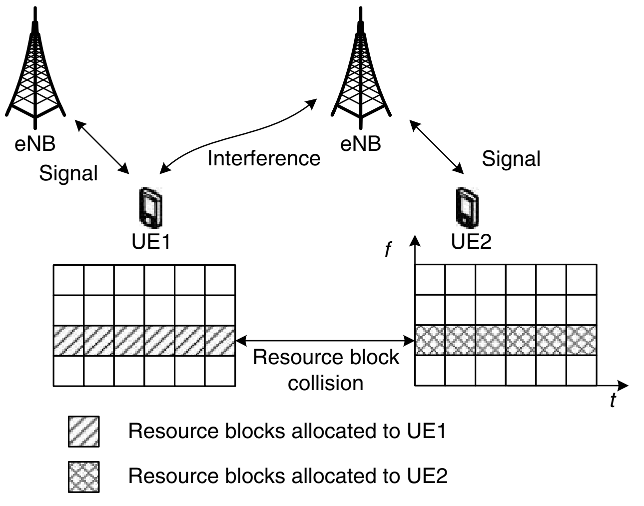

Figure 15.1 Inter-cell interference by radio resource block collision

the macro eNBs normally provide wide coverage by using high transmission power, and they are interconnected to neighboring eNBs as well as to core network nodes. The locations of cell sites in such a network are planned carefully so that mutual interference among these cells can be minimized.

While homogeneous network deployment was considered primarily by 3GPP, it was identified that the user's demand for data traffic is not homogeneous throughout the service area. For example, in a certain urban area crowded with many users, the required capacity is much higher than in other areas. In addition, it is a global trend that the amount of mobile traffic is continuing to increase sharply due to the extensive popularity of smartphones.

Facing heterogeneous demand and the explosive growth of data traffic on the mobile network, operators realized that homogeneous network deployment is not always the best solution, because obtaining new macro cell sites is a non-trivial job and the operational costs of macro cells are very high. Furthermore, the acquisition of new frequency spectrum to accommodate more data traffic typically requires prohibitive expense. Therefore, the operators started to take a look at the benefits of deploying a low-cost heterogeneous network in which small cells are collocated on the macro cell's overlaid coverage. These small cells are utilized to boost the capacity in high-density traffic hotspots such as airports, shopping malls, and large office buildings. In addition, the small cells are used to provide service coverage with lower cost in areas hardly reachable by the macro cells, for example, indoors and underground areas.

A network utilizing deployment of small cell(s) collocated on the macro cell's overlaid coverage is often called a *heterogeneous network*. When it comes to small cells, we normally think of pico cells and femto cells, as shown in Figure 15.2.

A pico cell is accessible by all users and is interconnected to neighboring cells via an X2 interface like the macro cell. However, a pico cell provides much less coverage than a macro cell since the transmission power of a pico cell is only up to a few watts.

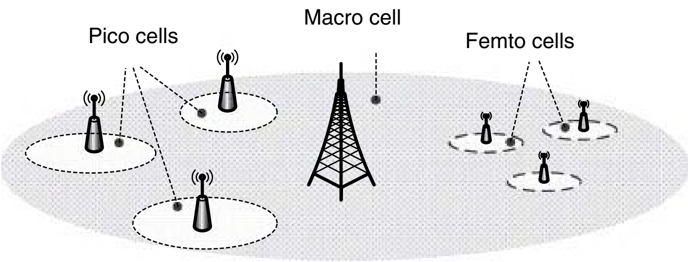

Figure 15.2 Heterogeneous network deployment

A femto cell is accessible only by Closed Subscriber Group (CSG) members. Note that when a femto cell is concerned in the context of enhanced ICIC, the femto cell normally refers to the CSG cell served by the HeNB. Since HeNBs are customer-premises pieces of equipment, an X2 interface may not be available (see Section 9.1.4). A femto cell provides even less coverage than a pico cell.

Table 15.1 describes the characteristics of heterogeneous cells in 3GPP.

15.1.1 Interference in Heterogeneous Networks

Although the deployment of heterogeneous networks was introduced to provide a low-cost solution to increasing network efficiency, it has brought with it a serious challenge to interference management in mobile network operations, which did not exist with a homogeneous network. Without managing interference carefully, the intended benefit of a heterogeneous network cannot be achieved.

Figure 15.3 illustrates one of the interference models in the heterogeneous network scenario, where a pico cell and a macro cell are deployed on the same frequency. The pico cell is used to offload the macro cell traffic. This scenario is often referred to as the macro–pico cell scenario. In this figure, the macro cell gives strong interference to the UE served by the pico cell, especially when the UE is positioned on the pico cell boundary. This UE is called a "pico UE". In this macro–pico cell scenario, the macro cell is an aggressor cell that causes interference to the pico UE and the pico cell is a victim cell that suffers from interference caused by the macro cell.

Table 15.1 Heterogeneous cells in 3GPP. Reproduced by permission of 3GPP, © 2010. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

Cell Type	Tx Power (dBm)	User Access	Interconnection with Neighboring Cells	Placement
Macro cell	46–49	Open to all users	X2	Outdoors and indoors
Pico cell	24–30	Open to all users	X2	Outdoors and indoors
Femto cell (CSG cell)	20	Open only to CSG members	No/limited X2	Normally indoors

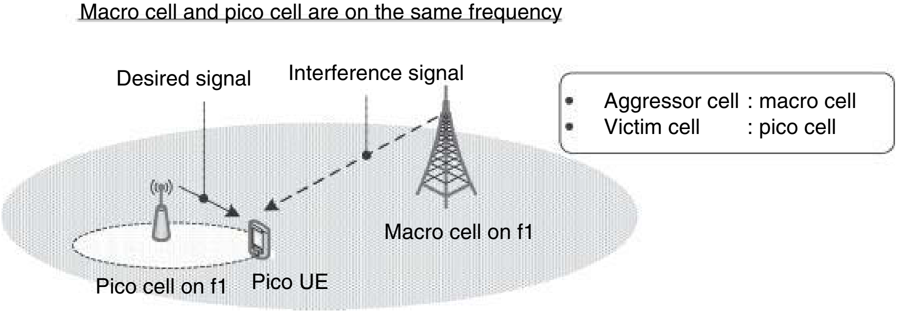

Figure 15.3 Macro–pico cell interference scenario

Figure 15.4 illustrates the other interference model in the heterogeneous network scenario, where a femto cell and a macro cell are deployed on the same frequency. This scenario is often called the macro–femto cell scenario. Since the femto cell is only accessible to CSG members, the coverage of the femto cell is seen as a coverage hole to a UE that is served by the macro cell and is not a CSG member of the femto cell (CSG cell). This UE is called a "macro UE". In this macro–femto cell scenario, the femto cell is an aggressor cell that causes interference to the macro UE and the macro cell is a victim cell that suffers from interference caused by the femto cell.

15.1.2 Limitation of Frequency Domain ICIC

In Releases 8/9, ICIC is based on frequency domain coordination, where one cell avoids using the resource blocks that can be used by neighboring cell(s). ICIC can be applied to the uplink and downlink, as illustrated in Figure 15.5.

For uplink ICIC, adjacent cells exchange uplink interference information by using a Load Information message over the X2 interface. The information exchanged can include a UL high interference indication and/or a UL interference overload indication.

A UL high interference indication provides the neighboring cells with information about the occurrence of uplink interference for each resource block. Upon receiving a UL high

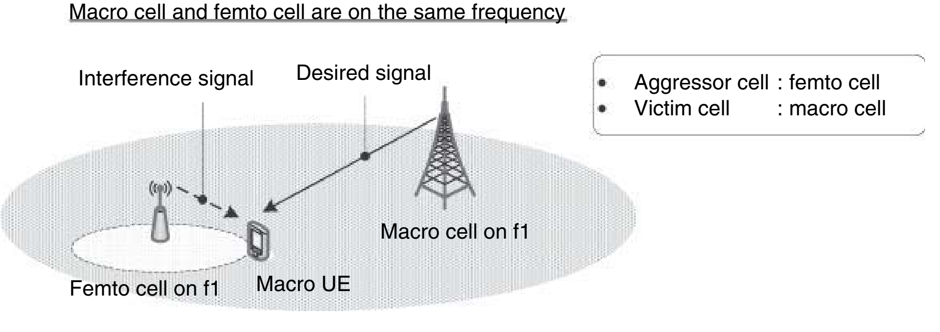

Figure 15.4 Macro–femto cell interference scenario

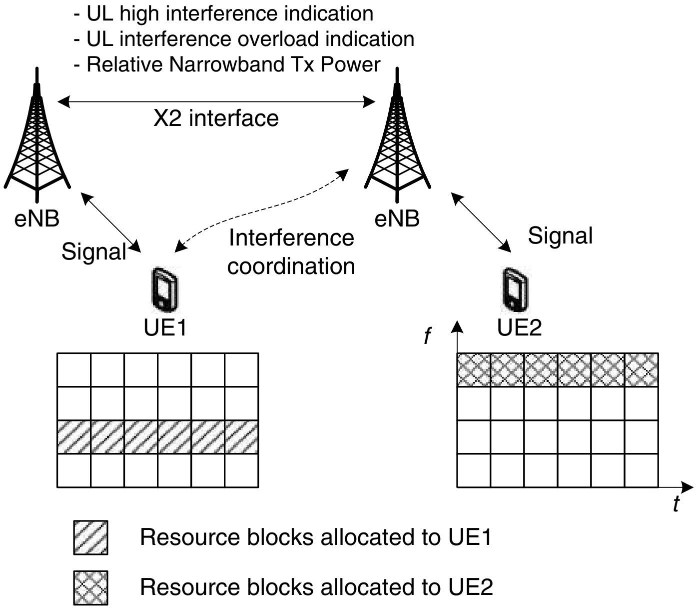

Figure 15.5 The frequency domain ICIC in Releases 8/9

interference indication, the neighboring cells may wish to avoid scheduling on the indicated resource blocks.

A UL interference overload indication provides information on the uplink interference level experienced in each resource block. A cell receiving a UL interference overload indication may reduce the interference generated on some of the indicated resource blocks.

For downlink ICIC, adjacent cells exchange potential downlink interference information, the Relative Narrowband Tx Power (RNTP), by using a Load Information message over the X2 interface. For each resource block, the RNTP indicates whether the transmission power of the resource block is greater than a certain threshold. Upon receiving this information, a neighboring cell may avoid scheduling on the resource blocks indicated as having high power transmission.

The frequency domain ICIC mechanism can be used to minimize inter-cell interference, particularly in cases of homogenous network deployment. Considering that the control channel is more robust than the traffic channel, the frequency domain ICIC mechanism focuses on resolving the mutual interference to/from the traffic channel, and does not provide any protection for the control channel.

However, protecting the control channel became a challenge when heterogeneous networks were being considered. For example, PDCCH reception performance is degraded severely, as shown in [2], which leads to both a decreased QoS and an increased rate of radio link failures. In addition, the simple extension of frequency domain ICIC, for example, PDCCH protection by applying the RNTP, is inappropriate because the PDCCH is already spread over the entire bandwidth.

Therefore, there was a demand to develop a new ICIC mechanism for management of control channel interference in a heterogeneous network environment.

In Release 10, two enhanced ICIC mechanisms were introduced for coordinating PDCCH interference: the CA-based ICIC mechanism and the time domain ICIC mechanism. Each mechanism is described in the subsequent sections.

15.2 CA-based ICIC

In Release 10, an eNB controlling multiple cells on multiple carrier frequencies may support CA. If CA is supported, the eNB can apply cross-carrier scheduling in such a way that the PDCCH from one serving cell of one carrier frequency schedules transmissions for another serving cell of another carrier frequency, as described in Section 12.2.4.

CA-based ICIC is applicable when the eNB controls both the aggressor cell and the victim cell. One use case is a scenario where the aggressor cell is the macro cell and the victim cell is the Remote Radio Head (RRH) cell (see Section 12.1.2). In such a scenario, the eNB can avoid using PDCCHs on the same carrier frequency by cross-carrier scheduling.

Figure 15.6 illustrates one example of the CA-based ICIC mechanism in the macro–pico cell scenario. In Figure 15.6, both the macro UE and the pico UE are configured with CA. For the macro UE that is close to the pico cell, the macro cell utilizes the PDCCH on f1 to schedule the data channels on both f1 and f2 by using cross-carrier scheduling. For the pico UE that is at the pico cell edge, the pico cell utilizes the PDCCH on f2 to schedule the data channels on both f1 and f2 by using cross-carrier scheduling.

Since the macro cell and the pico cell utilize PDCCHs on different frequencies, mutual interference on the PDCCH can be avoided. Note that data channel interference can be coordinated via the frequency domain ICIC mechanisms in Releases 8/9.

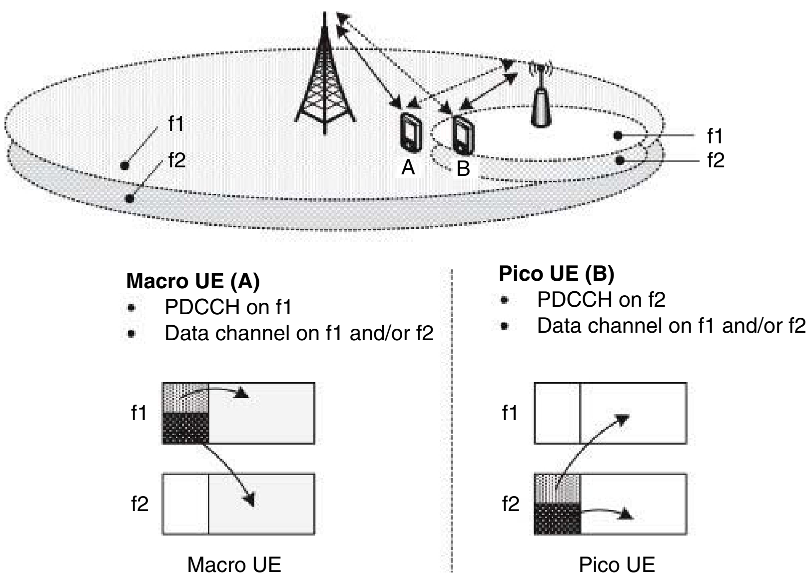

Figure 15.6 CA-based ICIC mechanism. Reproduced by permission of 3GPP, © 2010. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

15.3 Time Domain ICIC

The concept of time domain ICIC is that the subframes are partitioned into two sets so that the aggressor cell and the victim cell use different sets of subframes. Specifically, the aggressor cell mutes certain subframes by restraining any scheduling on those subframes to suppress generating interference to the victim cell. The victim cell utilizes the subframes muted by the aggressor cell for scheduling its UEs that are subject to strong interference from the aggressor cell. Since the interference level during a muted subframe is quite low, the UEs scheduled on the muted subframes in the victim cell can receive the PDCCH reliably. Figure 15.7 illustrates the concept of time domain ICIC.

During the muted subframes, the aggressor cell stops using the traffic channel but keeps broadcasting some essential signaling and information. This is why a muted subframe is called an Almost Blank Subframe (ABS) rather than just a blank subframe. The broadcasting of essential signaling and information during ABSs is to support backward compatibility; that is, to accommodate the UEs that cannot support time domain ICIC. The signaling and information sent on ABSs include the Primary Sync Signal (PSS), the Secondary Sync Signal (SSS), the Master Information Block (MIB), the System Information Block 1 (SIB1), and the Cell-specific Reference Signal (CRS).

The ABS pattern has a 40 ms periodicity (for FDD) to cover both the PHICH transmissions with 8 ms interval and PSS/SSS/MIB/SIB1 with 10 ms interval [3]. Within the ABS pattern, some subframes are defined as ABSs and the remainder of the subframes are defined as normal subframes.

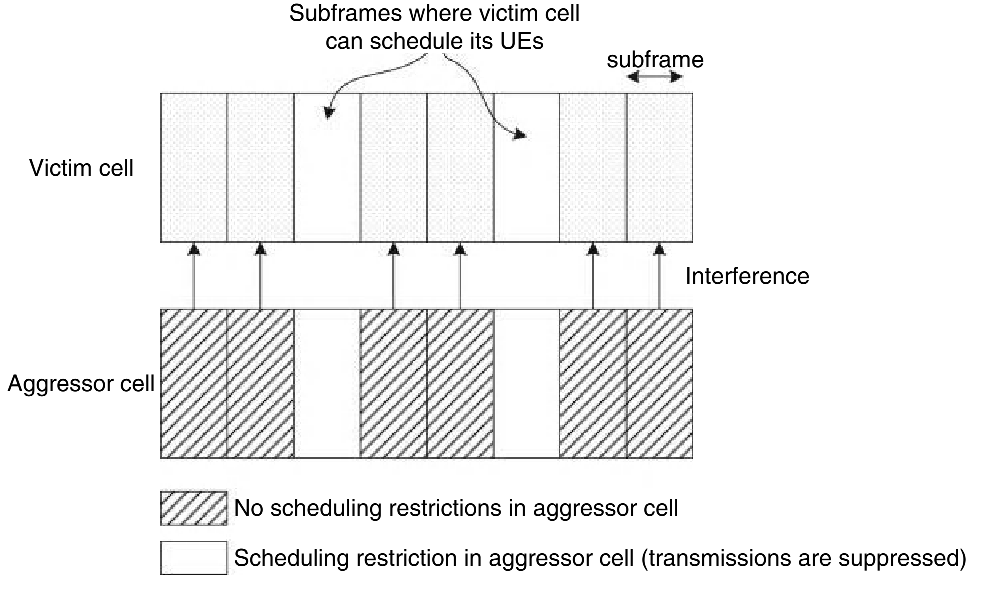

Figure 15.7 Concept of time domain ICIC

15.3.1 Restricted Measurements

When ABSs are configured and utilized by the network, the UEs in the interfered area of a victim cell can encounter serious SINR fluctuation from subframe to subframe. Such high fluctuation of SINR has a significant impact on UE measurements because, depending on whether the subframe taken for measurements is an ABS or a normal subframe, quite different measurement results can be acquired. The UE behaviors related to impacted measurement include Radio Link Monitoring (RLM), Radio Resource Management (RRM), and Channel Quality Indicator (CQI) measurements. To regulate the fluctuations in SINR during the UE measurements, the network can configure the UE to restrict measurements within certain subframes, for example, either ABSs only or normal subframes only. For restricted measurements, the measurement resource restriction pattern is signaled to the UE when time domain ICIC is applied. Different measurement resource restriction patterns are configured for each of the following measurement cases:

- RLM and RRM measurements of the serving cell;
- RRM measurements of neighboring cells;
- CQI measurements of the serving cell.

For restricted RLM and RRM measurements of the serving cell, the network can configure the UE with a measurement resource restriction pattern. The UE then performs restricted measurements using the resource restriction indicated by the restriction pattern for RLM and RRM of the serving cell.

For restricted RRM measurements of neighboring cells, the network can configure the UE with a measurement resource restriction pattern that is different from that used for restricted RLM and RRM measurements of the serving cell. The UE then performs restricted measurements using the resource restriction indicated by this restriction pattern. In Release 10, restricted measurements are applicable only to intra-frequency measurements. This limitation was introduced based on the assumption that inter-frequency restricted measurements are less critical than intra-frequency restricted measurements. This is because there is a workaround for inter-frequency restricted measurements such that the UE sequentially performs inter-frequency normal measurements, handover to the inter-frequency target cell, and intra-frequency restricted measurements on the new serving frequency.

When the measurement resource restriction pattern is configured for RRM measurements for neighboring cells, the list of each neighboring cell's physical cell identities is also provided to the UE. The UE applies the restricted measurements only for the listed cells and applies normal measurements for other cells. This selective restricted measurement means applying restricted measurements only to the neighboring cells for which interference is a problem, since it is unnecessary to apply restricted measurements to neighboring cells for which interference is not a problem. Figure 15.8 illustrates selective measurement restriction with a cell list.

For restricted CQI measurements, the network can configure two resource restriction patterns to the UE. The UE then sends two sets of CQI reports, each corresponding to the result of restricted measurement obtained by the respective resource restriction pattern.

An example of this configuration is where one pattern is configured for ABS and the other is configured for a normal subframe. With this configuration, the two sets of CQI reports can

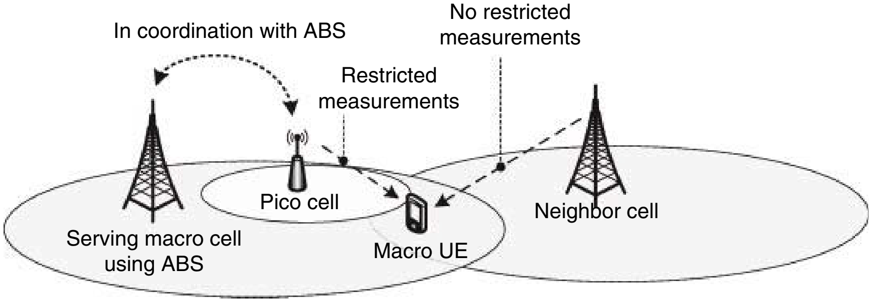

Figure 15.8 Selective RRM measurement resource restriction

assist the eNB to determine whether the UE needs to be scheduled only on ABSs or on normal subframes.

Another example is where one pattern is configured for ABS of the strongest aggressor cell and the other is configured for ABS of the second strongest aggressor. This configuration may be used when there are multiple aggressor cells, each having its own ABS on different subframes. In this case, the two sets of CQI reports can assist the eNB to ascertain the supportable Modulation and Coding Scheme (MCS) for transmissions to the UE by jointly considering these two CQI reports.

Table 15.2 summarizes the different types of restricted measurement configuration.

15.3.2 Macro–Pico Cell Scenario

When pico cells are deployed on the same frequency as the overlaid macro cell for offloading purposes, it is desirable to shift traffic load from the macro cell to the pico cells when necessary, for example, when the macro cell becomes overloaded. To expedite the offloading, the macro UEs in the vicinity of the pico cell had better perform handover to the pico cell and obtain services from the pico cell, even if the macro cell provides better radio conditions for those UEs. This kind of offloading results in so-called Cell Range Expansion (CRE) of the pico cell, since the coverage of the pico cell becomes virtually expanded. Figure 15.9 illustrates the CRE in the macro–pico cell scenario.

However, there is a big challenge associated with such offloading, in that the UEs in the CRE zone are highly vulnerable to interference because the UEs are close to the pico cell boundary. With strong interference from the pico cell, when the macro UE is positioned in the CRE zone, the reporting criterion to trigger the measurement report for handover to the pico cell is not likely to be met, and therefore handover to the pico cell

Table 15.2 Restricted measurement configuration

Purpose	Applicable Cell	Configuration Contents
RLM and RRM	Serving cell	One restriction pattern
RRM	Neighboring cell(s)	One restriction pattern with cell list
CQI	Serving cell	Two restriction patterns

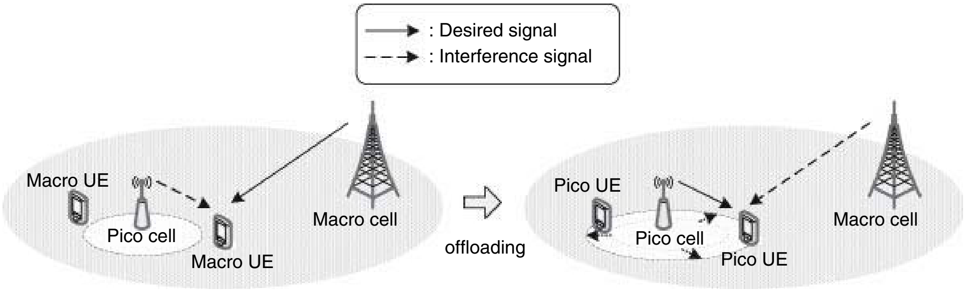

Figure 15.9 Cell range expansion of pico cell for offloading

cannot be initiated. Even if the UE is handed over successfully to the pico cell, the pico UE may suffer from serious degradation of QoS or even experience radio link failure sooner or later due to strong interference from the macro cell. According to [4], a higher rate of radio link failure in the pico cell is expected in the CRE zone than in the non-CRE zone.

In such a challenging interference environment, time domain ICIC can be used to enable the UE to trigger a measurement report for handover to the pico cell, and to help the UE survive after handover to the pico cell. The expected behaviors of the UE and the network are as follows:

- **Step 1:** The macro cell detects that the traffic demand in the cell may exceed its capacity and decides that it needs to offload some traffic to the overlaid pico cell by applying time domain ICIC.
- **Step 2:** The macro cell mutes some subframes by applying ABSs to provide low-interference subframes to the UEs in the CRE zone of the pico cell.
- **Step 3:** The pico cell is informed of the ABSs being configured in the macro cell.
- **Step 4:** The macro cell configures its macro UEs with restricted measurements to facilitate handover to the pico cell. The restricted measurement resource is desirably a subset of the ABSs being configured in the macro cell.
- **Step 5:** The macro UE then performs restricted measurements towards the pico cell. As a result of the restricted measurements, the macro UE can trigger a measurement report to the macro cell to indicate a chance of handover to the pico cell.
- **Step 6:** Upon receiving the measurement report, the macro cell triggers the handover procedure to the pico cell. Then the UE in the CRE zone can perform handover successfully to the pico cell, and the UE can now be served by the pico cell.
- **Step 7:** The pico cell configures the UE with the restricted measurements during or after the handover. The restricted measurement resource is desirably a subset of the ABSs from the macro cell.
- **Step 8:** The UE then applies restricted measurement for RLM and RRM measurement of the pico cell, by which the UE can avoid declaring radio link failure and giving an incorrect indication in the measurement report that the measured level of the pico cell is bad. The RRM measurement on neighboring cells (including the macro cell) and the CQI measurement on the pico cell can also be performed with restricted measurement.

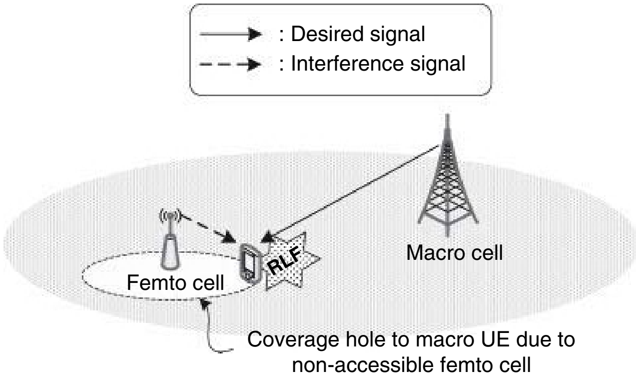

Figure 15.10 Interference from femto cell due to non-allowed access

15.3.3 Macro–Femto Cell Scenario

Femto cells deployed on the same frequency as a macro cell become a source of interference to the macro UEs that are not CSG members of those femto cells. If a macro UE is close to femto cell coverage to which the UE is not allowed to access, the UE will be interfered by the femto cell. The UE may suffer from severe degradation of QoS or even radio link failure due to the sharply increased level of interference. This case is illustrated in Figure 15.10.

In this scenario, time domain ICIC can be used to assist the UE to survive the strong interference from the non-accessible femto cell. The expected behaviors of the UE and the network are as follows:

- **Step 1:** The femto cell mutes some subframes by applying ABSs to provide low interference subframes to the macro UEs near the femto cell.
- **Step 2:** The macro cell configures its concerned UEs with restricted measurements for RLM/RRM/CQI measurements of the serving cell.
- **Step 3:** The macro UE then performs restricted measurements for the macro cell. Since the RLM measurements utilize the low interference subframes, the UE can avoid declaring radio link failure even when it is within the coverage of a femto cell that is not accessible to the UE. Applying restricted measurements for the RRM measurement and the CQI measurement on the macro cell would also give acceptable results.

15.3.4 Network Configuration

The ABS information configured by an aggressor cell should be known to a victim cell so that the victim cell can properly configure its UEs that are subject to strong interference from the aggressor cell.

In the macro–pico cell scenario, the macro cell determines which subframes to configure as ABSs, and then transfers the ABS information to the pico cell via a Load Information message on the X2 interface, as shown in Figure 15.11. The ABS information indicates the ABS pattern expressed in bitmap format. Each bit of the bitmap indicates whether the corresponding subframe is an ABS or a normal subframe.

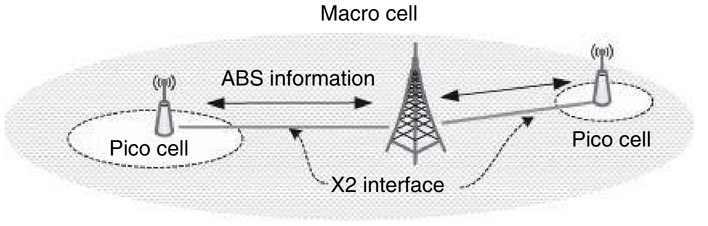

Figure 15.11 ABS information transfer from macro eNB to pico eNB.

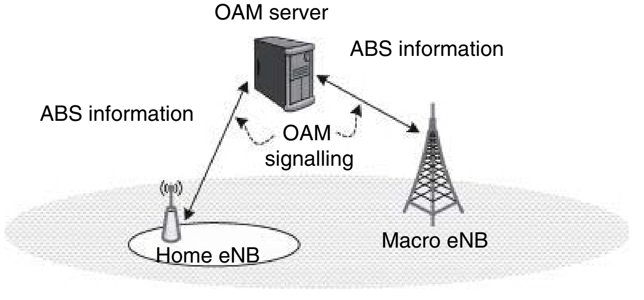

Figure 15.12 ABS information transfer between the macro eNB and the femto eNB.

The aggressor cell can request the victim cell to provide the usage status of ABS resources, in order to evaluate whether or not current ABS resources are suitable for the victim cell. The usage status of ABS resources indicates the percentage of used resource blocks within the ABSs. The usage status of ABS resources is transferred via a Resource Status Update message.

In the femto–macro cell scenario, an X2 interface between the macro cell and the femto cell is not available. Thus, the exchange of ABS information between the macro cell and the femto cell relies on Operations, Administration, and Maintenance (OAM) signaling, as shown in Figure 15.12. That is, the OAM server configures the HeNB with the proper ABSs, and provides the ABS information to the macro eNB.

References

1. Racz, A., Reider, N., and Fodor, G. (2008) On the impact of inter-cell interference in LTE. *IEEE Globecom*, **27** (1), 5436–5441.
2. Fujitsu, “Enhancing Cell-Edge Performance with PDCCH ICIC”, www.fujitsu.com, accessed on 14 March 2012.
3. NTT DoCoMo, “R1-105724: Views on eICIC Schemes for Rel-10”, 3GPP TSG RAN WG1, meeting 62bis, Xian, China, October 2010.
4. Qualcomm Inc., “R1-100700: Interference Conditions in Heterogeneous Networks”, 3GPP TSG RAN WG1, meeting 59bis, Valencia, Spain, January 2010.

16

Machine Type Communication (MTC)

3GPP systems including the GERAN, the UTRAN, and the E-UTRAN support Machine Type Communication (MTC), which is also known as Machine-to-Machine communication (M2M). The 3GPP system provides services which are used for end-to-end communications between the MTC application on the external network and the MTC device that is a UE supporting MTC.

Several use cases for MTC were identified in [1,2]. One of the use cases is smart metering. Metering devices could monitor electric power, gas, water, or heating to report information on energy consumption to service providers. Metering devices may autonomously report the information to the centralized node, or the centralized node may poll metering devices when reporting is needed. The reported information could be used to facilitate more efficient use of energy.

Another use case is road security. For instance, when a car accident occurs, an in-vehicle emergency call service could be initiated autonomously to report location information for the car accident to the emergency center. The reported information could be used to facilitate prompt assistance. MTC use cases for road security could include intelligent traffic management, automatic ticketing, fleet management, and so on.

Consumer electronics and devices such as eBook readers, digital cameras, personal computers, and navigation systems also form one of the potential use cases. Such devices could use MTC to upgrade firmware or to upload and download online contents.

The E-UTRAN and the EPC have been improved for MTC in Release 10, with a new overload control mechanism specialized in MTC. It is expected that 3GPP will continue to upgrade the systems in Release 11 and beyond. Several optimizations and improvements could be specified in the future to support various MTC applications.

16.1 Overload Control for MTC

Overload control for MTC is the main subject on which 3GPP has focused for the E-UTRAN in Releases 10 and 11. 3GPP improved mechanisms to control overload that might occur

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.

when a tremendous number of MTC devices access the network simultaneously. The improvements stem from the existing overload control mechanisms.

The overload control mechanisms that have been introduced to the E-UTRAN for MTC are as follows:

- core network overload control in Release 10;
- RAN overload control in Release 11.

16.1.1 Overload Control in Release 8

It is generally required that the network should provide overload control mechanisms for avoiding and handling overload situations for normal UEs, for example when an unexpected number of UEs access the network. To protect the core network from overload situations, when a number of UEs send NAS requests, the MME is allowed to reject the NAS requests. However, such rejection cannot prevent eNBs from transferring a number of NAS requests received from UEs to the overloaded MME.

Thus, to protect the core network from a number of NAS requests from a number of UEs, the MME can also control overload by using the S1AP protocol on the S1 interface between the MME and the eNB in Release 8. The following S1AP procedures specified in [3] are used for MME control of overload:

- the Overload Start procedure;
- the Overload Stop procedure.

The MME initiates the Overload Start procedure by sending an Overload Start message to the eNB, as illustrated in Figure 16.1. The Overload Start message informs the eNB what it should do for overload control by including *OverloadAction* in the message. The *Overload-Action* indicates one of the following actions that the eNB should take:

- reject all RRC Connection Establishments for non-emergency mobile originated data transfer;
- reject all RRC Connection Establishments for signaling;
- only permit RRC Connection Establishment for emergency sessions and mobile terminated services;
- only permit RRC Connection Establishment for high priority sessions and mobile terminated services.

After receiving the Overload Start message from the MME, the eNB is ready to apply the *OverloadAction* specified in the Overload Start message to RRC connection requests received from UEs. When the UE sends an *RRCConnectionRequest* message, the UE includes the *EstablishmentCause* in the message, as described in Section 3.4. The *Establish-mentCause* indicates one of the following: emergency, high priority access, mobile terminating access, mobile originating signaling, or mobile originating data.

The eNB uses the indicated *EstablishmentCause* to apply the *OverloadAction* to the corresponding *RRCConnectionRequest* message. For instance, when the *OverloadAction* indicates that RRC Connection Establishment should be permitted only for emergency sessions

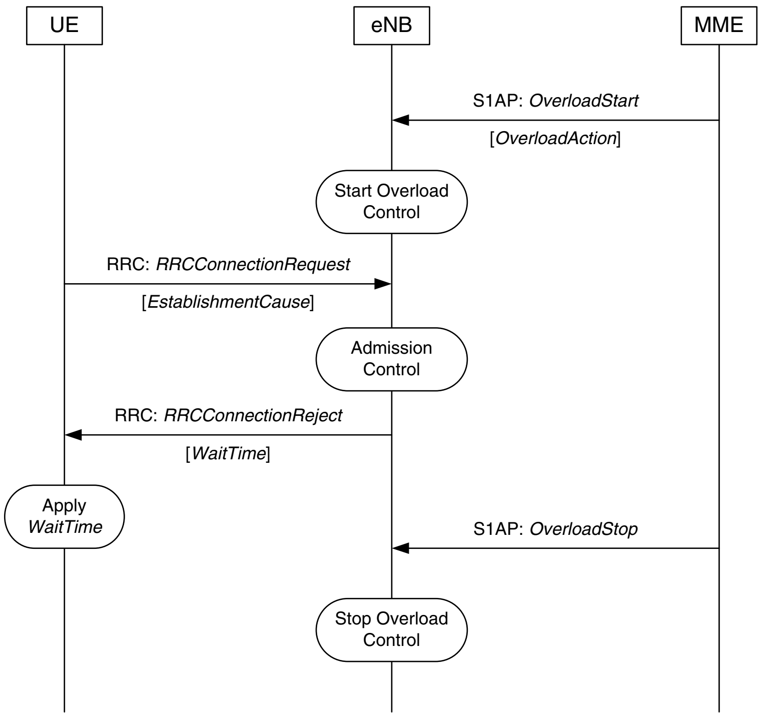

Figure 16.1 An example of MME overload control in Release 8. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

and mobile terminated services, the eNB rejects all the *RRCConnectionRequest* messages that include an *EstablishmentCause* other than emergency and mobile terminating access, by sending *RRCConnectionReject* messages to the UEs.

If the eNB rejects an RRC Connection Establishment, the eNB does not send the Initial UE message to the MME. Thus, by using the Overload Start procedure, the MME filters out Initial UE messages sent from the eNB to the MME.

The eNB includes *waitTime* in the *RRCConnectionReject* message. When the UE receives *waitTime*, the UE starts a timer. The *waitTime* is set from 1 second to 16 seconds. While the timer is running, the UE considers the cell barred except for the case of RRC Connection Establishment for an emergency call, as specified in [4].

Other than the MME controlling core network overload, the eNB may be able to use the Access Class Barring in the RRC layer or the Random Access Backoff mechanism in the MAC layer. Access Class Barring is explained in Section 3.4.1 and the Random Access Backoff mechanism is explained in Section 6.9.

16.1.2 Core Network Overload Control in Release 10

When MTC devices supporting LTE are prevalent on the market, it is expected that a tremendous number of MTC devices will be operating on the network. If many MTC devices try to connect to the same MME in a short period, the MME will be heavily overloaded. In particular, we can imagine that when one network collapses, for example, due to a heavy load, a tremendous number of MTC devices will lose their connections to the collapsed network and, as a result, try to access a neighboring network in a short period. In this scenario, the neighboring network is also expected to become overloaded quickly. To protect networks from such a scenario that could occur with MTC devices, 3GPP improved the MME overload control for the MTC in Release 10.

From Release 10, delay-tolerant access is defined as a new *EstablishmentCause* in the *RRCConnectionRequest* message to control overload of MTC devices selectively. This new *EstablishmentCause* is intended for MTC devices that use delay-tolerant MTC applications. For example, smart metering applications may accept long delays before sending reports to the MTC server. Thus, the network may be able to de-prioritize access for metering devices. Such MTC devices are configured for NAS signaling low priority as a device property in the NAS layer. The NAS signaling low priority indication specified in [5] is mapped to delay-tolerant access of the *EstablishmentCause* in Release 10.

When the eNB applies overload control for delay-tolerant access, the eNB rejects an RRC connection request by sending an *RRCConnectionReject* message, or releases an existing RRC connection by sending an *RRCConnectionRelease* message. In this case, the eNB can include *extendedWaitTime* in such messages.

The NAS layer of the UE uses this *extendedWaitTime* to back off the NAS request initiating RRC Connection Establishment. The *extendedWaitTime* is set up to 30 minutes because it is used for delay-tolerant MTC applications. For instance, metering devices would be allowed to delay reporting metering information to the centralized node for 30 minutes.

Figure 16.2 explains one example of the RRC connection rejection for overload control of delay-tolerant access. When overload occurs, the MME can de-prioritize access for MTC devices configured for NAS signaling low priority, and send an Overload Start message including the *OverloadAction* indicating rejection of delay-tolerant traffic. When the *OverloadAction* indicates rejection of delay-tolerant traffic, the eNB starts to reject all RRC Connection Establishments with *EstablishmentCause* set to delay-tolerant access.

In the meantime, when an NAS request message is triggered in the NAS layer of the UE, the NAS layer initiates the RRC Connection Establishment with the NAS signaling low priority indication. This indication triggers the RRC Connection Establishment procedure with *EstablishmentCause* set to delay-tolerant access.

After overload control has started, the eNB will reject RRC Connection Establishment with *EstablishmentCause* set to delay-tolerant access by sending an *RRCConnectionReject* message with the *extendedWaitTime*. When the UE supports this delay-tolerant access feature, the RRC layer of the UE forwards the received *extendedWaitTime* to the NAS layer of the UE.

It should be noted that the RRC layer of the UE is not required to memorize which *EstablishmentCause* was set in the *RRCConnectionRequest* message. Thus, upon receipt of the *extendedWaitTime*, the RRC layer of the UE forwards the *extendedWaitTime* to the NAS layer even if the RRC Connection Establishment used a cause other than delay-tolerant access. When receiving the *extendedWaitTime* from the RRC layer, the NAS layer can use the *extendedWaitTime* to back off the NAS request.

The eNB is also not required to memorize the *EstablishmentCause* sent by the UE for the RRC Connection Establishment. Thus, the eNB cannot transfer the *EstablishmentCause* to another eNB via the handover procedure.

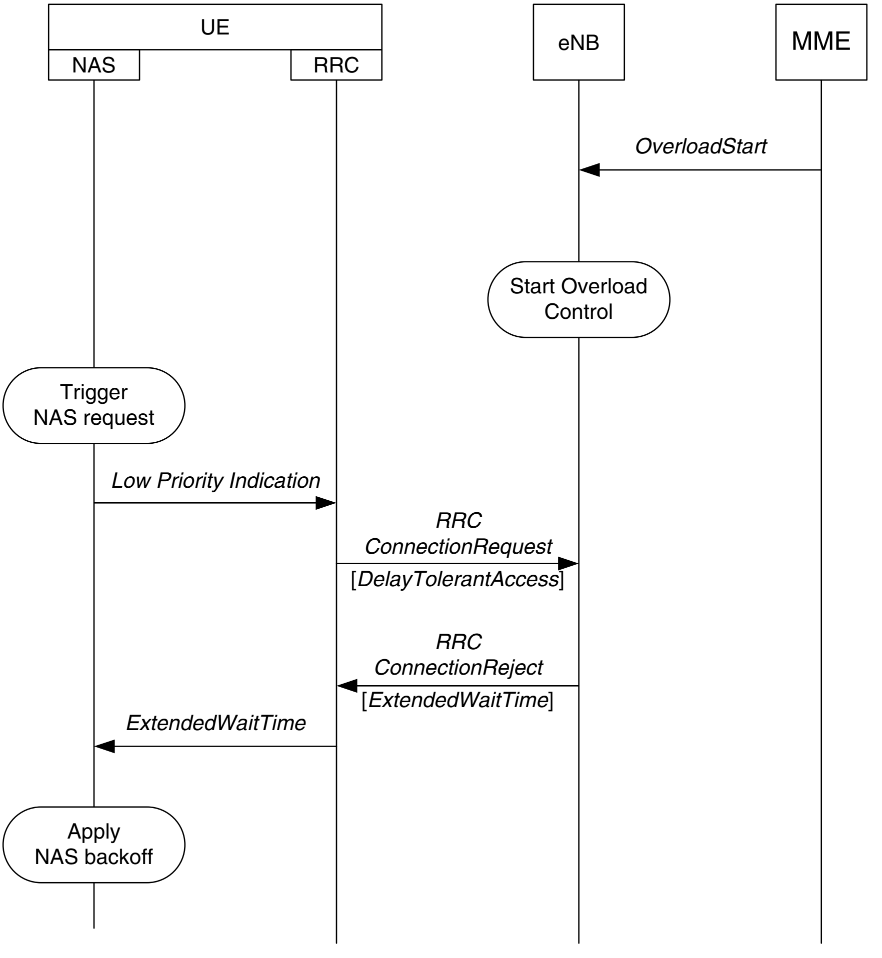

Figure 16.2 An example of the *RRCConnectionReject* message for overload control of delay-tolerant access. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

The eNB may be linked to multiple MMEs. In this case, while overload is ongoing only at a particular MME, when the UE requests an RRC Connection Establishment and an S1 connection, the eNB needs to check to which MME the UE wishes to be connected over the S1 interface. However, since the UE includes information on MME selection in the *RRCConnectionSetupComplete* message (see Section 3.4), the eNB can check information on MME selection only after the UE enters RRC_CONNECTED. As a result, the eNB cannot use the *RRCConnectionReject* message in this case. Instead, the eNB uses the *RRCConnectionRelease* message to release the RRC connection after the RRC connection has been established. Therefore, *extendedWaitTime* is also included in the *RRCConnectionRelease* message. Figure 16.3 shows an example of the RRC Connection Release for overload control of delay-tolerant access.

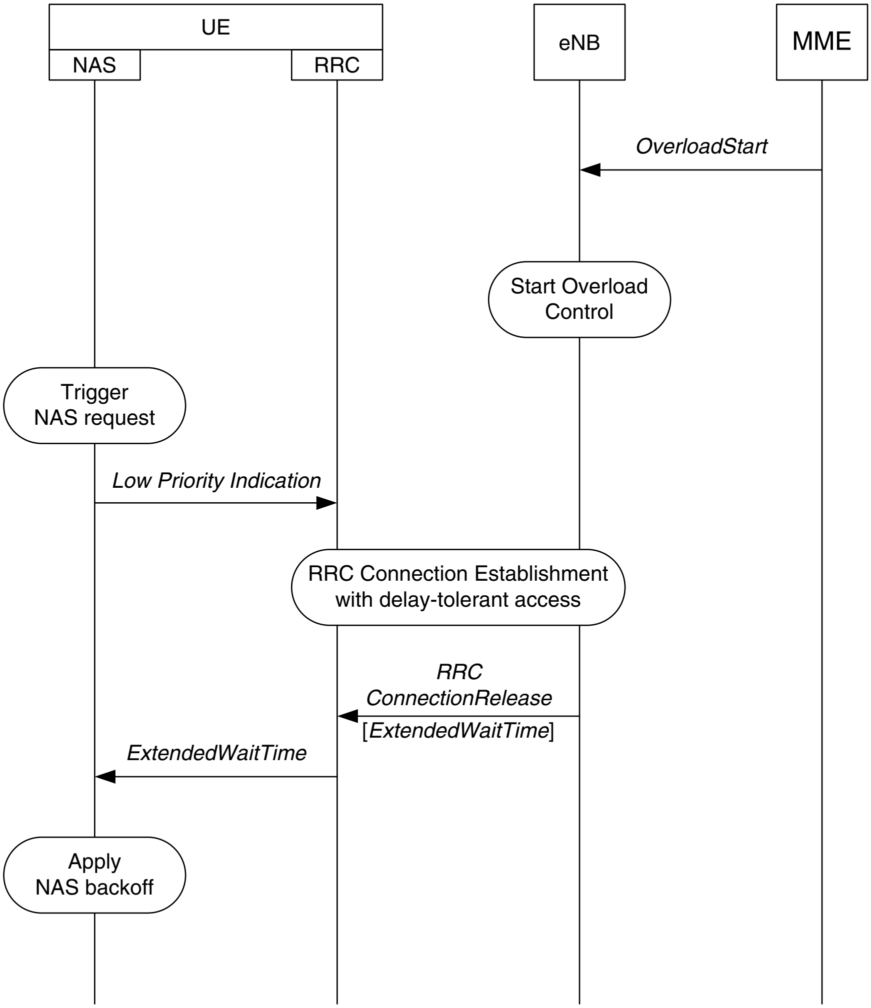

Figure 16.3 An example of the *RRCConnectionRelease* message for overload control of delay-tolerant access. Reproduced by permission of 3GPP, © 2011. 3GPP™ TSs and TRs are the property of ARIB, ATIS, CCSA, ETSI, TTA and TTC who jointly own the copyright in them

Overload control in Release 10 aims to send no Initial UE messages to the overloaded MME. Thus, use of the *extendedWaitTime* in the *RRCConnectionRelease* message is most likely to be used in the early stage of the RRC connection.

16.1.3 RAN Overload Control in Release 11

Overload may occur at the eNB, for example due to a surge of RRC connection requests by many MTC devices in a short period. To protect the eNB from such an overload situation, it would be beneficial for the eNB to control the number of requests for RRC Connection Establishment. For this purpose, the GERAN introduced the Extended Access Barring (EAB) mechanism in Release 10. The E-UTRAN and the UTRAN also introduced the EAB in Release 11, based on the EAB requirements specified in [6].

The eNB broadcasts the EAB information via system information on the BCCH. The EAB information includes a bitmap. Each bit of the bitmap corresponds to one of the Access Classes (ACs) from 0 to 9, and indicates whether or not a member of the corresponding AC is barred. All UEs are members of one out of ten ACs. When the AC allocated to the UE is indicated as being barred in the EAB information, the UE does not send a request for RRC Connection Establishment. Thus, the eNB can bar a number of requests with a granularity of 10%.

When the UE is configured for EAB, the UE applies EAB for RRC Connection Establishment. The UE uses up-to-date EAB information for EAB. When the UE is barred due to EAB, the UE will be barred until the bitmap is updated in the system information. While the UE is barred, the UE will not initiate the Random Access procedure to carry an *RRCConnectionRequest* message. When the UE is not barred by EAB, the UE continues the RRC Connection Establishment procedure.

16.2 MTC Features in 3GPP

MTC applications cannot be categorized by the same characteristics. In other words, one single option about system optimization cannot be applicable for all MTC applications. Thus, 3GPP defined several MTC features in [1] to categorize different options for potential system optimization.

The following service-specific MTC features were defined in 3GPP for various MTC applications. It is expected that 3GPP will gradually specify solutions to cover these MTC features in Release 11 and beyond.

- **Low Mobility:** This feature is intended for use when MTC devices do not move, move infrequently, or move only within a certain area.
- **Time Controlled:** This feature is intended for use when MTC applications tolerate sending or receiving data only for defined time intervals.
- **Time Tolerant:** This feature is intended for use when MTC devices can delay their data transfer.
- **Small Data Transmissions:** This feature is intended for use when MTC devices send or receive small amounts of data.
- **Mobile Originated Only:** This feature is intended for use when MTC devices utilize mobile-originated communications only.
- **Infrequent Mobile Terminated:** This feature is intended for use when MTC devices utilize mainly mobile-originated communications.
- **MTC Monitoring:** This feature is intended for use when monitoring events related to MTC devices.
- **Priority Alarm:** This feature is intended for use when MTC devices generate a priority alarm for an event requiring immediate attention, such as a theft or vandalism.
- **Secure Connection:** This feature is intended for use when a secure connection is needed between the MTC device and the MTC server.
- **Location Specific Trigger:** This feature is intended for use when a trigger is initiated based on where the MTC device is located.

- **Infrequent Transmission:** This feature is intended for use when MTC devices are expected to send or receive data infrequently.
- **Group Based MTC Features:**
  - **Group Based Policing:** This feature is intended for use when the network operator wants to enforce a combined QoS policy for the MTC group members.
  - **Group Based Addressing:** This feature is intended for use when the network operator wants to optimize the message volume for the case where many MTC devices should receive the same message.

References

1. 3GPP Technical Specification 22.368, "Service requirements for Machine-Type Communications (MTC); Stage 1 (Release 11)", www.3gpp.org, accessed on 14 March 2012.
2. 3GPP Technical Report 37.868, "Study on RAN Improvements for Machine-type Communications (Release 10)", www.3gpp.org, accessed on 14 March 2012.
3. 3GPP Technical Specification 36.413, "S1 Application Protocol (S1AP) (Release 10)", www.3gpp.org, accessed on 14 March 2012.
4. 3GPP Technical Specification 36.331, "Radio Resource Control (RRC); Protocol specification (Release 10)", www.3gpp.org, accessed on 14 March 2012.
5. 3GPP Technical Specification 24.301, "Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS); Stage 3 (Release 10)", www.3gpp.org, accessed on 14 March 2012.
6. 3GPP Technical Specification 22.011, "Service accessibility (Release 11)", www.3gpp.org, accessed on 14 March 2012.

Index

Radio Protocols for LTE and LTE-Advanced, First Edition.
SeungJune Yi, SungDuck Chun, YoungDae Lee, SungJun Park and SungHoon Jung.