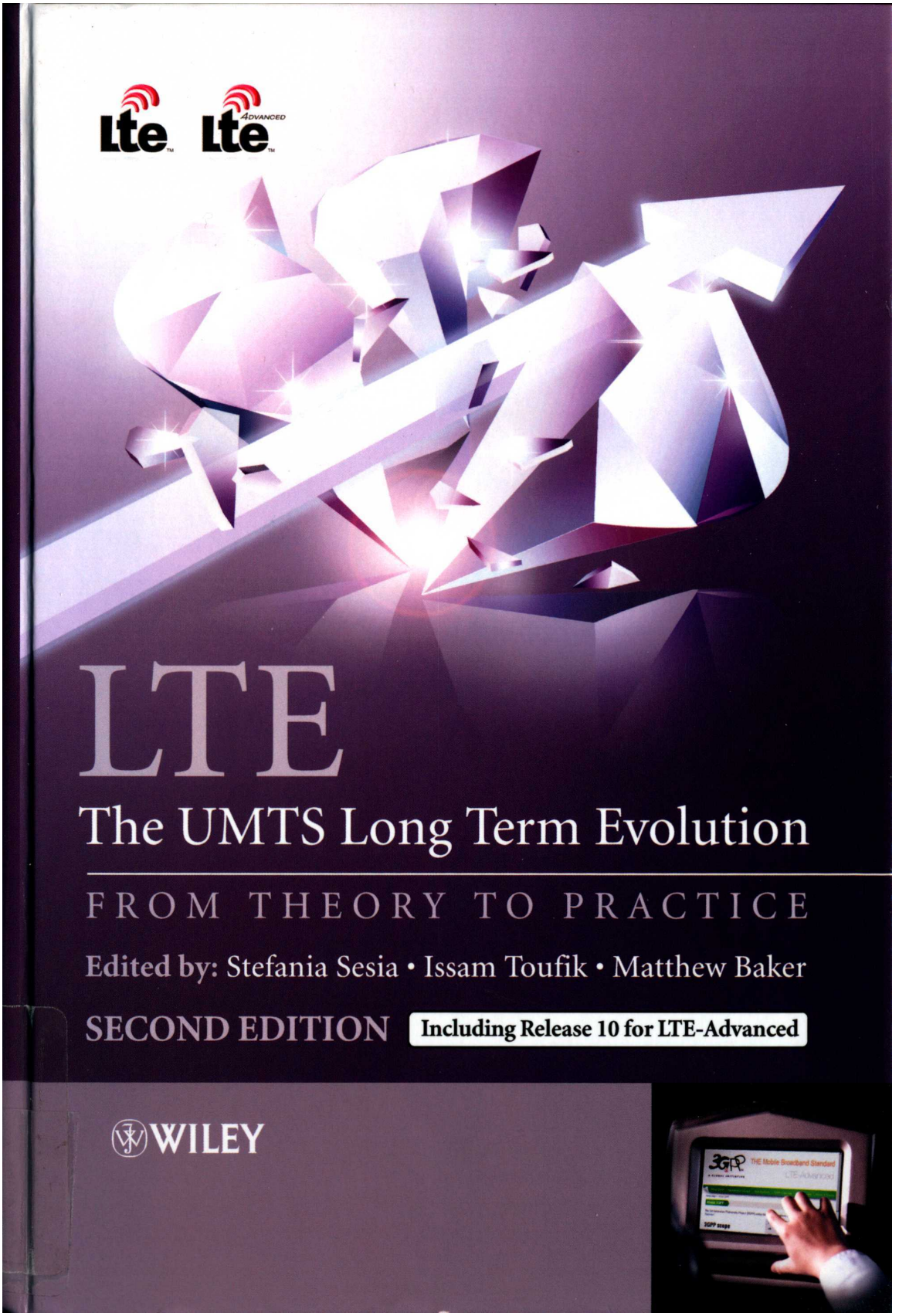


Samsung Ex. 1008
1 of 126

LTE – The UMTS Long Term Evolution

From Theory to Practice

Second Edition

Stefania Sesia
ST-Ericsson, France

Issam Toufik
ETSI, France

Matthew Baker
Alcatel-Lucent, UK

WILEY
A John Wiley & Sons, Ltd., Publication

Samsung Ex. 1008
2 of 126

This edition first published 2011
© 2011 John Wiley & Sons Ltd.

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

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.

The rights of the authors to be identified as the authors of this work have been asserted in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Photograph on cover courtesy of Alcatel-Lucent, from the *ngConnect* LTE-equipped car.
3GPP website reproduced by permission of © 3GPP™.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data

Sesia, Stefania.
LTE–the UMTS long term evolution : from theory to practice / Stefania Sesia, Issam Toufik, Matthew Baker. – 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-66025-6 (hardback)
1. Universal Mobile Telecommunications System. 2. Long-Term Evolution (Telecommunications)
I. Toufik, Issam. II. Baker, Matthew (Matthew P.J.) III. Title.
TK5103.4883.S47 2011
621.3845'6–dc22 2010039466

A catalogue record for this book is available from the British Library.

Print ISBN: 9780470660256 (H/B)
ePDF ISBN: 9780470978511
oBook ISBN: 9780470978504
epub ISBN: 9780470978641

Printed in Great Britain by CPI Antony Rowe, Chippenham, Wiltshire.

Contents

Editors' Biographies xxi

List of Contributors xxiii

Foreword xxvii

Preface xxix

Acknowledgements xxxi

List of Acronyms xxxiii

1 Introduction and Background 1
Thomas Sälzer and Matthew Baker

1.1 The Context for the Long Term Evolution of UMTS 1
1.1.1 Historical Context 1
1.1.2 LTE in the Mobile Radio Landscape 2
1.1.3 The Standardization Process in 3GPP 5
1.2 Requirements and Targets for the Long Term Evolution 7
1.2.1 System Performance Requirements 7
1.2.2 Deployment Cost and Interoperability 12
1.3 Technologies for the Long Term Evolution 14
1.3.1 Multicarrier Technology 14
1.3.2 Multiple Antenna Technology 15
1.3.3 Packet-Switched Radio Interface 16
1.3.4 User Equipment Categories 17
1.3.5 From the First LTE Release to LTE-Advanced 19
1.4 From Theory to Practice 20
References 21

Part I Network Architecture and Protocols 23

2 Network Architecture 25
Sudeep Palat and Philippe Godin

2.1 Introduction . . . 25
2.2 Overall Architectural Overview . . . 26
2.2.1 The Core Network . . . 27
2.2.2 The Access Network . . . 30
2.2.3 Roaming Architecture . . . 31
2.3 Protocol Architecture . . . 32
2.3.1 User Plane . . . 32
2.3.2 Control Plane . . . 33
2.4 Quality of Service and EPS Bearers . . . 34
2.4.1 Bearer Establishment Procedure . . . 37
2.4.2 Inter-Working with other RATs . . . 38
2.5 The E-UTRAN Network Interfaces: S1 Interface . . . 40
2.5.1 Protocol Structure over S1 . . . 41
2.5.2 Initiation over S1 . . . 43
2.5.3 Context Management over S1 . . . 43
2.5.4 Bearer Management over S1 . . . 44
2.5.5 Paging over S1 . . . 44
2.5.6 Mobility over S1 . . . 45
2.5.7 Load Management over S1 . . . 47
2.5.8 Trace Function . . . 48
2.5.9 Delivery of Warning Messages . . . 48
2.6 The E-UTRAN Network Interfaces: X2 Interface . . . 49
2.6.1 Protocol Structure over X2 . . . 49
2.6.2 Initiation over X2 . . . 49
2.6.3 Mobility over X2 . . . 51
2.6.4 Load and Interference Management Over X2 . . . 54
2.6.5 UE Historical Information Over X2 . . . 54
2.7 Summary . . . 55
References . . . 55

3 Control Plane Protocols 57
Himke van der Velde

3.1 Introduction . . . 57
3.2 Radio Resource Control (RRC) . . . 58
3.2.1 Introduction . . . 58
3.2.2 System Information . . . 59
3.2.3 Connection Control within LTE . . . 63
3.2.4 Connected Mode Inter-RAT Mobility . . . 73
3.2.5 Measurements . . . 75
3.2.6 Other RRC Signalling Aspects . . . 78
3.3 PLMN and Cell Selection . . . 78

Samsung Ex. 1008
5 of 126

3.3.1 Introduction . . . 78
3.3.2 PLMN Selection . . . 79
3.3.3 Cell Selection . . . 79
3.3.4 Cell Reselection . . . 80
3.4 Paging . . . 84
3.5 Summary . . . 86
References . . . 86

4 User Plane Protocols **87**
Patrick Fischer, SeungJune Yi, SungDuck Chun and YoungDae Lee

4.1 Introduction to the User Plane Protocol Stack . . . 87
4.2 Packet Data Convergence Protocol (PDCP) . . . 89
4.2.1 Functions and Architecture . . . 89
4.2.2 Header Compression . . . 90
4.2.3 Security . . . 92
4.2.4 Handover . . . 93
4.2.5 Discard of Data Packets . . . 95
4.2.6 PDCP PDU Formats . . . 97
4.3 Radio Link Control (RLC) . . . 98
4.3.1 RLC Entities . . . 99
4.3.2 RLC PDU Formats . . . 105
4.4 Medium Access Control (MAC) . . . 108
4.4.1 MAC Architecture . . . 108
4.4.2 MAC Functions . . . 111
4.5 Summary of the User Plane Protocols . . . 120
References . . . 120

Part II Physical Layer for Downlink **121**

5 Orthogonal Frequency Division Multiple Access (OFDMA) **123**
Andrea Ancora, Issam Toufik, Andreas Bury and Dirk Slock

5.1 Introduction . . . 123
5.1.1 History of OFDM Development . . . 124
5.2 OFDM . . . 125
5.2.1 Orthogonal Multiplexing Principle . . . 125
5.2.2 Peak-to-Average Power Ratio and Sensitivity to Non-Linearity . . . 131
5.2.3 Sensitivity to Carrier Frequency Offset and Time-Varying Channels . 133
5.2.4 Timing Offset and Cyclic Prefix Dimensioning . . . 135
5.3 OFDMA . . . 137
5.4 Parameter Dimensioning . . . 139
5.4.1 Physical Layer Parameters for LTE . . . 140
5.5 Summary . . . 142
References . . . 142

Samsung Ex. 1008
6 of 126

6 Introduction to Downlink Physical Layer Design **145**
Matthew Baker

6.1 Introduction . . . 145
6.2 Transmission Resource Structure . . . 145
6.3 Signal Structure . . . 148
6.4 Introduction to Downlink Operation . . . 149
References . . . 150

7 Synchronization and Cell Search **151**
Fabrizio Tomatis and Stefania Sesia

7.1 Introduction . . . 151
7.2 Synchronization Sequences and Cell Search in LTE . . . 151
7.2.1 Zadoff–Chu Sequences . . . 155
7.2.2 Primary Synchronization Signal (PSS) Sequences . . . 157
7.2.3 Secondary Synchronization Signal (SSS) Sequences . . . 158
7.3 Coherent Versus Non-Coherent Detection . . . 161
References . . . 163

8 Reference Signals and Channel Estimation **165**
Andrea Ancora, Stefania Sesia and Alex Gorokhov

8.1 Introduction . . . 165
8.2 Design of Reference Signals in the LTE Downlink . . . 167
8.2.1 Cell-Specific Reference Signals . . . 168
8.2.2 UE-Specific Reference Signals in Release 8 . . . 171
8.2.3 UE-Specific Reference Signals in Release 9 . . . 171
8.3 RS-Aided Channel Modelling and Estimation . . . 174
8.3.1 Time-Frequency-Domain Correlation: The WSSUS Channel Model . 175
8.3.2 Spatial-Domain Correlation: The Kronecker Model . . . 176
8.4 Frequency-Domain Channel Estimation . . . 178
8.4.1 Channel Estimate Interpolation . . . 178
8.4.2 General Approach to Linear Channel Estimation . . . 179
8.4.3 Performance Comparison . . . 180
8.5 Time-Domain Channel Estimation . . . 181
8.5.1 Finite and Infinite Length MMSE . . . 182
8.5.2 Normalized Least-Mean-Square . . . 184
8.6 Spatial-Domain Channel Estimation . . . 184
8.7 Advanced Techniques . . . 185
References . . . 186

9 Downlink Physical Data and Control Channels **189**
Matthew Baker and Tim Moulsley

9.1 Introduction . . . 189
9.2 Downlink Data-Transporting Channels . . . 189
9.2.1 Physical Broadcast Channel (PBCH) . . . 189
9.2.2 Physical Downlink Shared CHannel (PDSCH) . . . 192

Samsung Ex. 1008
7 of 126

9.2.3 Physical Multicast Channel (PMCH) . . . 196
9.3 Downlink Control Channels . . . 196
9.3.1 Requirements for Control Channel Design . . . 196
9.3.2 Control Channel Structure . . . 198
9.3.3 Physical Control Format Indicator CHannel (PCFICH) . . . 198
9.3.4 Physical Hybrid ARQ Indicator Channel (PHICH) . . . 200
9.3.5 Physical Downlink Control CHannel (PDCCH) . . . 202
9.3.6 PDCCH Scheduling Process . . . 212
References . . . 214

10 Link Adaptation and Channel Coding **215**
Brian Classon, Ajit Nimbalker, Stefania Sesia and Issam Toufik

10.1 Introduction . . . 215
10.2 Link Adaptation and CQI Feedback . . . 217
10.2.1 CQI Feedback in LTE . . . 218
10.3 Channel Coding . . . 223
10.3.1 Theoretical Aspects of Channel Coding . . . 223
10.3.2 Channel Coding for Data Channels in LTE . . . 232
10.3.3 Channel Coding for Control Channels in LTE . . . 244
10.4 Conclusions . . . 245
References . . . 246

11 Multiple Antenna Techniques **249**
Thomas Sälzer, David Gesbert, Cornelius van Rensburg, Filippo Tosato, Florian Kaltenberger and Tetsushi Abe

11.1 Fundamentals of Multiple Antenna Theory . . . 249
11.1.1 Overview . . . 249
11.1.2 MIMO Signal Model . . . 252
11.1.3 Single-User MIMO Techniques . . . 253
11.1.4 Multi-User MIMO Techniques . . . 258
11.2 MIMO Schemes in LTE . . . 262
11.2.1 Practical Considerations . . . 263
11.2.2 Single-User Schemes . . . 264
11.2.3 Multi-User MIMO . . . 274
11.2.4 MIMO Performance . . . 276
11.3 Summary . . . 276
References . . . 277

12 Multi-User Scheduling and Interference Coordination **279**
Issam Toufik and Raymond Knopp

12.1 Introduction . . . 279
12.2 General Considerations for Resource Allocation Strategies . . . 280
12.3 Scheduling Algorithms . . . 283
12.3.1 Ergodic Capacity . . . 283
12.3.2 Delay-Limited Capacity . . . 285

12.4 Considerations for Resource Scheduling in LTE . . . 286
12.5 Interference Coordination and Frequency Reuse . . . 287
12.5.1 Inter-eNodeB Signalling to Support Downlink Frequency-Domain ICIC in LTE . . . 290
12.5.2 Inter-eNodeB Signalling to Support Uplink Frequency-Domain ICIC in LTE . . . 290
12.5.3 Static versus Semi-Static ICIC . . . 291
12.6 Summary . . . 291
References . . . 292

13 Broadcast Operation 293
Himke van der Velde, Olivier Hus and Matthew Baker

13.1 Introduction . . . 293
13.2 Broadcast Modes . . . 293
13.3 Overall MBMS Architecture . . . 295
13.3.1 Reference Architecture . . . 295
13.3.2 Content Provision . . . 295
13.3.3 Core Network . . . 296
13.3.4 Radio Access Network – E-UTRAN/UTRAN/GERAN and UE . . . 296
13.3.5 MBMS Interfaces . . . 297
13.4 MBMS Single Frequency Network Transmission . . . 297
13.4.1 Physical Layer Aspects . . . 297
13.4.2 MBSFN Areas . . . 301
13.5 MBMS Characteristics . . . 303
13.5.1 Mobility Support . . . 303
13.5.2 UE Capabilities and Service Prioritization . . . 303
13.6 Radio Access Protocol Architecture and Signalling . . . 304
13.6.1 Protocol Architecture . . . 304
13.6.2 Session Start Signalling . . . 305
13.6.3 Radio Resource Control (RRC) Signalling Aspects . . . 306
13.6.4 Content Synchronization . . . 308
13.6.5 Counting Procedure . . . 310
13.7 Public Warning Systems . . . 312
13.8 Comparison of Mobile Broadcast Modes . . . 312
13.8.1 Delivery by Cellular Networks . . . 312
13.8.2 Delivery by Broadcast Networks . . . 313
13.8.3 Services and Applications . . . 313
References . . . 314

Part III Physical Layer for Uplink 315

14 Uplink Physical Layer Design 317
Robert Love and Vijay Nangia

14.1 Introduction . . . 317
14.2 SC-FDMA Principles . . . 318

14.2.1 SC-FDMA Transmission Structure . . . 318
14.2.2 Time-Domain Signal Generation . . . 318
14.2.3 Frequency-Domain Signal Generation (DFT-S-OFDM) . . . 320
14.3 SC-FDMA Design in LTE . . . 321
14.3.1 Transmit Processing for LTE . . . 321
14.3.2 SC-FDMA Parameters for LTE . . . 322
14.3.3 d.c. Subcarrier in SC-FDMA . . . 324
14.3.4 Pulse Shaping . . . 324
14.4 Summary . . . 325
References . . . 326

15 Uplink Reference Signals **327**
Robert Love and Vijay Nangia

15.1 Introduction . . . 327
15.2 RS Signal Sequence Generation . . . 328
15.2.1 Base RS Sequences and Sequence Grouping . . . 330
15.2.2 Orthogonal RS via Cyclic Time-Shifts of a Base Sequence . . . 330
15.3 Sequence-Group Hopping and Planning . . . 332
15.3.1 Sequence-Group Hopping . . . 332
15.3.2 Sequence-Group Planning . . . 333
15.4 Cyclic Shift Hopping . . . 333
15.5 Demodulation Reference Signals (DM-RS) . . . 335
15.6 Uplink Sounding Reference Signals (SRS) . . . 337
15.6.1 SRS Subframe Configuration and Position . . . 337
15.6.2 Duration and Periodicity of SRS Transmissions . . . 337
15.6.3 SRS Symbol Structure . . . 338
15.7 Summary . . . 340
References . . . 341

16 Uplink Physical Channel Structure **343**
Robert Love and Vijay Nangia

16.1 Introduction . . . 343
16.2 Physical Uplink Shared Data Channel Structure . . . 344
16.2.1 Scheduling on PUSCH . . . 345
16.2.2 PUSCH Transport Block Sizes . . . 347
16.3 Uplink Control Channel Design . . . 348
16.3.1 Physical Uplink Control Channel (PUCCH) Structure . . . 348
16.3.2 Types of Control Signalling Information and PUCCH Formats . . . 352
16.3.3 Channel State Information Transmission on PUCCH (Format 2) . . . 353
16.3.4 Multiplexing of CSI and HARQ ACK/NACK from a UE on PUCCH 355
16.3.5 HARQ ACK/NACK Transmission on PUCCH (Format 1a/1b) . . . 356
16.3.6 Multiplexing of CSI and HARQ ACK/NACK in the Same (Mixed) PUCCH RB . . . 363
16.3.7 Scheduling Request (SR) Transmission on PUCCH (Format 1) . . . 363
16.4 Multiplexing of Control Signalling and UL-SCH Data on PUSCH . . . 365
16.5 ACK/NACK Repetition . . . 367

16.6 Multiple-Antenna Techniques . . . 367
16.6.1 Closed-Loop Switched Antenna Diversity . . . 367
16.6.2 Multi-User 'Virtual' MIMO or SDMA . . . 368
16.7 Summary . . . 369
References . . . 369

17 Random Access **371**
Pierre Bertrand and Jing Jiang

17.1 Introduction . . . 371
17.2 Random Access Usage and Requirements in LTE . . . 371
17.3 Random Access Procedure . . . 372
17.3.1 Contention-Based Random Access Procedure . . . 373
17.3.2 Contention-Free Random Access Procedure . . . 376
17.4 Physical Random Access Channel Design . . . 376
17.4.1 Multiplexing of PRACH with PUSCH and PUCCH . . . 376
17.4.2 The PRACH Structure . . . 377
17.4.3 Preamble Sequence Theory and Design . . . 385
17.5 PRACH Implementation . . . 396
17.5.1 UE Transmitter . . . 397
17.5.2 eNodeB PRACH Receiver . . . 398
17.6 Time Division Duplex (TDD) PRACH . . . 404
17.6.1 Preamble Format 4 . . . 404
17.7 Concluding Remarks . . . 405
References . . . 406

18 Uplink Transmission Procedures **407**
Matthew Baker

18.1 Introduction . . . 407
18.2 Uplink Timing Control . . . 407
18.2.1 Overview . . . 407
18.2.2 Timing Advance Procedure . . . 408
18.3 Power Control . . . 411
18.3.1 Overview . . . 411
18.3.2 Detailed Power Control Behaviour . . . 412
18.3.3 UE Power Headroom Reporting . . . 419
18.3.4 Summary of Uplink Power Control Strategies . . . 420
References . . . 420

Part IV Practical Deployment Aspects **421**

19 User Equipment Positioning **423**
Karri Ranta-aho and Zukang Shen

19.1 Introduction . . . 423
19.2 Assisted Global Navigation Satellite System (A-GNSS) Positioning . . . 425
19.3 Observed Time Difference Of Arrival (OTDOA) Positioning . . . 426

Samsung Ex. 1008
11 of 126

19.3.1 Positioning Reference Signals (PRS) . . . 427
19.3.2 OTDOA Performance and Practical Considerations . . . 430
19.4 Cell-ID-based Positioning . . . 431
19.4.1 Basic CID Positioning . . . 431
19.4.2 Enhanced CID Positioning using Round Trip Time and UE Receive Level Measurements . . . 431
19.4.3 Enhanced CID Positioning using Round Trip Time and Angle of Arrival . . . 432
19.5 LTE Positioning Protocols . . . 433
19.6 Summary and Future Techniques . . . 435
References . . . 436

20 The Radio Propagation Environment **437**
Juha Ylitalo and Tommi Jämsä

20.1 Introduction . . . 437
20.2 SISO and SIMO Channel Models . . . 438
20.2.1 ITU Channel Model . . . 439
20.2.2 3GPP Channel Model . . . 440
20.2.3 Extended ITU Models . . . 440
20.3 MIMO Channel Models . . . 441
20.3.1 SCM Channel Model . . . 442
20.3.2 SCM-Extension Channel Model . . . 444
20.3.3 WINNER Model . . . 445
20.3.4 LTE Evaluation Model . . . 446
20.3.5 Extended ITU Models with Spatial Correlation . . . 448
20.3.6 ITU Channel Models for IMT-Advanced . . . 449
20.3.7 Comparison of MIMO Channel Models . . . 453
20.4 Radio Channel Implementation for Conformance Testing . . . 454
20.4.1 Performance and Conformance Testing . . . 454
20.4.2 Future Testing Challenges . . . 454
20.5 Concluding Remarks . . . 455
References . . . 455

21 Radio Frequency Aspects **457**
Moray Rumney, Takaharu Nakamura, Stefania Sesia, Tony Sayers and Adrian Payne

21.1 Introduction . . . 457
21.2 Frequency Bands and Arrangements . . . 459
21.3 Transmitter RF Requirements . . . 462
21.3.1 Requirements for the Intended Transmissions . . . 462
21.3.2 Requirements for Unwanted Emissions . . . 467
21.3.3 Power Amplifier Considerations . . . 471
21.4 Receiver RF Requirements . . . 474
21.4.1 Receiver General Requirements . . . 474
21.4.2 Transmit Signal Leakage . . . 475
21.4.3 Maximum Input Level . . . 477
21.4.4 Small Signal Requirements . . . 478

21.4.5 Selectivity and Blocking Specifications . . . 482
21.4.6 Spurious Emissions . . . 488
21.4.7 Intermodulation Requirements . . . 489
21.4.8 Dynamic Range . . . 491
21.5 RF Impairments . . . 492
21.5.1 Transmitter RF Impairments . . . 492
21.5.2 Model of the Main RF Impairments . . . 495
21.6 Summary . . . 500
References . . . 501

22 Radio Resource Management **503**
Muhammad Kazmi

22.1 Introduction . . . 503
22.2 Cell Search Performance . . . 505
22.2.1 Cell Search within E-UTRAN . . . 505
22.2.2 E-UTRAN to E-UTRAN Cell Global Identifier Reporting Requirements . . . 509
22.2.3 E-UTRAN to UTRAN Cell Search . . . 510
22.2.4 E-UTRAN to GSM Cell Search . . . 511
22.2.5 Enhanced Inter-RAT Measurement Requirements . . . 512
22.3 Mobility Measurements . . . 513
22.3.1 E-UTRAN Measurements . . . 513
22.3.2 UTRAN Measurements . . . 514
22.3.3 GSM Measurements: GSM Carrier RSSI . . . 516
22.3.4 CDMA2000 Measurements . . . 516
22.4 UE Measurement Reporting Mechanisms and Requirements . . . 516
22.4.1 E-UTRAN Event Triggered Reporting Requirements . . . 517
22.4.2 Inter-RAT Event-Triggered Reporting . . . 517
22.5 Mobility Performance . . . 518
22.5.1 Mobility Performance in RRC_IDLE State . . . 518
22.5.2 Mobility Performance in RRC_CONNECTED State . . . 522
22.6 RRC Connection Mobility Control Performance . . . 525
22.6.1 RRC Connection Re-establishment . . . 525
22.6.2 Random Access . . . 525
22.7 Radio Link Monitoring Performance . . . 526
22.7.1 In-sync and Out-of-sync Thresholds . . . 526
22.7.2 Requirements without DRX . . . 527
22.7.3 Requirements with DRX . . . 527
22.7.4 Requirements during Transitions . . . 527
22.8 Concluding Remarks . . . 528
References . . . 529

23 Paired and Unpaired Spectrum **531**
Nicholas Anderson

23.1 Introduction . . . 531
23.2 Duplex Modes . . . 532

23.3 Interference Issues in Unpaired Spectrum . . . 533
23.3.1 Adjacent Carrier Interference Scenarios . . . 535
23.3.2 Summary of Interference Scenarios . . . 543
23.4 Half-Duplex System Design Aspects . . . 544
23.4.1 Accommodation of Transmit–Receive Switching . . . 544
23.4.2 Coexistence between Dissimilar Systems . . . 547
23.4.3 HARQ and Control Signalling for TDD Operation . . . 548
23.4.4 Half-Duplex FDD (HD-FDD) Physical Layer Operation . . . 551
23.5 Reciprocity . . . 552
23.5.1 Conditions for Reciprocity . . . 554
23.5.2 Applications of Reciprocity . . . 558
23.5.3 Summary of Reciprocity Considerations . . . 561
References . . . 562

24 Picocells, Femtocells and Home eNodeBs **563**
Philippe Godin and Nick Whinnett

24.1 Introduction . . . 563
24.2 Home eNodeB Architecture . . . 564
24.2.1 Architecture Overview . . . 564
24.2.2 Functionalities . . . 565
24.2.3 Mobility . . . 566
24.2.4 Local IP Access Support . . . 568
24.3 Interference Management for Femtocell Deployment . . . 569
24.3.1 Interference Scenarios . . . 570
24.3.2 Network Listen Mode . . . 574
24.4 RF Requirements for Small Cells . . . 574
24.4.1 Transmitter Specifications . . . 575
24.4.2 Receiver Specifications . . . 576
24.4.3 Demodulation Performance Requirements . . . 578
24.4.4 Time Synchronization for TDD Operation . . . 579
24.5 Summary . . . 580
References . . . 580

25 Self-Optimizing Networks **581**
Philippe Godin

25.1 Introduction . . . 581
25.2 Automatic Neighbour Relation Function (ANRF) . . . 582
25.2.1 Intra-LTE ANRF . . . 582
25.2.2 Automatic Neighbour Relation Table . . . 583
25.2.3 Inter-RAT or Inter-Frequency ANRF . . . 583
25.3 Self-Configuration of eNodeB and MME . . . 584
25.3.1 Self-Configuration of eNodeB/MME over S1 . . . 585
25.3.2 Self-Configuration of IP address and X2 interface . . . 585
25.4 Automatic Configuration of Physical Cell Identity . . . 587
25.5 Mobility Load Balancing Optimization . . . 587

25.5.1 Intra-LTE Load Exchange . . . 588
25.5.2 Intra-LTE Handover Parameter Optimization . . . 589
25.5.3 Inter-RAT Load Exchange . . . 590
25.5.4 Enhanced Inter-RAT Load Exchange . . . 590
25.6 Mobility Robustness Optimization . . . 591
25.6.1 Too-Late Handover . . . 591
25.6.2 Coverage Hole Detection . . . 591
25.6.3 Too-Early Handover . . . 592
25.6.4 Handover to an Inappropriate Cell . . . 592
25.6.5 MRO Verdict Improvement . . . 593
25.6.6 Handover to an Unprepared Cell . . . 594
25.6.7 Unnecessary Inter-RAT Handovers . . . 594
25.6.8 Potential Remedies for Identified Mobility Problems . . . 595
25.7 Random Access CHannel (RACH) Self-Optimization . . . 595
25.8 Energy Saving . . . 596
25.9 Emerging New SON Use Cases . . . 597
References . . . 598

26 LTE System Performance **599**
Tetsushi Abe

26.1 Introduction . . . 599
26.2 Factors Contributing to LTE System Capacity . . . 599
26.2.1 Multiple Access Techniques . . . 600
26.2.2 Frequency Reuse and Interference Management . . . 600
26.2.3 Multiple Antenna Techniques . . . 601
26.2.4 Semi-Persistent Scheduling . . . 601
26.2.5 Short Subframe Duration and Low HARQ Round Trip Time . . . 602
26.2.6 Advanced Receivers . . . 602
26.2.7 Layer 1 and Layer 2 Overhead . . . 602
26.3 LTE Capacity Evaluation . . . 603
26.3.1 Downlink and Uplink Spectral Efficiency . . . 605
26.3.2 VoIP Capacity . . . 608
26.4 LTE Coverage and Link Budget . . . 608
26.5 Summary . . . 610
References . . . 611

Part V LTE-Advanced **613**

27 Introduction to LTE-Advanced **615**
Dirk Gerstenberger

27.1 Introduction and Requirements . . . 615
27.2 Overview of the Main Features of LTE-Advanced . . . 618
27.3 Backward Compatibility . . . 619
27.4 Deployment Aspects . . . 620
27.5 UE Categories for LTE-Advanced . . . 621

References 622

28 Carrier Aggregation **623**
Juan Montojo and Jelena Damnjanovic

28.1 Introduction 623
28.2 Protocols for Carrier Aggregation 624
28.2.1 Initial Acquisition, Connection Establishment and CC Management . 624
28.2.2 Measurements and Mobility 625
28.2.3 User Plane Protocols 628
28.3 Physical Layer Aspects 631
28.3.1 Downlink Control Signalling 631
28.3.2 Uplink Control Signalling 636
28.3.3 Sounding Reference Signals 642
28.3.4 Uplink Timing Advance 642
28.3.5 Uplink Power Control 642
28.3.6 Uplink Multiple Access Scheme Enhancements 644
28.4 UE Transmitter and Receiver Aspects 648
28.4.1 UE Transmitter Aspects of Carrier Aggregation 648
28.4.2 UE Receiver Aspects of Carrier Aggregation 648
28.4.3 Prioritized Carrier Aggregation Scenarios 649
28.5 Summary 650
References 650

29 Multiple Antenna Techniques for LTE-Advanced **651**
Alex Gorokhov, Amir Farajidana, Kapil Bhattad, Xiliang Luo and Stefan Geirhofer

29.1 Downlink Reference Signals 651
29.1.1 Downlink Reference Signals for Demodulation 652
29.1.2 Downlink Reference Signals for Estimation of Channel State Information (CSI-RS) 654
29.2 Uplink Reference Signals 657
29.2.1 Uplink DeModulation Reference Signals (DM-RS) 657
29.2.2 Sounding Reference Signals (SRSs) 658
29.3 Downlink MIMO Enhancements 659
29.3.1 Downlink 8-Antenna Transmission 659
29.3.2 Enhanced Downlink Multi-User MIMO 661
29.3.3 Enhanced CSI Feedback 662
29.4 Uplink Multiple Antenna Transmission 666
29.4.1 Uplink SU-MIMO for PUSCH 666
29.4.2 Uplink Transmit Diversity for PUCCH 668
29.5 Coordinated MultiPoint (CoMP) Transmission and Reception 669
29.5.1 Cooperative MIMO Schemes and Scenarios 669
29.6 Summary 671
References 671

30 Relaying 673
Eric Hardouin, J. Nicholas Laneman, Alexander Golitschek, Hidetoshi Suzuki, Osvaldo Gonsa
30.1 Introduction . . . 673
30.1.1 What is Relaying? . . . 673
30.1.2 Characteristics of Relay Nodes . . . 675
30.1.3 Protocol Functionality of Relay Nodes . . . 676
30.1.4 Relevant Deployment Scenarios . . . 677
30.2 Theoretical Analysis of Relaying . . . 679
30.2.1 Relaying Strategies and Benefits . . . 679
30.2.2 Duplex Constraints and Resource Allocation . . . 683
30.3 Relay Nodes in LTE-Advanced . . . 684
30.3.1 Types of RN . . . 684
30.3.2 Backhaul and Access Resource Sharing . . . 685
30.3.3 Relay Architecture . . . 687
30.3.4 RN Initialization and Configuration . . . 689
30.3.5 Random Access on the Backhaul Link . . . 690
30.3.6 Radio Link Failure on the Backhaul Link . . . 690
30.3.7 RN Security . . . 690
30.3.8 Backhaul Physical Channels . . . 691
30.3.9 Backhaul Scheduling . . . 696
30.3.10 Backhaul HARQ . . . 698
30.4 Summary . . . 699
References . . . 699

31 Additional Features of LTE Release 10 701
Teck Hu, Philippe Godin and Sudeep Palat
31.1 Introduction . . . 701
31.2 Enhanced Inter-Cell Interference Coordination . . . 701
31.2.1 LTE Interference Management . . . 703
31.2.2 Almost Blank Subframes . . . 703
31.2.3 X2 Interface Enhancements for Time-Domain ICIC . . . 705
31.2.4 UE Measurements in Time-Domain ICIC Scenarios . . . 706
31.2.5 RRC Signalling for Restricted Measurements . . . 708
31.2.6 ABS Deployment Considerations . . . 709
31.3 Minimization of Drive Tests . . . 710
31.3.1 Logged MDT . . . 711
31.3.2 Immediate MDT . . . 712
31.4 Machine-Type Communications . . . 712
References . . . 714

32 LTE-Advanced Performance and Future Developments 715
Takehiro Nakamura and Tetsushi Abe
32.1 LTE-Advanced System Performance . . . 715
32.2 Future Developments . . . 718
References . . . 720

Index 721

Samsung Ex. 1008
17 of 126

1

Introduction and Background

Thomas Sälzer and Matthew Baker

1.1 The Context for the Long Term Evolution of UMTS

1.1.1 Historical Context

The Long Term Evolution of UMTS is one of the latest steps in an advancing series of mobile telecommunications systems. Arguably, at least for land-based systems, the series began in 1947 with the development of the concept of *cells* by Bell Labs, USA. The use of cells enabled the capacity of a mobile communications network to be increased substantially, by dividing the coverage area up into small cells each with its own base station operating on a different frequency.

The early systems were confined within national boundaries. They attracted only a small number of users, as the equipment on which they relied was expensive, cumbersome and power-hungry, and therefore was only really practical in a car.

The first mobile communication systems to see large-scale commercial growth arrived in the 1980s and became known as the 'First Generation' systems. The First Generation used analogue technology and comprised a number of independently developed systems worldwide (e.g. AMPS (Analogue Mobile Phone System, used in America), TACS (Total Access Communication System, used in parts of Europe), NMT (Nordic Mobile Telephone, used in parts of Europe) and J-TACS (Japanese Total Access Communication System, used in Japan and Hong Kong)).

Global roaming first became a possibility with the development of the 'Second Generation' system known as GSM (Global System for Mobile communications), which was based on digital technology. The success of GSM was due in part to the collaborative spirit in which it was developed. By harnessing the creative expertise of a number of companies working

LTE – The UMTS Long Term Evolution: From Theory to Practice, Second Edition.
Stefania Sesia, Issam Toufik and Matthew Baker.
© 2011 John Wiley & Sons, Ltd. Published 2011 by John Wiley & Sons, Ltd.

together under the auspices of the European Telecommunications Standards Institute (ETSI), GSM became a robust, interoperable and widely accepted standard.

Fuelled by advances in mobile handset technology, which resulted in small, fashionable terminals with a long battery life, the widespread acceptance of the GSM standard exceeded initial expectations and helped to create a vast new market. The resulting near-universal penetration of GSM phones in the developed world provided an ease of communication never previously possible, first by voice and text message, and later also by more advanced data services. Meanwhile in the developing world, GSM technology had begun to connect communities and individuals in remote regions where fixed-line connectivity was non-existent and would be prohibitively expensive to deploy.

This ubiquitous availability of user-friendly mobile communications, together with increasing consumer familiarity with such technology and practical reliance on it, thus provides the context for new systems with more advanced capabilities. In the following section, the series of progressions which have succeeded GSM is outlined, culminating in the development of the system known as LTE – the Long Term Evolution of UMTS (Universal Mobile Telecommunications System).

1.1.2 LTE in the Mobile Radio Landscape

In contrast to transmission technologies using media such as copper lines and optical fibres, the radio spectrum is a medium shared between different, and potentially interfering, technologies.

As a consequence, regulatory bodies – in particular, ITU-R (International Telecommunication Union – Radiocommunication Sector) [1], but also regional and national regulators – play a key role in the evolution of radio technologies since they decide which parts of the spectrum and how much bandwidth may be used by particular types of service and technology. This role is facilitated by the *standardization* of families of radio technologies – a process which not only provides specified interfaces to ensure interoperability between equipment from a multiplicity of vendors, but also aims to ensure that the allocated spectrum is used as efficiently as possible, so as to provide an attractive user experience and innovative services.

The complementary functions of the regulatory authorities and the standardization organizations can be summarized broadly by the following relationship:

$$\text{Aggregated data rate} = \underbrace{\text{bandwidth}}_{\substack{\text{regulation and licences} \\ \text{(ITU-R, regional regulators)}}} \times \underbrace{\text{spectral efficiency}}_{\substack{\text{technology} \\ \text{and standards}}}$$

On a worldwide basis, ITU-R defines technology families and associates specific parts of the spectrum with these families. Facilitated by ITU-R, spectrum for mobile radio technologies is identified for the radio technologies which meet ITU-R's requirements to be designated as members of the *International Mobile Telecommunications* (IMT) family. Effectively, the IMT family comprises systems known as 'Third Generation' (for the first time providing data rates up to 2 Mbps) and beyond.

From the technology and standards angle, three main organizations have recently been developing standards relevant to IMT requirements, and these organisations continue to shape the landscape of mobile radio systems as shown in Figure 1.1.

Samsung Ex. 1008
19 of 126

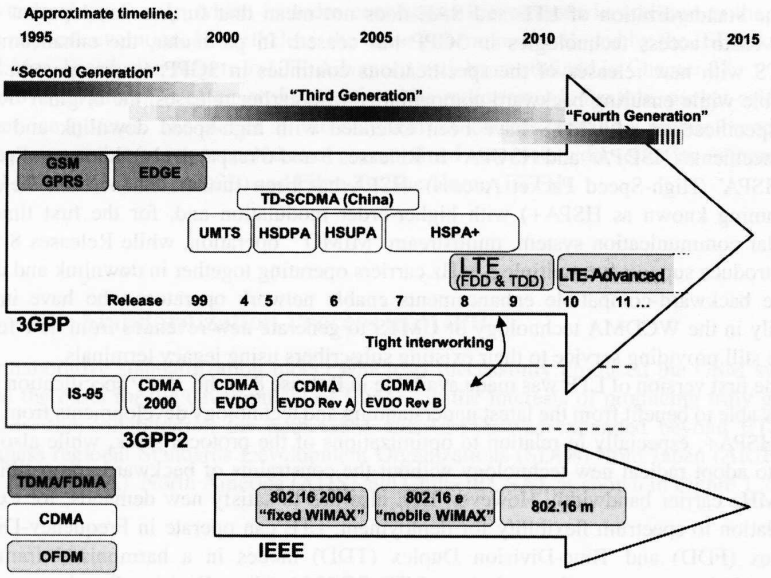


Figure 1.1: Approximate timeline of the mobile communications standards landscape.

The uppermost evolution track shown in Figure 1.1 is that developed in the 3rd Generation Partnership Project (3GPP), which is currently the dominant standards development group for mobile radio systems and is described in more detail below.

Within the 3GPP evolution track, three multiple access technologies are evident: the 'Second Generation' GSM/GPRS/EDGE family[1] was based on Time- and Frequency-Division Multiple Access (TDMA/FDMA); the 'Third Generation' UMTS family marked the entry of Code Division Multiple Access (CDMA) into the 3GPP evolution track, becoming known as *Wideband* CDMA (owing to its 5 MHz carrier bandwidth) or simply WCDMA; finally LTE has adopted Orthogonal Frequency-Division Multiplexing (OFDM), which is the access technology dominating the latest evolutions of all mobile radio standards.

In continuing the technology progression from the GSM and UMTS technology families within 3GPP, the LTE system can be seen as completing the trend of expansion of service provision beyond voice calls towards a multiservice air interface. This was already a key aim of UMTS and GPRS/EDGE, but LTE was designed from the start with the goal of evolving the radio access technology under the assumption that all services would be packet-switched, rather than following the circuit-switched model of earlier systems. Furthermore, LTE is accompanied by an evolution of the non-radio aspects of the complete system, under the term 'System Architecture Evolution' (SAE) which includes the Evolved Packet Core (EPC) network. Together, LTE and SAE comprise the Evolved Packet System (EPS), where both the core network and the radio access are fully packet-switched.

[1]The maintenance and development of specifications for the GSM family was passed to 3GPP from ETSI.

Samsung Ex. 1008
20 of 126

The standardization of LTE and SAE does not mean that further development of the other radio access technologies in 3GPP has ceased. In particular, the enhancement of UMTS with new releases of the specifications continues in 3GPP, to the greatest extent possible while ensuring backward compatibility with earlier releases: the original 'Release 99' specifications of UMTS have been extended with high-speed downlink and uplink enhancements (HSDPA[2] and HSUPA[3] in Releases 5 and 6 respectively), known collectively as 'HSPA' (High-Speed Packet Access). HSPA has been further enhanced in Release 7 (becoming known as HSPA+) with higher-order modulation and, for the first time in a cellular communication system, multistream 'MIMO'[4] operation, while Releases 8, 9 and 10 introduce support for multiple 5 MHz carriers operating together in downlink and uplink. These backward-compatible enhancements enable network operators who have invested heavily in the WCDMA technology of UMTS to generate new revenues from new features while still providing service to their existing subscribers using legacy terminals.

The first version of LTE was made available in Release 8 of the 3GPP specification series. It was able to benefit from the latest understanding and technology developments from HSPA and HSPA+, especially in relation to optimizations of the protocol stack, while also being free to adopt radical new technology without the constraints of backward compatibility or a 5 MHz carrier bandwidth. However, LTE also has to satisfy new demands, for example in relation to spectrum flexibility for deployment. LTE can operate in Frequency-Division Duplex (FDD) and Time-Division Duplex (TDD) modes in a harmonized framework designed also to support the evolution of TD-SCDMA (Time-Division Synchronous Code Division Multiple Access), which was developed in 3GPP as an additional branch of the UMTS technology path, essentially for the Chinese market.

A second version of LTE was developed in Release 9, and Release 10 continues the progression with the beginning of the next significant step known as LTE-Advanced.

A second evolution track shown in Figure 1.1 is led by a partnership organization similar to 3GPP and known as 3GPP2. CDMA2000 was developed based on the American 'IS-95' standard, which was the first mobile cellular communication system to use CDMA technology; it was deployed mainly in the USA, Korea and Japan. Standardization in 3GPP2 has continued with parallel evolution tracks towards data-oriented systems (EV-DO), to a certain extent taking a similar path to the evolutions in 3GPP. It is important to note that LTE will provide tight interworking with systems developed by 3GPP2, which allows a smooth migration to LTE for operators who previously followed the 3GPP2 track.

The third path of evolution has emerged from the IEEE 802 LAN/MAN[5] standards committee, which created the '802.16' family as a broadband wireless access standard. This family is also fully packet-oriented. It is often referred to as *WiMAX*, on the basis of a so-called 'System Profile' assembled from the 802.16 standard and promoted by the WiMAX Forum. The WiMAX Forum also ensures the corresponding product certification. While the first version, known as 802.16-2004, was restricted to fixed access, the following version 802.16e includes basic support of mobility and is therefore often referred to as 'mobile WiMAX'. However, it can be noted that in general the WiMAX family has not been designed with the same emphasis on mobility and compatibility with operators' core networks as the

[2]High-Speed Downlink Packet Access.
[3]High-Speed Uplink Packet Access.
[4]Multiple-Input Multiple-Output antenna system.
[5]Local Area Network / Metropolitan Area Network.

Samsung Ex. 1008
21 of 126

3GPP technology family, which includes core network evolutions in addition to the radio access network evolution. Nevertheless, the latest generation developed by the IEEE, known as 802.16m, has similar targets to LTE-Advanced which are outlined in Chapter 27.

The overall pattern is of an evolution of mobile radio towards flexible, packet-oriented, multiservice systems. The aim of all these systems is towards offering a mobile broadband user experience that can approach that of current fixed access networks such as Asymmetric Digital Subscriber Line (ADSL) and Fibre-To-The-Home (FTTH).

1.1.3 The Standardization Process in 3GPP

The collaborative standardization model which so successfully produced the GSM system became the basis for the development of UMTS. In the interests of producing truly global standards, the collaboration for both GSM and UMTS was expanded beyond ETSI to encompass regional Standards Development Organizations (SDOs) from Japan (ARIB and TTC), Korea (TTA), North America (ATIS) and China (CCSA), as shown in Figure 1.2.

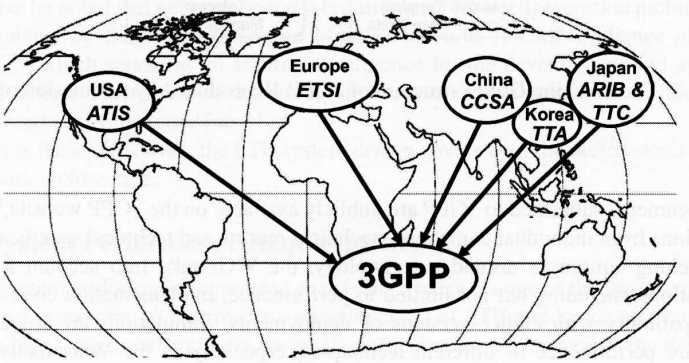


Figure 1.2: 3GPP is a global partnership of six regional SDOs.

So the 3GPP was born and by 2011 boasted 380 individual member companies.

The successful creation of such a large and complex system specification as that for UMTS or LTE requires a well-structured organization with pragmatic working procedures. 3GPP is divided into four Technical Specification Groups (TSGs), each of which is comprised of a number of Working Groups (WGs) with responsibility for a specific aspect of the specifications as shown in Figure 1.3.

A distinctive feature of the working methods of these groups is the consensus-driven approach to decision-making.

Samsung Ex. 1008
22 of 126

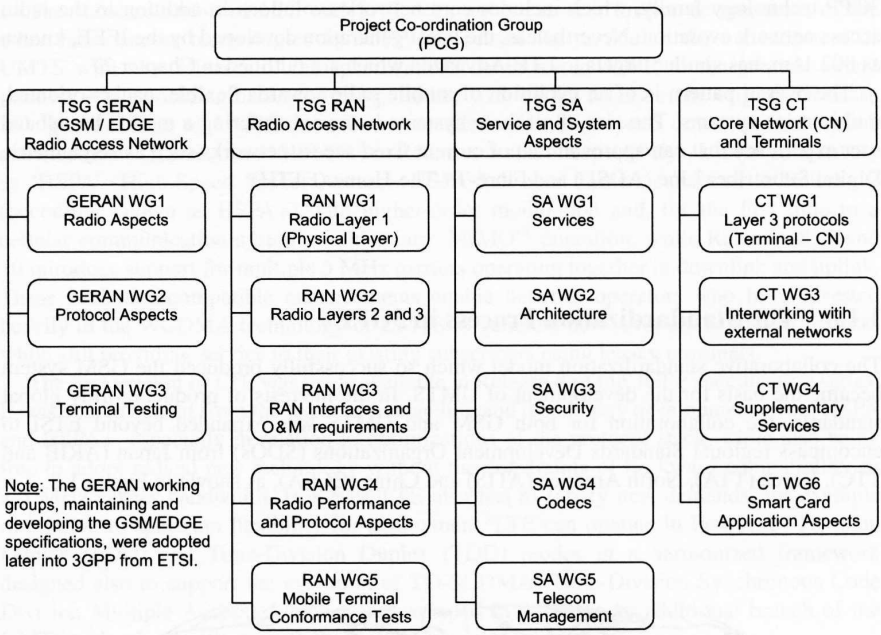


Figure 1.3: The Working Group structure of 3GPP. Reproduced by permission of © 3GPP.

All documents submitted to 3GPP are publicly available on the 3GPP website,[6] including contributions from individual companies, technical reports and technical specifications.

In reaching consensus around a technology, the WGs take into account a variety of considerations, including but not limited to performance, implementation cost, complexity and compatibility with earlier versions or deployments. Simulations are frequently used to compare performance of different techniques, especially in the WGs focusing on the physical layer of the air interface and on performance requirements. This requires consensus first to be reached around the simulation assumptions to be used for the comparison, including, in particular, understanding and defining the scenarios of interest to network operators.

The LTE standardization process was inaugurated at a workshop in Toronto in November 2004, when a broad range of companies involved in the mobile communications business presented their visions for the future evolution of the specifications to be developed in 3GPP. These visions included both initial perceptions of the *requirements* which needed to be satisfied, and proposals for *suitable technologies* to meet those requirements.

The requirements are reviewed in detail in Section 1.2, while the key technologies are introduced in Section 1.3.

[6] www.3gpp.org.

Samsung Ex. 1008
23 of 126

1.2 Requirements and Targets for the Long Term Evolution

Discussion of the key requirements for the new LTE system led to the creation of a formal 'Study Item' in 3GPP with the specific aim of 'evolving' the 3GPP radio access technology to ensure competitiveness over a ten-year time-frame. Under the auspices of this Study Item, the requirements for LTE Release 8 were refined and crystallized, being finalized in June 2005.

They can be summarized as follows:

- reduced delays, in terms of both connection establishment and transmission latency;
- increased user data rates;
- increased cell-edge bit-rate, for uniformity of service provision;
- reduced cost per bit, implying improved spectral efficiency;
- greater flexibility of spectrum usage, in both new and pre-existing bands;
- simplified network architecture;
- seamless mobility, including between different radio-access technologies;
- reasonable power consumption for the mobile terminal.

It can also be noted that network operator requirements for next generation mobile systems were formulated by the Next Generation Mobile Networks (NGMN) alliance of network operators [2], which served as an additional reference for the development and assessment of the LTE design. Such operator-driven requirements have also guided the development of LTE-Advanced (see Chapters 27 to 31).

To address these objectives, the LTE system design covers both the radio interface and the radio network architecture.

1.2.1 System Performance Requirements

Improved system performance compared to existing systems is one of the main requirements from network operators, to ensure the competitiveness of LTE and hence to arouse market interest. In this section, we highlight the main performance metrics used in the definition of the LTE requirements and its performance assessment.

Table 1.1 summarizes the main performance requirements to which the first release of LTE was designed. Many of the figures are given relative to the performance of the most advanced available version of UMTS, which at the time of the definition of the LTE requirements was HSDPA/HSUPA Release 6 – referred to here as the *reference baseline*. It can be seen that the target requirements for LTE represent a significant step from the capacity and user experience offered by the third generation mobile communications systems which were being deployed at the time when the first version of LTE was being developed.

As mentioned above, HSPA technologies are also continuing to be developed to offer higher spectral efficiencies than were assumed for the reference baseline. However, LTE has been able to benefit from avoiding the constraints of backward compatibility, enabling the inclusion of advanced MIMO schemes in the system design from the beginning, and highly flexible spectrum usage built around new multiple access schemes.

Samsung Ex. 1008
24 of 126

Table 1.1: Summary of key performance requirement targets for LTE Release 8.

		Absolute requirement	Release 6 (for comparison)	Comments
Downlink	Peak transmission rate	> 100 Mbps	14.4 Mbps	LTE in 20 MHz FDD, 2 × 2 spatial multiplexing. Reference: HSDPA in 5 MHz FDD, single antenna transmission
	Peak spectral efficiency	> 5 bps/Hz	3 bps/Hz	
	Average cell spectral efficiency	> 1.6–2.1 bps/Hz/cell	0.53 bps/Hz/cell	LTE: 2 × 2 spatial multiplexing, Interference Rejection Combining (IRC) receiver [3]. Reference: HSDPA, Rake receiver [4], 2 receive antennas
	Cell edge spectral efficiency	> 0.04–0.06 bps/Hz/user	0.02 bps/Hz/user	As above, 10 users assumed per cell
	Broadcast spectral efficiency	> 1 bps/Hz	N/A	Dedicated carrier for broadcast mode
Uplink	Peak transmission rate	> 50 Mbps	11 Mbps	LTE in 20 MHz FDD, single antenna transmission. Reference: HSUPA in 5 MHz FDD, single antenna transmission
	Peak spectral efficiency	> 2.5 bps/Hz	2 bps/Hz	
	Average cell spectral efficiency	> 0.66–1.0 bps/Hz/cell	0.33 bps/Hz/cell	LTE: single antenna transmission, IRC receiver [3]. Reference: HSUPA, Rake receiver [4], 2 receive antennas
	Cell edge spectral efficiency	> 0.02–0.03 bps/Hz/user	0.01 bps/Hz/user	As above, 10 users assumed per cell
System	User plane latency (two way radio delay)	< 10 ms		LTE target approximately one fifth of Reference.
	Connection set-up latency	< 100 ms		Idle state → active state
	Operating bandwidth	1.4–20 MHz	5 MHz	(initial requirement started at 1.25 MHz)
	VoIP capacity	NGMN preferred target expressed in [2] is > 60 sessions/MHz/cell		

The requirements shown in Table 1.1 are discussed and explained in more detail below. Chapter 26 shows how the overall performance of the LTE system meets these requirements.

1.2.1.1 Peak Rates and Peak Spectral Efficiency

For marketing purposes, the first parameter by which different radio access technologies are usually compared is the peak per-user data rate which can be achieved. This peak data rate generally scales according to the amount of spectrum used, and, for MIMO systems, according to the minimum of the number of transmit and receive antennas (see Section 11.1).

The peak data rate can be defined as the maximum throughput per user assuming the whole bandwidth being allocated to a single user with the highest modulation and coding scheme and the maximum number of antennas supported. Typical radio interface overhead (control channels, pilot signals, guard intervals, etc.) is estimated and taken into account for a given operating point. For TDD systems, the peak data rate is generally calculated for the downlink

Samsung Ex. 1008
25 of 126

and uplink periods separately. This makes it possible to obtain a single value independent of the uplink/downlink ratio and a fair system comparison that is agnostic of the duplex mode. The maximum spectral efficiency is then obtained simply by dividing the peak rate by the used spectrum allocation.

The target peak data rates for downlink and uplink in LTE Release 8 were set at 100 Mbps and 50 Mbps respectively within a 20 MHz bandwidth,[7] corresponding to respective peak spectral efficiencies of 5 and 2.5 bps/Hz. The underlying assumption here is that the terminal has two receive antennas and one transmit antenna. The number of antennas used at the base station is more easily upgradeable by the network operator, and the first version of the LTE specifications was therefore designed to support downlink MIMO operation with up to four transmit and receive antennas. The MIMO techniques enabling high peak data rates are described in detail in Chapter 11.

When comparing the capabilities of different radio communication technologies, great emphasis is often placed on the peak data rate capabilities. While this is one indicator of how technologically advanced a system is and can be obtained by simple calculations, it may not be a key differentiator in the usage scenarios for a mobile communication system in practical deployment. Moreover, it is relatively easy to design a system that can provide very high peak data rates for users close to the base station, where interference from other cells is low and techniques such as MIMO can be used to their greatest extent. It is much more challenging to provide high data rates with good coverage and mobility, but it is exactly these latter aspects which contribute most strongly to user satisfaction.

In typical deployments, individual users are located at varying distances from the base stations, the propagation conditions for radio signals to individual users are rarely ideal, and the available resources must be shared between many users. Consequently, although the claimed peak data rates of a system are genuinely achievable in the right conditions, it is rare for a single user to be able to experience the peak data rates for a sustained period, and the envisaged applications do not usually require this level of performance.

A differentiator of the LTE system design compared to some other systems has been the recognition of these 'typical deployment constraints' from the beginning. During the design process, emphasis was therefore placed not only on providing a competitive peak data rate for use when conditions allow, but also importantly on *system level performance*, which was evaluated during several performance verification steps.

System-level evaluations are based on simulations of multicell configurations where data transmission from/to a population of mobiles is considered in a typical deployment scenario. The sections below describe the main metrics used as requirements for system level performance. In order to make these metrics meaningful, parameters such as the deployment scenario, traffic models, channel models and system configuration need to be defined.

The key definitions used for the system evaluations of LTE Release 8 can be found in an input document from network operators addressing the performance verification milestone in the LTE development process [5]. This document takes into account deployment scenarios and channel models agreed during the LTE Study Item [6], and is based on an evaluation methodology elaborated by NGMN operators in [7]. The reference deployment scenarios which were given special consideration for the LTE performance evaluation covered macrocells with base station separations of between 500 m and 1.7 km, as well as microcells using MIMO with base station separations of 130 m. A range of mobile terminal

[7] Four times the bandwidth of a WCDMA carrier.

Samsung Ex. 1008
26 of 126

speeds were studied, focusing particularly on the range 3–30 km/h, although higher mobile speeds were also considered important.

1.2.1.2 Cell Throughput and Spectral Efficiency

Performance at the cell level is an important criterion, as it relates directly to the number of cell sites that a network operator requires, and hence to the capital cost of deploying the system. For LTE Release 8, it was chosen to assess the cell level performance with full-queue traffic models (i.e. assuming that there is never a shortage of data to transmit if a user is given the opportunity) and a relatively high system load, typically 10 users per cell.

The requirements at the cell level were defined in terms of the following metrics:

- Average cell throughput [bps/cell] and spectral efficiency [bps/Hz/cell];
- Aaverage user throughput [bps/user] and spectral efficiency [bps/Hz/user];
- Cell-edge user throughput [bps/user] and spectral efficiency [bps/Hz/user] (the metric used for this assessment is the 5-percentile user throughput, obtained from the cumulative distribution function of the user throughput).

For the UMTS Release 6 reference baseline, it was assumed that both the terminal and the base station use a single transmit antenna and two receive antennas; for the terminal receiver the assumed performance corresponds to a two-branch Rake receiver [4] with linear combining of the signals from the two antennas.

For the LTE system, the use of two transmit and receive antennas was assumed at the base station. At the terminal, two receive antennas were assumed, but still only a single transmit antenna. The receiver for both downlink and uplink is assumed to be a linear receiver with optimum combining of the signals from the antenna branches [3].

The original requirements for the cell level metrics were only expressed as relative gains compared to the Release 6 reference baseline. The absolute values provided in Table 1.1 are based on evaluations of the reference system performance that can be found in [8] and [9] for downlink and uplink respectively.

1.2.1.3 Voice Capacity

Unlike full queue traffic (such as file download) which is typically delay-tolerant and does not require a guaranteed bit-rate, real-time traffic such as Voice over IP (VoIP) has tight delay constraints. It is important to set system capacity requirements for such services – a particular challenge in fully packet-based systems like LTE which rely on adaptive scheduling.

The system capacity requirement is defined as the number of satisfied VoIP users, given a particular traffic model and delay constraints. The details of the traffic model used for evaluating LTE can be found in [5]. Here, a VoIP user is considered to be in outage (i.e. not satisfied) if more than 2% of the VoIP packets do not arrive successfully at the radio receiver within 50 ms and are therefore discarded. This assumes an overall end-to-end delay (from mobile terminal to mobile terminal) below 200 ms. The system capacity for VoIP can then be defined as the number of users present per cell when more than 95% of the users are satisfied.

The NGMN group of network operators expressed a preference for the ability to support 60 satisfied VoIP sessions per MHz – an increase of two to four times what can typically be achieved in the Release 6 reference case.

Samsung Ex. 1008
27 of 126

1.2.1.4 Mobility and Cell Ranges

LTE is required to support communication with terminals moving at speeds of up to 350 km/h, or even up to 500 km/h depending on the frequency band. The primary scenario for operation at such high speeds is usage on high-speed trains – a scenario which is increasing in importance across the world as the number of high-speed rail lines increases and train operators aim to offer an attractive working environment to their passengers. These requirements mean that handover between cells has to be possible without interruption – in other words, with imperceptible delay and packet loss for voice calls, and with reliable transmission for data services.

These targets are to be achieved by the LTE system in typical cells of radius up to 5 km, while operation should continue to be possible for cell ranges of 100 km and more, to enable wide-area deployments.

1.2.1.5 Broadcast Mode Performance

The requirements for LTE included the integration of an efficient broadcast mode for high rate Multimedia Broadcast/Multicast Services (MBMS) such as mobile TV, based on a Single Frequency Network mode of operation as explained in detail in Chapter 13. The spectral efficiency requirement is given in terms of a carrier dedicated to broadcast transmissions – i.e. not shared with unicast transmissions.

In broadcast systems, the system throughput is limited to what is achievable for the users in the worst conditions. Consequently, the broadcast performance requirement was defined in terms of an achievable system throughput (bps) and spectral efficiency (bps/Hz) assuming a coverage of 98% of the nominal coverage area of the system. This means that only 2% of the locations in the nominal coverage area are in outage – where outage for broadcast services is defined as experiencing a packet error rate higher than 1%. This broadcast spectral efficiency requirement was set to 1 bps/Hz [10].

While the broadcast mode was not available in Release 8 due to higher prioritization of other service modes, Release 9 incorporates a broadcast mode employing Single Frequency Network operation on a mixed unicast-broadcast carrier.

1.2.1.6 User Plane Latency

User plane latency is an important performance metric for real-time and interactive services. On the radio interface, the minimum user plane latency can be calculated based on signalling analysis for the case of an unloaded system. It is defined as the average time between the first transmission of a data packet and the reception of a physical layer acknowledgement. The calculation should include typical HARQ[8] retransmission rates (e.g. 0–30%). This definition therefore considers the capability of the system design, without being distorted by the scheduling delays that would appear in the case of a loaded system. The round-trip latency is obtained simply by multiplying the one-way user plane latency by a factor of two.

LTE is also required to be able to operate with an IP-layer one-way data-packet latency across the radio access network as low as 5 ms in optimal conditions. However, it is recognized that the actual delay experienced in a practical system will be dependent on system loading and radio propagation conditions. For example, HARQ plays a key role in

[8]Hybrid Automatic Repeat reQuest – see Section 10.3.2.5.

Samsung Ex. 1008
28 of 126

maximizing spectral efficiency at the expense of increased delay while retransmissions take place, whereas maximal spectral efficiency may not be essential in situations when minimum latency is required.

1.2.1.7 Control Plane Latency and Capacity

In addition to the user plane latency requirement, call setup delay was required to be significantly reduced compared to previous cellular systems. This not only enables a good user experience but also affects the battery life of terminals, since a system design which allows a fast transition from an idle state to an active state enables terminals to spend more time in the low-power idle state.

Control plane latency is measured as the time required for performing the transitions between different LTE states. LTE is based on only two main states, 'RRC_IDLE' and 'RRC_CONNECTED' (i.e. 'active') (see Section 3.1).

LTE is required to support transition from idle to active in less than 100 ms (excluding paging delay and Non-Access Stratum (NAS) signalling delay).

The LTE system capacity is dependent not only on the supportable throughput but also on the number of users simultaneously located within a cell which can be supported by the control signalling. For the latter aspect, LTE is required to support at least 200 active-state users per cell for spectrum allocations up to 5 MHz, and at least 400 users per cell for wider spectrum allocations; only a small subset of these users would be actively receiving or transmitting data at any given time instant, depending, for example, on the availability of data to transmit and the prevailing radio channel conditions. An even larger number of non-active users may also be present in each cell, and therefore able to be paged or to start transmitting data with low latency.

1.2.2 Deployment Cost and Interoperability

Besides the system performance aspects, a number of other considerations are important for network operators. These include reduced deployment cost, spectrum flexibility and enhanced interoperability with legacy systems – essential requirements to enable deployment of LTE networks in a variety of scenarios and to facilitate migration to LTE.

1.2.2.1 Spectrum Allocations and Duplex Modes

As demand for suitable radio spectrum for mobile communications increases, LTE is required to be able to operate in a wide range of frequency bands and sizes of spectrum allocations in both uplink and downlink. LTE can use spectrum allocations ranging from 1.4 to 20 MHz with a single carrier and addresses all frequency bands currently identified for IMT systems by ITU-R [1] including those below 1 GHz.

This will include deploying LTE in spectrum currently occupied by older radio access technologies – a practice often known as 'spectrum refarming'.

New frequency bands are continually being introduced for LTE in a release-independent way, meaning that any of the LTE Releases can be deployed in a new frequency band once the Radio-Frequency (RF) requirements have been specified [11].

The ability to operate in both paired and unpaired spectrum is required, depending on spectrum availability (see Chapter 23). LTE provides support for FDD, TDD and half-duplex

Samsung Ex. 1008
29 of 126

FDD operation in a unified design, ensuring a high degree of commonality which facilitates implementation of multimode terminals and allows worldwide roaming.

Starting from Release 10, LTE also provides means for flexible spectrum use via aggregation of contiguous and non-contiguous spectrum assets for high data rate services using a total bandwidth of up to 100 MHz (see Chapter 28).

1.2.2.2 Inter-Working with Other Radio Access Technologies

Flexible interoperation with other radio access technologies is essential for service continuity, especially during the migration phase in early deployments of LTE with partial coverage, where handover to legacy systems will often occur.

LTE relies on an evolved packet core network which allows interoperation with various access technologies, in particular earlier 3GPP technologies (GSM/EDGE and UTRAN[9]) as well as non-3GPP technologies (e.g. WiFi, CDMA2000 and WiMAX).

However, service continuity and short interruption times can only be guaranteed if measurements of the signals from other systems and fast handover mechanisms are integrated in the LTE radio access design. LTE therefore supports tight inter-working with all legacy 3GPP technologies and some non-3GPP technologies such as CDMA2000.

1.2.2.3 Terminal Complexity and Cost

A key consideration for competitive deployment of LTE is the availability of low-cost terminals with long battery life, both in stand-by and during activity. Therefore, low terminal complexity has been taken into account where relevant throughout the LTE system, as well as designing the system wherever possible to support low terminal power consumption.

1.2.2.4 Network Architecture Requirements

LTE is required to allow a cost-effective deployment by an improved radio access network architecture design including:

- Flat architecture consisting of just one type of node, the base station, known in LTE as the *eNodeB* (see Chapter 2);
- Effective protocols for the support of packet-switched services (see Chapters 3 to 4);
- Open interfaces and support of multivendor equipment interoperability;
- efficient mechanisms for operation and maintenance, including self-optimization functionalities (see Chapter 25);
- Support of easy deployment and configuration, for example for so-called home base stations (otherwise known as femto-cells) (see Chapter 24).

[9]Universal Terrestrial Radio Access Network.

Samsung Ex. 1008
30 of 126

1.3 Technologies for the Long Term Evolution

The fulfilment of the extensive range of requirements outlined above is only possible thanks to advances in the underlying mobile radio technology. As an overview, we outline here three fundamental technologies that have shaped the LTE radio interface design: *multicarrier* technology, *multiple-antenna* technology, and the application of *packet-switching* to the radio interface. Finally, we summarize the combinations of capabilities that are supported by different categories of LTE mobile terminal in Releases 8 and 9.

1.3.1 Multicarrier Technology

Adopting a multicarrier approach for multiple access in LTE was the first major design choice. After initial consolidation of proposals, the candidate schemes for the downlink were Orthogonal Frequency-Division Multiple Access (OFDMA)[10] and Multiple WCDMA, while the candidate schemes for the uplink were Single-Carrier Frequency-Division Multiple Access (SC-FDMA), OFDMA and Multiple WCDMA. The choice of multiple-access schemes was made in December 2005, with OFDMA being selected for the downlink, and SC-FDMA for the uplink. Both of these schemes open up the frequency domain as a new dimension of flexibility in the system, as illustrated schematically in Figure 1.4.

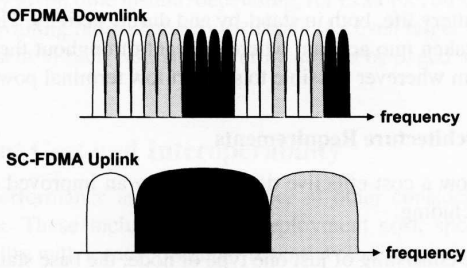


Figure 1.4: Frequency-domain view of the LTE multiple-access technologies.

OFDMA extends the multicarrier technology of OFDM to provide a very flexible multiple-access scheme. OFDM subdivides the bandwidth available for signal transmission into a multitude of narrowband subcarriers, arranged to be mutually orthogonal, which either individually or in groups can carry independent information streams; in OFDMA, this subdivision of the available bandwidth is exploited in sharing the subcarriers among multiple users.[11]

This resulting flexibility can be used in various ways:

[10] OFDM technology was already well understood in 3GPP as a result of an earlier study of the technology in 2003–4.

[11] The use of the frequency domain comes in addition to the well-known time-division multiplexing which continues to play an important role in LTE.

Samsung Ex. 1008
31 of 126

- Different spectrum bandwidths can be utilized without changing the fundamental system parameters or equipment design;
- Transmission resources of variable bandwidth can be allocated to different users and scheduled freely in the frequency domain;
- Fractional frequency re-use and interference coordination between cells are facilitated.

Extensive experience with OFDM has been gained in recent years from deployment of digital audio and video broadcasting systems such as DAB, DVB and DMB.[12] This experience has highlighted some of the key advantages of OFDM, which include:

- Robustness to time-dispersive radio channels, thanks to the subdivision of the wide-band transmitted signal into multiple narrowband subcarriers, enabling inter-symbol interference to be largely constrained within a guard interval at the beginning of each symbol;
- Low-complexity receivers, by exploiting frequency-domain equalization;
- Simple combining of signals from multiple transmitters in broadcast networks.

These advantages, and how they arise from the OFDM signal design, are explained in detail in Chapter 5.

By contrast, the transmitter design for OFDM is more costly, as the Peak-to-Average Power Ratio (PAPR) of an OFDM signal is relatively high, resulting in a need for a highly-linear RF power amplifier. However, this limitation is not inconsistent with the use of OFDM for *downlink* transmissions, as low-cost implementation has a lower priority for the base station than for the mobile terminal.

In the uplink, however, the high PAPR of OFDM is difficult to tolerate for the transmitter of the mobile terminal, since it is necessary to compromise between the output power required for good outdoor coverage, the power consumption, and the cost of the power amplifier. SC-FDMA, which is explained in detail in Chapter 14, provides a multiple-access technology which has much in common with OFDMA – in particular the flexibility in the frequency domain, and the incorporation of a guard interval at the start of each transmitted symbol to facilitate low-complexity frequency-domain equalization at the receiver. At the same time, SC-FDMA has a significantly lower PAPR. It therefore resolves to some extent the dilemma of how the uplink can benefit from the advantages of multicarrier technology while avoiding excessive cost for the mobile terminal transmitter and retaining a reasonable degree of commonality between uplink and downlink technologies.

In Release 10, the uplink multiple access scheme is extended to allow multiple clusters of subcarriers in the frequency domain, as explained in Section 28.3.6.

1.3.2 Multiple Antenna Technology

The use of multiple antenna technology allows the exploitation of the spatial-domain as another new dimension. This becomes essential in the quest for higher spectral efficiencies. As will be detailed in Chapter 11, with the use of multiple antennas the theoretically achievable spectral efficiency scales linearly with the minimum of the number of transmit and receive antennas employed, at least in suitable radio propagation environments.

[12] Digital Audio Broadcasting, Digital Video Broadcasting and Digital Mobile Broadcasting.

Samsung Ex. 1008
32 of 126

Multiple antenna technology opens the door to a large variety of features, but not all of them easily deliver their theoretical promises when it comes to implementation in practical systems. Multiple antennas can be used in a variety of ways, mainly based on three fundamental principles, schematically illustrated in Figure 1.5:

- **Diversity gain.** Use of the spatial diversity provided by the multiple antennas to improve the robustness of the transmission against multipath fading.
- **Array gain.** Concentration of energy in one or more given directions via precoding or beamforming. This also allows multiple users located in different directions to be served simultaneously (so-called multi-user MIMO).
- **Spatial multiplexing gain.** Transmission of multiple signal streams to a single user on multiple spatial layers created by combinations of the available antennas.

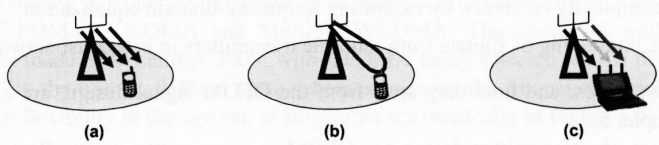


Figure 1.5: Three fundamental benefits of multiple antennas:
(a) diversity gain; (b) array gain; (c) spatial multiplexing gain.

A large part of the LTE Study Item phase was therefore dedicated to the selection and design of the various multiple antenna features to be included in the first release of LTE. The final system includes several complementary options which allow for adaptability according to the network deployment and the propagation conditions of the different users.

1.3.3 Packet-Switched Radio Interface

As has already been noted, LTE has been designed as a completely packet-oriented multi-service system, without the reliance on circuit-switched connection-oriented protocols prevalent in its predecessors. In LTE, this philosophy is applied across all the layers of the protocol stack.

The route towards fast packet scheduling over the radio interface was already opened by HSDPA, which allowed the transmission of short packets having a duration of the same order of magnitude as the coherence time of the fast fading channel, as shown in Figure 1.6. This calls for a joint optimization of the physical layer configuration and the resource management carried out by the link layer protocols according to the prevailing propagation conditions. This aspect of HSDPA involves tight coupling between the lower two layers of the protocol stack – the MAC (Medium Access Control layer – see Chapter 4) and the physical layer.

In HSDPA, this coupling already included features such as fast channel state feedback, dynamic link adaptation, scheduling exploiting multi-user diversity, and fast retransmission protocols. In LTE, in order to improve the system latency, the packet duration was further reduced from the 2 ms used in HSDPA down to just 1 ms. This short transmission interval,

Samsung Ex. 1008
33 of 126

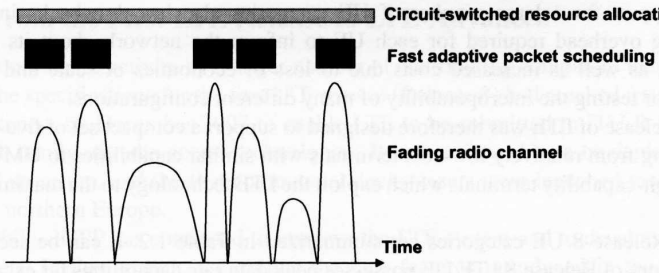


Figure 1.6: Fast scheduling and link adaptation.

together with the new dimensions of frequency and space, has further extended the field of cross-layer techniques between the MAC and physical layers to include the following techniques in LTE:

- Adaptive scheduling in both the frequency and spatial dimensions;
- Adaptation of the MIMO configuration including the selection of the number of spatial layers transmitted simultaneously;
- Link adaptation of modulation and code-rate, including the number of transmitted codewords;
- Several modes of fast channel state reporting.

These different levels of optimization are combined with very sophisticated control signalling.

1.3.4 User Equipment Categories

In practice it is important to recognize that the market for UEs is large and diverse, and there is therefore a need for LTE to support a range of categories of UE with different capabilities to satisfy different market segments. In general, each market segment attaches different priorities to aspects such as peak data rate, UE size, cost and battery life. Some typical trade-offs include the following:

- Support for the highest data rates is key to the success of some applications, but generally requires large amounts of memory for data processing, which increases the cost of the UE.
- UEs which may be embedded in large devices such as laptop computers are often not significantly constrained in terms of acceptable power consumption or the number of antennas which may be used; on the other hand, other market segments require ultra-slim hand-held terminals which have little space for multiple antennas or large batteries.

The wider the range of UE categories supported, the closer the match which may be made between a UE's supported functionality and the requirements of a particular market segment.

Samsung Ex. 1008
34 of 126

However, support for a large number of UE categories also has drawbacks in terms of the signalling overhead required for each UE to inform the network about its supported functionality, as well as increased costs due to loss of economies of scale and increased complexity for testing the interoperability of many different configurations.

The first release of LTE was therefore designed to support a compact set of five categories of UE, ranging from relatively low-cost terminals with similar capabilities to UMTS HSPA, up to very high-capability terminals which exploit the LTE technology to the maximum extent possible.

The five Release 8 UE categories are summarized in Table 1.2. It can be seen that the highest category of Release 8 LTE UE possesses peak data rate capabilities far exceeding the LTE Release 8 targets. Full details are specified in [12].

Table 1.2: Categories of LTE user equipment in Releases 8 and 9.

	UE category				
	1	2	3	4	5
Supported downlink data rate (Mbps)	10	50	100	150	300
Supported uplink data rate (Mbps)	5	25	50	50	75
Number of receive antennas required	2	2	2	2	4
Number of downlink MIMO layers supported	1	2	2	2	4
Support for 64QAM modulation in downlink	✔	✔	✔	✔	✔
Support for 64QAM modulation in uplink	✖	✖	✖	✖	✔
Relative memory requirement for physical layer processing (normalized to category 1 level)	1	4.9	4.9	7.3	14.6

Additional UE categories are introduced in Release 10, and these are explained in Section 27.5.

The LTE specifications deliberately avoid large numbers of optional features for the UEs, preferring to take the approach that if a feature is sufficiently useful to be worth including in the specifications then support of it should be mandatory. Nevertheless, a very small number of optional Release 8 features, whose support is indicated by each UE by specific signalling, are listed in [12]; such features are known as 'UE capabilities'. Some additional UE capabilities are added in later releases.

In addition, it is recognized that it is not always possible to complete conformance testing and Inter-Operability Testing (IOT) of every mandatory feature simultaneously for early deployments of LTE. Therefore, the development of conformance test cases for LTE was prioritized according to the likelihood of early deployment of each feature. Correspondingly, *Feature Group Indicators* (FGIs) are used for certain groups of lower priority mandatory features, to enable a UE to indicate whether IOT has been successfully completed for those features; the grouping of features corresponding to each FGI can be found in Annex B.1 of [13]. For UEs of Release 9 and later, it becomes mandatory for certain of these FGIs to be set to indicate that the corresponding feature(s) have been implemented and successfully tested.

Samsung Ex. 1008
35 of 126

1.3.5 From the First LTE Release to LTE-Advanced

As a result of intense activity by a larger number of contributing companies than ever before in 3GPP, the specifications for the first LTE release (Release 8) had reached a sufficient level of completeness by December 2007 to enable LTE to be submitted to ITU-R as a member of the IMT family of radio access technologies. It is therefore able to be deployed in IMT-designated spectrum, and the first commercial deployments were launched towards the end of 2009 in northern Europe.

Meanwhile, 3GPP has continued to improve the LTE system and to develop it to address new markets. In this section, we outline the new features introduced in the second LTE release, Release 9, and those provided by LTE Release 10, which begins the next significant step known as LTE-Advanced.

Increasing LTE's suitability for different markets and deployments was the first goal of Release 9. One important market with specific regulatory requirements is North America. LTE Release 9 therefore provides improved support for Public Warning Systems (PWS) and some accurate positioning methods (see Chapter 19). One positioning method uses the Observed Time Difference of Arrival (OTDOA) principle, supported by specially designed new reference signals inserted in the LTE downlink transmissions. Measurements of these positioning reference signals received from different base stations allow a UE to calculate its position very accurately, even in locations where other positioning means such as GPS fail (e.g. indoors). Enhanced Cell-ID-based techniques are also supported.

Release 9 also introduces support for a broadcast mode based on Single Frequency Network type transmissions (see Chapter 13).

The MIMO transmission modes are further developed in Release 9, with an extension of the Release 8 beamforming mode to support two orthogonal spatial layers that can be transmitted to a single user or multiple users, as described in Section 11.2.2.3. The design of this mode is forward-compatible for extension to more than two spatial layers in Release 10.

Release 9 also addresses specific deployments and, in particular, low power nodes (see Chapter 24). It defines new requirements for pico base stations and home base stations, in addition to improving support for Closed Subscriber Groups (CSG). Support for self-optimization of the networks is also enhanced in Release 9, as described in Chapter 25.

1.3.5.1 LTE-Advanced

The next version of LTE, Release 10, develops LTE to LTE-Advanced. While LTE Releases 8 and 9 already satisfy to a large extent the requirements set by ITU-R for the IMT-Advanced designation [14] (see Section 27.1), Release 10 will fully satisfy them and even exceed them in several aspects where 3GPP has set more demanding performance targets than those of ITU-R. The requirements for LTE-Advanced are discussed in detail in Chapter 27.

The main Release 10 features that are directly related to fulfilment of the IMT-Advanced requirements are:

- Carrier aggregation, allowing the total transmission bandwidth to be increased up to 100 MHz (see Chapter 28);
- Uplink MIMO transmission for peak spectral efficiencies greater than 7.5 bps/Hz and targeting up to 15 bps/Hz (see Chapter 29);

Samsung Ex. 1008
36 of 126

- Downlink MIMO enhancements, targeting peak spectral efficiencies up to 30 bps/Hz (see Chapter 29).

Besides addressing the IMT-Advanced requirements, Release 10 also provides some new features to enhance LTE deployment, such as support for relaying (see Chapter 30), enhanced inter-cell interference coordination (see Chapter 31) and mechanisms to minimize the need for drive tests by supporting extended measurement reports from the terminals (see Chapters 25 and 31).

1.4 From Theory to Practice

With commercial deployment of LTE now a reality, the advances in theoretical understanding and technology which underpin the LTE specifications are being exploited practically. This book is written with the primary aim of illuminating the transition from the underlying academic progress to the realization of useful advances in the provision of mobile communication services. Particular focus is given to the physical layer of the Radio Access Network (RAN), as it is here that many of the most dramatic technical advances are manifested. This should enable the reader to develop an understanding of the background to the technology choices in the LTE system, and hence to understand better the LTE specifications and how they may be implemented.

Parts I to IV of the book describe the features of LTE Releases 8 and 9, including indications of the aspects that are further enhanced in Release 10, while the details of the major new features of Release 10 are explained in Part V.

Part I sets the radio interface in the context of the network architecture and protocols, including radio resource management aspects, as well as explaining the new developments in these areas which distinguish LTE from previous systems.

In Part II, the physical layer of the RAN downlink is covered in detail, beginning with an explanation of the theory of the new downlink multiple access technology, OFDMA, in Chapter 5. This sets the context for the details of the LTE downlink design in Chapters 6 to 9. As coding, link adaptation and multiple antenna operation are of fundamental importance in fulfilling the LTE requirements, two chapters are then devoted to these topics, covering both the theory and the practical implementation in LTE.

Chapter 12 shows how these techniques can be applied to the system-level operation of the LTE system, focusing on applying the new degrees of freedom to multi-user scheduling and interference coordination.

Finally for the downlink, Chapter 13 covers broadcast operation – a mode which has its own unique challenges in a cellular system but which is nonetheless important in enabling a range of services to be provided to the end user.

Part III addresses the physical layer of the RAN uplink, beginning in Chapter 14 with an introduction to the theory behind the new uplink multiple access technology, SC-FDMA. This is followed in Chapters 15 to 18 with an analysis of the detailed uplink structure and operation, including the design of the associated procedures for random access, timing control and power control which are essential to the efficient operation of the uplink.

This leads on to Part IV, which examines a number of aspects of LTE related to its deployment as a mobile cellular system. Chapter 19 explains the UE positioning techniques introduced in Release 9. Chapter 20 provides a thorough analysis of the characteristics

Samsung Ex. 1008
37 of 126

of the radio propagation environments in which LTE systems will be deployed, since an understanding of the propagation environment underpins much of the technology adopted for the LTE specifications. The new technologies and bandwidths adopted in LTE also have implications for the radio-frequency implementation of the mobile terminals in particular, and some of these are analysed in Chapter 21. The LTE system is designed to operate not just in wide bandwidths but also in a diverse range of spectrum allocation scenarios, and Chapter 23 therefore addresses the different duplex modes applicable to LTE and the effects that these may have on system design and operation. Chapter 24 addresses aspects of special relevance to deployment of low-power base stations such as Home eNodeBs and picocells, while Chapter 25 explains the advanced techniques for self-optimization of the network. Part IV concludes with a dedicated chapter examining a wide range of aspects of the overall system performance achievable with the first release of LTE.

Finally, Part V explains in detail the major new features included in Release 10 for LTE-Advanced, as 3GPP continues to respond to the ever-higher expectations of end-users. Chapters 28 to 30 address the technologies of carrier aggregation, enhanced MIMO and relaying respectively, and Chapter 31 covers enhanced Inter-Cell Interference Coordination, Minimization of Drive Tests and Machine-Type Communications. Chapter 32 provides an evaluation of the system performance achievable with LTE-Advanced Release 10, and concludes with a further look into the future.

References[13]

[1] ITU, International Telecommunications Union, www.itu.int/itu-r.

[2] NGMN, 'Next Generation Mobile Networks Beyond HSPA & EVDO – A white paper', www.ngmn.org, December 2006.

[3] J. H. Winters, 'Optimum Combining in Digital Mobile Radio with Cochannel Interference'. *IEEE Journal on Selected Areas in Communications*, Vol. 2, July 1984.

[4] R. Price and P. E. Green, 'A Communication Technique for Multipath Channels' in *Proceedings of the IRE*, Vol. 46, March 1958.

[5] Orange, China Mobile, KPN, NTT DoCoMo, Sprint, T-Mobile, Vodafone, and Telecom Italia, 'R1-070674: LTE Physical Layer Framework for Performance Verification', www.3gpp.org, 3GPP TSG RAN WG1, meeting 48, St Louis, USA, February 2007.

[6] 3GPP Technical Report 25.814, 'Physical Layer Aspects for Evolved UTRA', www.3gpp.org.

[7] NGMN, 'Next Generation Mobile Networks Radio Access Performance Evaluation Methodology', www.ngmn.org, June 2007.

[8] Ericsson, 'R1-072578: Summary of Downlink Performance Evaluation', www.3gpp.org, 3GPP TSG RAN WG1, meeting 49, Kobe, Japan, May 2007.

[9] Nokia, 'R1-072261: LTE Performance Evaluation – Uplink Summary', www.3gpp.org, 3GPP TSG RAN WG1, meeting 49, Kobe, Japan, May 2007.

[10] 3GPP Technical Report 25.913, 'Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN)', www.3gpp.org.

[11] 3GPP Technical Specification 36.307, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Requirements on User Equipments (UEs) supporting a release-independent frequency band', www.3gpp.org.

[13] All web sites confirmed 1st March 2011.

Samsung Ex. 1008
38 of 126

[12] 3GPP Technical Specification 36.306, 'Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio access capabilities', www.3gpp.org.

[13] 3GPP Technical Specification 36.331, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification', www.3gpp.org.

[14] ITU-R Report M.2134, 'Requirements related to technical performance for IMT-Advanced radio interface(s)', www.itu.int/itu-r.

Samsung Ex. 1008
39 of 126

2

Network Architecture

Sudeep Palat and Philippe Godin

2.1 Introduction

As mentioned in the preceding chapter, LTE has been designed to support only Packet-Switched (PS) services, in contrast to the Circuit-Switched (CS) model of previous cellular systems. It aims to provide seamless Internet Protocol (IP) connectivity between User Equipment (UE) and the Packet Data Network (PDN), without any disruption to the end users' applications during mobility. While the term 'LTE' encompasses the evolution of the radio access through the Evolved-UTRAN[1] (E-UTRAN), it is accompanied by an evolution of the non-radio aspects under the term 'System Architecture Evolution' (SAE) which includes the Evolved Packet Core (EPC) network. Together LTE and SAE comprise the Evolved Packet System (EPS).

EPS uses the concept of *EPS bearers* to route IP traffic from a gateway in the PDN to the UE. A bearer is an IP packet flow with a defined Quality of Service (QoS). The E-UTRAN and EPC together set up and release bearers as required by applications. EPS natively supports voice services over the IP Multimedia Subsystem (IMS) using Voice over IP (VoIP), but LTE also supports interworking with legacy systems for traditional CS voice support.

This chapter presents the overall EPS network architecture, giving an overview of the functions provided by the Core Network (CN) and E-UTRAN. The protocol stack across the different interfaces is then explained, along with an overview of the functions provided by the different protocol layers. Section 2.4 outlines the end-to-end bearer path including QoS aspects, provides details of a typical procedure for establishing a bearer and discusses the inter-working with legacy systems for CS voice services. The remainder of the chapter presents the network interfaces in detail, with particular focus on the E-UTRAN interfaces

[1] Universal Terrestrial Radio Access Network.

LTE – The UMTS Long Term Evolution: From Theory to Practice, Second Edition.
Stefania Sesia, Issam Toufik and Matthew Baker.
© 2011 John Wiley & Sons, Ltd. Published 2011 by John Wiley & Sons, Ltd.

and associated procedures, including those for the support of user mobility. The network elements and interfaces used solely to support broadcast services are covered in Chapter 13, and the aspects related to UE positioning in Chapter 19.

2.2 Overall Architectural Overview

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

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

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

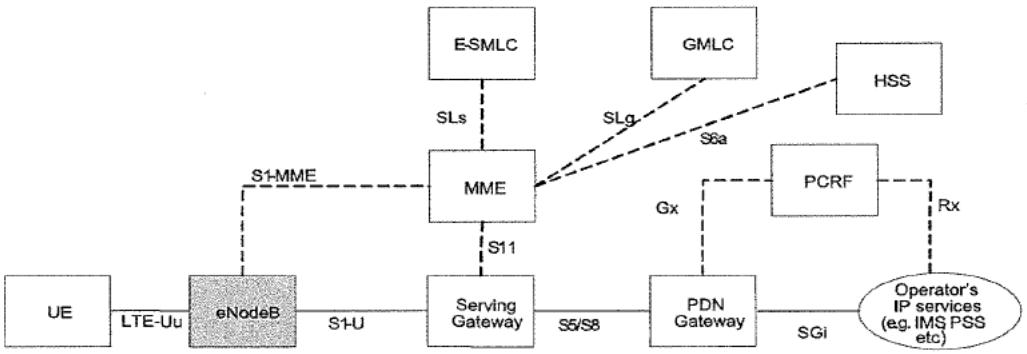


Figure 2.1: The EPS network elements.

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

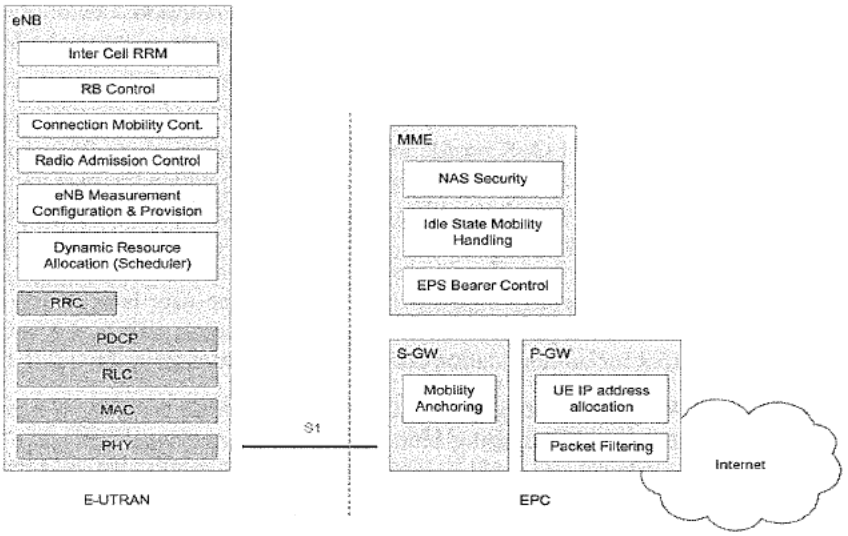


Figure 2.2: Functional split between E-UTRAN and EPC.
Reproduced by permission of © 3GPP.

2.2.1 The Core Network

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

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

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

- **PCRF.** The PCRF is responsible for policy control decision-making, as well as for controlling the flow-based charging functionalities in the Policy Control Enforcement Function (PCEF) which resides in the P-GW. The PCRF provides the QoS authorization (QoS class identifier and bit rates) that decides how a certain data flow will be treated in the PCEF and ensures that this is in accordance with the user's subscription profile.

- **GMLC.** The GMLC contains functionalities required to support LoCation Services (LCS). After performing authorization, it sends positioning requests to the MME and receives the final location estimates.

- **Home Subscriber Server (HSS).** The HSS contains users' SAE subscription data such as the EPS-subscribed QoS profile and any access restrictions for roaming (see Section 2.2.3). It also holds information about the PDNs to which the user can connect. This could be in the form of an Access Point Name (APN) (which is a label according to DNS[2] naming conventions describing the access point to the PDN), or a PDN Address (indicating subscribed IP address(es)). In addition, the HSS holds dynamic information such as the identity of the MME to which the user is currently attached or registered. The HSS may also integrate the Authentication Centre (AuC) which generates the vectors for authentication and security keys (see Section 3.2.3.1).

- **P-GW.** The P-GW is responsible for IP address allocation for the UE, as well as QoS enforcement and flow-based charging according to rules from the PCRF. The P-GW is responsible for the filtering of downlink user IP packets into the different QoS-based bearers. This is performed based on Traffic Flow Templates (TFTs) (see Section 2.4). The P-GW performs QoS enforcement for Guaranteed Bit Rate (GBR) bearers. It also serves as the mobility anchor for inter-working with non-3GPP technologies such as CDMA2000 and WiMAX networks (see Section 2.4.2 and Chapter 22 for more information about mobility).

- **S-GW.** All user IP packets are transferred through the S-GW, which serves as the local mobility anchor for the data bearers when the UE moves between eNodeBs. It also retains the information about the bearers when the UE is in idle state (known as EPS Connection Management IDLE (ECM-IDLE), see Section 2.2.1.1) and temporarily buffers downlink data while the MME initiates paging of the UE to re-establish the bearers. In addition, the S-GW performs some administrative functions in the visited network, such as collecting information for charging (e.g. the volume of data sent to or received from the user) and legal interception. It also serves as the mobility anchor for inter-working with other 3GPP technologies such as GPRS[3] and UMTS[4] (see Section 2.4.2 and Chapter 22 for more information about mobility).

- **MME.** The MME is the control node which processes the signalling between the UE and the CN. The protocols running between the UE and the CN are known as the *Non-Access Stratum* (NAS) protocols.

  The main functions supported by the MME are classified as:

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

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

[2]Domain Name System.
[3]General Packet Radio Service.
[4]Universal Mobile Telecommunications System.

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

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

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

2.2.1.1 Non-Access Stratum (NAS) Procedures

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

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

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

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

When there is a need to deliver downlink data to an ECM-IDLE UE, the MME sends a paging message to all the eNodeBs in its current TA, and the eNodeBs page the UE over the radio interface. On receipt of a paging message, the UE performs a service request procedure which results in moving the UE to the ECM-CONNECTED state. UE-related information is thereby created in the E-UTRAN, and the bearers are re-established. The MME is responsible for the re-establishment of the radio bearers and updating the UE context in the eNodeB. This transition between the UE states is called an 'idle-to-active transition'. To speed up the idle-to-active transition and bearer establishment, EPS supports concatenation of the NAS and AS[5] procedures for bearer activation (see also Section 2.4.1). Some inter-relationship between the NAS and AS protocols is intentionally used to allow procedures to run simultaneously, rather than sequentially as in UMTS. For example, the bearer establishment procedure can be executed by the network without waiting for the completion of the security procedure.

[5]Access Stratum – the protocols which run between the eNodeBs and the UE.

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

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

2.2.2 The Access Network

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

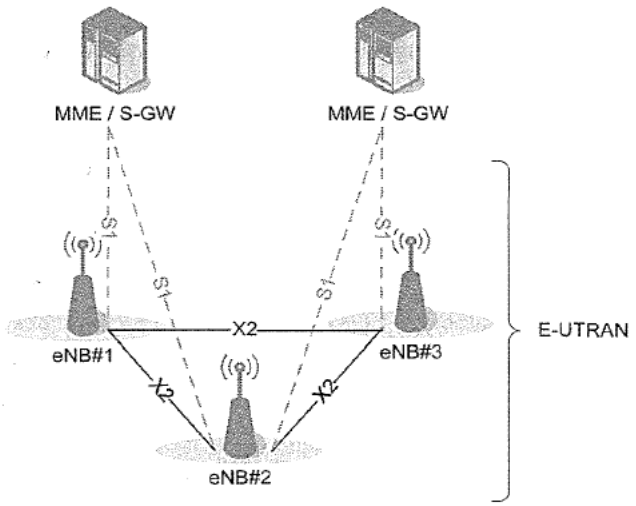


Figure 2.3: Overall E-UTRAN architecture. Reproduced by permission of © 3GPP.

The eNodeBs are normally inter-connected with each other by means of an interface known as *X2*, and to the EPC by means of the *S1* interface – more specifically, to the MME by means of the S1-MME interface and to the S-GW by means of the S1-U interface.

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

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

Samsung Ex. 1008
45 of 126

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

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

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

An important feature of the S1 interface linking the access network to the CN is known as *S1-flex*. This is a concept whereby multiple CN nodes (MME/S-GWs) can serve a common geographical area, being connected by a mesh network to the set of eNodeBs in that area (see Section 2.5). An eNodeB may thus be served by multiple MME/S-GWs, as is the case for eNodeB#2 in Figure 2.3. The set of MME/S-GW nodes serving a common area is called an *MME/S-GW pool* , and the area covered by such a pool of MME/S-GWs is called a *pool area*. This concept allows UEs in the cell(s) controlled by one eNodeB to be shared between multiple CN nodes, thereby providing a possibility for load sharing and also eliminating single points of failure for the CN nodes. The UE context normally remains with the same MME as long as the UE is located within the pool area.

2.2.3 Roaming Architecture

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

the home operator's services even while in a visited network. A P-GW in the visited network allows a 'local breakout' to the Internet in the visited network.

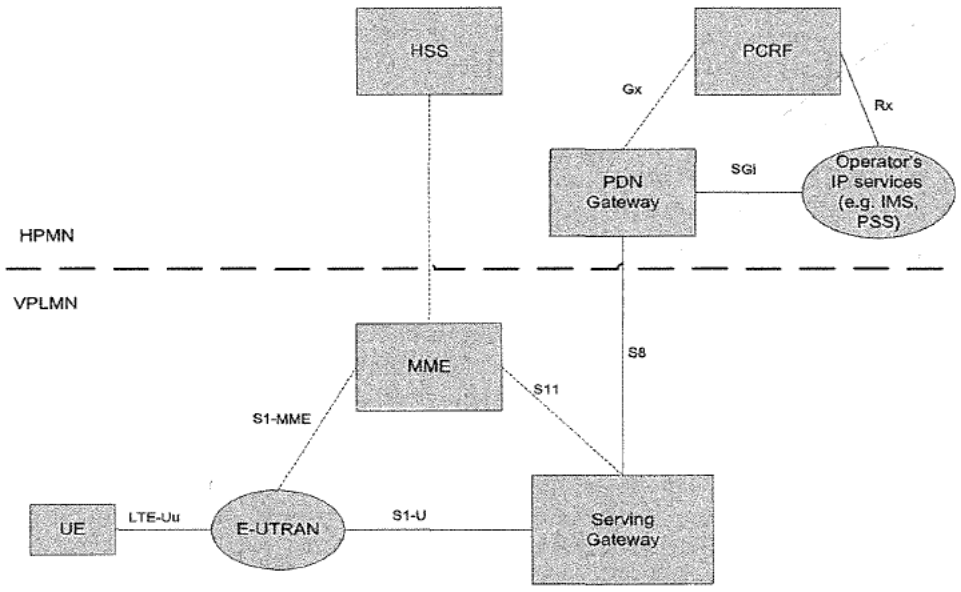


Figure 2.4: Roaming architecture for 3GPP accesses with P-GW in home network.

2.3 Protocol Architecture

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

2.3.1 User Plane

An IP packet for a UE is encapsulated in an EPC-specific protocol and tunnelled between the P-GW and the eNodeB for transmission to the UE. Different tunnelling protocols are used across different interfaces. A 3GPP-specific tunnelling protocol called the GPRS Tunnelling Protocol (GTP) [4] is used over the core network interfaces, S1 and S5/S8.[6]

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

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

Samsung Ex. 1008
47 of 126

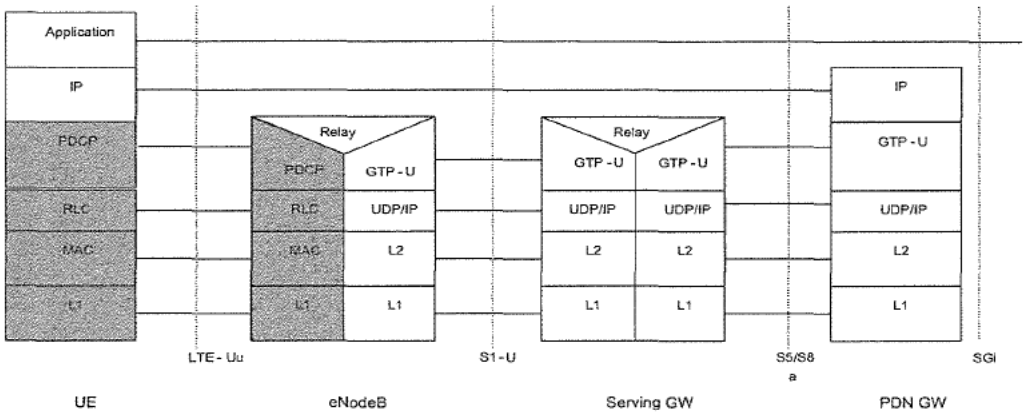


Figure 2.5: The E-UTRAN user plane protocol stack. Reproduced by permission of © 3GPP.

2.3.1.1 Data Handling During Handover

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

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

2.3.2 Control Plane

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

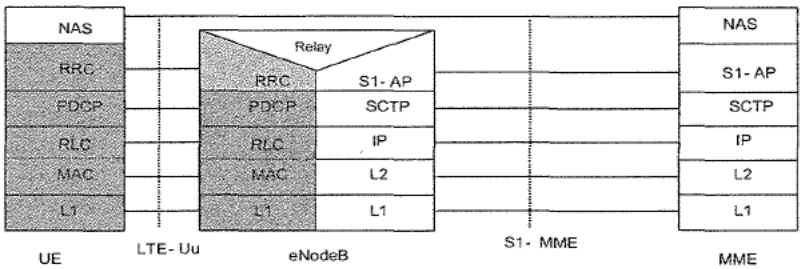


Figure 2.6: Control plane protocol stack. Reproduced by permission of © 3GPP.

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

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

2.4 Quality of Service and EPS Bearers

In a typical case, multiple applications may be running in a UE at the same time, each one having different QoS requirements. For example, a UE can be engaged in a VoIP call while at the same time browsing a web page or downloading an FTP file. VoIP has more stringent requirements for QoS in terms of delay and delay jitter than web browsing and FTP, while the latter requires a much lower packet loss rate. In order to support multiple QoS requirements, different bearers are set up within EPS, each being associated with a QoS.

Broadly, bearers can be classified into two categories based on the nature of the QoS they provide:

- **Minimum Guaranteed Bit Rate (GBR) bearers** which can be used for applications such as VoIP. These have an associated GBR value for which dedicated transmission resources are permanently allocated (e.g. by an admission control function in the eNodeB) at bearer establishment/modification. Bit rates higher than the GBR may be allowed for a GBR bearer if resources are available. In such cases, a Maximum Bit Rate (MBR) parameter, which can also be associated with a GBR bearer, sets an upper limit on the bit rate which can be expected from a GBR bearer.

- **Non-GBR bearers** which do not guarantee any particular bit rate. These can be used for applications such as web browsing or FTP transfer. For these bearers, no bandwidth resources are allocated permanently to the bearer.

In the access network, it is the eNodeB's responsibility to ensure that the necessary QoS for a bearer over the radio interface is met. Each bearer has an associated Class Identifier (QCI), and an Allocation and Retention Priority (ARP).

Each QCI is characterized by priority, packet delay budget and acceptable packet loss rate. The QCI label for a bearer determines the way it is handled in the eNodeB. Only a dozen such QCIs have been standardized so that vendors can all have the same understanding of the underlying service characteristics and thus provide the corresponding treatment, including queue management, conditioning and policing strategy. This ensures that an LTE operator can expect uniform traffic handling behaviour throughout the network regardless of the manufacturers of the eNodeB equipment. The set of standardized QCIs and their characteristics (from which the PCRF in an EPS can select) is provided in Table 2.1 (from Section 6.1.7 in [5]).

The priority and packet delay budget (and, to some extent, the acceptable packet loss rate) from the QCI label determine the RLC mode configuration (see Section 4.3.1), and how the scheduler in the MAC (see Section 4.4.2.1) handles packets sent over the bearer (e.g. in terms of scheduling policy, queue management policy and rate shaping policy). For example, a packet with a higher priority can be expected to be scheduled before a packet with lower priority. For bearers with a low acceptable loss rate, an Acknowledged Mode (AM) can be

Samsung Ex. 1008
49 of 126

Table 2.1: Standardized QoS Class Identifiers (QCIs) for LTE.

QCI	Resource type	Priority	Packet delay budget (ms)	Packet error loss rate	Example services
1	GBR	2	100	10^{-2}	Conversational voice
2	GBR	4	150	10^{-3}	Conversational video (live streaming)
3	GBR	5	300	10^{-6}	Non-conversational video (buffered streaming)
4	GBR	3	50	10^{-3}	Real time gaming
5	Non-GBR	1	100	10^{-6}	IMS signalling
6	Non-GBR	7	100	10^{-3}	Voice, video (live streaming), interactive gaming
7	Non-GBR	6	300	10^{-6}	Video (buffered streaming)
8	Non-GBR	8	300	10^{-6}	TCP-based (e.g. WWW, e-mail) chat, FTP, p2p file sharing, progressive video, etc.
9	Non-GBR	9	300	10^{-6}	

used within the RLC protocol layer to ensure that packets are delivered successfully across the radio interface (see Section 4.3.1.3).

The ARP of a bearer is used for call admission control – i.e. to decide whether or not the requested bearer should be established in case of radio congestion. It also governs the prioritization of the bearer for pre-emption with respect to a new bearer establishment request. Once successfully established, a bearer's ARP does not have any impact on the bearer-level packet forwarding treatment (e.g. for scheduling and rate control). Such packet forwarding treatment should be solely determined by the other bearer-level QoS parameters such as QCI, GBR and MBR.

An EPS bearer has to cross multiple interfaces as shown in Figure 2.7 – the S5/S8 interface from the P-GW to the S-GW, the S1 interface from the S-GW to the eNodeB, and the radio interface (also known as the LTE-Uu interface) from the eNodeB to the UE. Across each interface, the EPS bearer is mapped onto a lower layer bearer, each with its own bearer identity. Each node must keep track of the binding between the bearer IDs across its different interfaces.

An S5/S8 bearer transports the packets of an EPS bearer between a P-GW and an S-GW. The S-GW stores a one-to-one mapping between an S1 bearer and an S5/S8 bearer. The bearer is identified by the GTP tunnel ID across both interfaces.

An S1 bearer transports the packets of an EPS bearer between an S-GW and an eNodeB. A radio bearer [6] transports the packets of an EPS bearer between a UE and an eNodeB. An E-UTRAN Radio Access Bearer (E-RAB) refers to the concatenation of an S1 bearer and the corresponding radio bearer. An eNodeB stores a one-to-one mapping between a radio bearer ID and an S1 bearer to create the mapping between the two. The overall EPS bearer service architecture is shown in Figure 2.8.

Samsung Ex. 1008
50 of 126

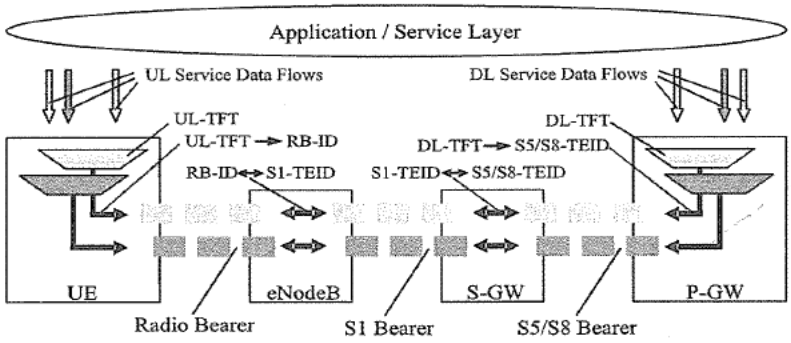


Figure 2.7: LTE/SAE bearers across the different interfaces. Reproduced by permission of © 3GPP.

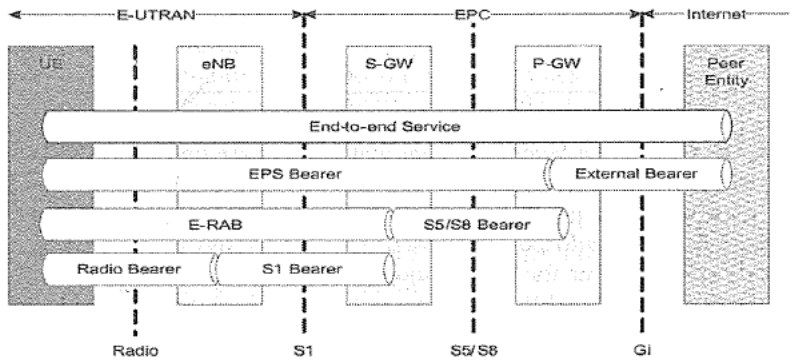


Figure 2.8: The overall EPS bearer service architecture. Reproduced by permission of © 3GPP.

IP packets mapped to the same EPS bearer receive the same bearer-level packet forwarding treatment (e.g. scheduling policy, queue management policy, rate shaping policy, RLC configuration). Providing different bearer-level QoS thus requires that a separate EPS bearer is established for each QoS flow, and user IP packets must be filtered into the different EPS bearers.

Packet filtering into different bearers is based on Traffic Flow Templates (TFTs). The TFTs use IP header information such as source and destination IP addresses and Transmission Control Protocol (TCP) port numbers to filter packets such as VoIP from web browsing traffic so that each can be sent down the respective bearers with appropriate QoS. An UpLink TFT (UL TFT) associated with each bearer in the UE filters IP packets to EPS bearers in the uplink direction. A DownLink TFT (DL TFT) in the P-GW is a similar set of downlink packet filters.

Samsung Ex. 1008
51 of 126

As part of the procedure by which a UE attaches to the network, the UE is assigned an IP address by the P-GW and at least one bearer is established, called the default bearer, and it remains established throughout the lifetime of the PDN connection in order to provide the UE with always-on IP connectivity to that PDN. The initial bearer-level QoS parameter values of the default bearer are assigned by the MME, based on subscription data retrieved from the HSS. The PCEF may change these values in interaction with the PCRF or according to local configuration. Additional bearers called dedicated bearers can also be established at any time during or after completion of the attach procedure. A dedicated bearer can be either GBR or non-GBR (the default bearer always has to be a non-GBR bearer since it is permanently established). The distinction between default and dedicated bearers should be transparent to the access network (e.g. E-UTRAN). Each bearer has an associated QoS, and if more than one bearer is established for a given UE, then each bearer must also be associated with appropriate TFTs. These dedicated bearers could be established by the network, based for example on a trigger from the IMS domain, or they could be requested by the UE. The dedicated bearers for a UE may be provided by one or more P-GWs.

The bearer-level QoS parameter values for dedicated bearers are received by the P-GW from the PCRF and forwarded to the S-GW. The MME only transparently forwards those values received from the S-GW over the S11 reference point to the E-UTRAN.

2.4.1 Bearer Establishment Procedure

This section describes an example of the end-to-end bearer establishment procedure across the network nodes using the functionality described in the previous sections.

A typical bearer establishment flow is shown in Figure 2.9. Each of the messages is described below.

When a bearer is established, the bearers across each of the interfaces discussed above are established.

The PCRF sends a 'PCC[7] Decision Provision' message indicating the required QoS for the bearer to the P-GW. The P-GW uses this QoS policy to assign the bearer-level QoS parameters. The P-GW then sends to the S-GW a 'Create Dedicated Bearer Request' message including the QoS and UL TFT to be used in the UE.

The S-GW forwards the Create Dedicated Bearer Request message (including bearer QoS, UL TFT and S1-bearer ID) to the MME (message 3 in Figure 2.9).

The MME then builds a set of session management configuration information including the UL TFT and the EPS bearer identity, and includes it in the 'Bearer Setup Request' message which it sends to the eNodeB (message 4 in Figure 2.9). The session management configuration is NAS information and is therefore sent transparently by the eNodeB to the UE.

The Bearer Setup Request also provides the QoS of the bearer to the eNodeB; this information is used by the eNodeB for call admission control and also to ensure the necessary QoS by appropriate scheduling of the user's IP packets. The eNodeB maps the EPS bearer QoS to the radio bearer QoS. It then signals a 'RRC Connection Reconfiguration' message (including the radio bearer QoS, session management configuration and EPS radio bearer identity) to the UE to set up the radio bearer (message 5 in Figure 2.9). The RRC Connection Reconfiguration message contains all the configuration parameters for the radio interface.

[7] Policy Control and Charging.

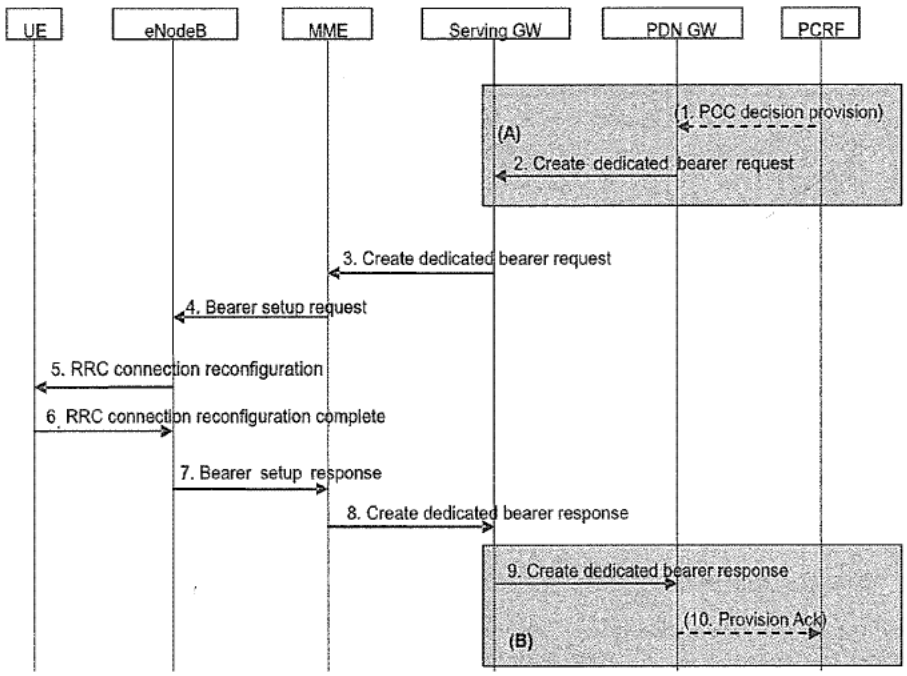


Figure 2.9: An example message flow for an LTE/SAE bearer establishment. Reproduced by permission of © 3GPP.

This is mainly for the configuration of the Layer 2 (PDCP, RLC and MAC) parameters, but also the Layer 1 parameters required for the UE to initialize the protocol stack.

Messages 6 to 10 are the corresponding response messages to confirm that the bearers have been set up correctly.

2.4.2 Inter-Working with other RATs

EPS also supports inter-working and mobility (handover) with networks using other Radio Access Technologies (RATs), notably GSM[8], UMTS, CDMA2000 and WiMAX. The architecture for inter-working with 2G and 3G GPRS/UMTS networks is shown in Figure 2.10. The S-GW acts as the mobility anchor for inter-working with other 3GPP technologies such as GSM and UMTS, while the P-GW serves as an anchor allowing seamless mobility to non-3GPP networks such as CDMA2000 or WiMAX. The P-GW may also support a Proxy Mobile Internet Protocol (PMIP) based interface. While VoIP is the primary mechanism for voice services, LTE also supports inter-working with legacy systems for CS voice services.

[8]Global System for Mobile Communications.

Samsung Ex. 1008
53 of 126

This is controlled by the MME and is based on two procedures outlined in Sections 2.4.2.1 and 2.4.2.2.

More details of the radio interface procedures for inter-working with other RATs are specified in [3] and covered in Sections 2.5.6.2 and 3.2.4.

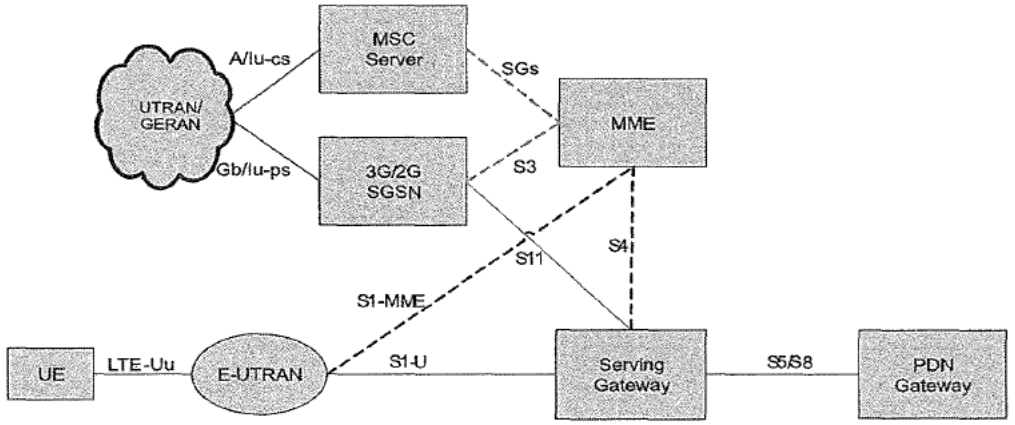


Figure 2.10: Architecture for 3G UMTS interworking.

2.4.2.1 Circuit-Switched Fall Back (CSFB)

LTE natively supports VoIP only using IMS services. However, in case IMS services are not deployed from the start, LTE also supports a Circuit-Switched FallBack (CSFB) mechanism which allows CS voice calls to be handled via legacy RATs for UEs that are camped on LTE.

CSFB allows a UE in LTE to be handed over to a legacy RAT to originate a CS voice call. This is supported by means of an interface, referred to as SGs[9], between the MME and the Mobile Switching Centre (MSC) of the legacy RAT shown in Figure 2.10. This interface allows the UE to attach with the MSC and register for CS services while still in LTE. Moreover it carries paging messages from the MSC for incoming voice calls so that UEs can be paged over LTE. The network may choose a handover, cell change order, or redirection procedure to move the UE to the legacy RAT.

Figure 2.11 shows the message flow for a CSFB call from LTE to UMTS, including paging from the MSC via the SGs interface and MME in the case of UE-terminated calls, and the sending of an Extended Service Request NAS message from the UE to the MME to trigger either a handover or redirection to the target RAT in the case of a UE-originated call. In the latter case, the UE then originates the CS call over the legacy RAT using the procedure defined in the legacy RAT specification. Further details of CSFB can be found in [7].

[9]SGs is an extension of the Gs interface between the Serving GPRS Support Node (SGSN) and the Mobile Switching Centre (MSC)

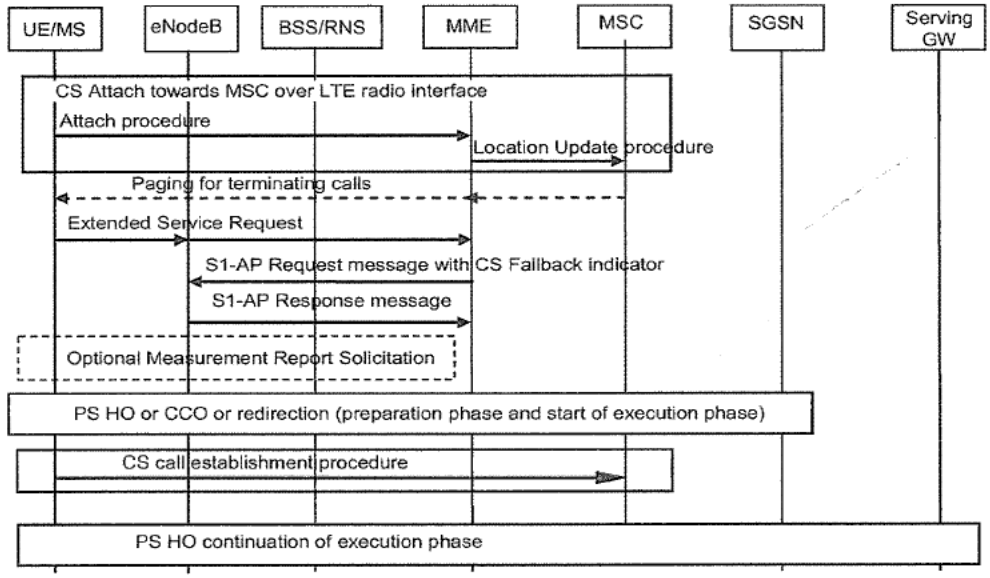


Figure 2.11: Message sequence diagram for CSFB from LTE to UMTS/GERAN.

2.4.2.2 Single Radio Voice Call Continuity (SRVCC)

If ubiquitous coverage of LTE is not available, it is possible that a UE involved in a VoIP call over LTE might then move out of LTE coverage to enter a legacy RAT cell which only offers CS voice services. The Single Radio Voice Call Continuity (SRVCC) procedure is designed for handover of a Packet Switched (PS) VoIP call over LTE to a CS voice call in the legacy RAT, involving the transfer of a PS bearer into a CS bearer.

Figure 2.12 shows an overview of the functions involved in SRVCC. The eNodeB may detect that the UE is moving out of LTE coverage and trigger a handover procedure towards the MME by means of an SRVCC indication. The MME is responsible for the SRVCC procedure and also for the transfer of the PS E-RAB carrying VoIP into a CS bearer. The MSC Server then initiates the session transfer procedure to IMS and coordinates it with the CS handover procedure to the target cell. The handover command provided to the UE to request handover to the legacy RAT also provides the information to set up the CS and PS radio bearers. The UE can continue with the call over the CS domain on completion of the handover. Further details of SRVCC can be found in [8].

2.5 The E-UTRAN Network Interfaces: S1 Interface

The S1 interface connects the eNodeB to the EPC. It is split into two interfaces, one for the control plane and the other for the user plane. The protocol structure for the S1 and the functionality provided over S1 are discussed in more detail below.

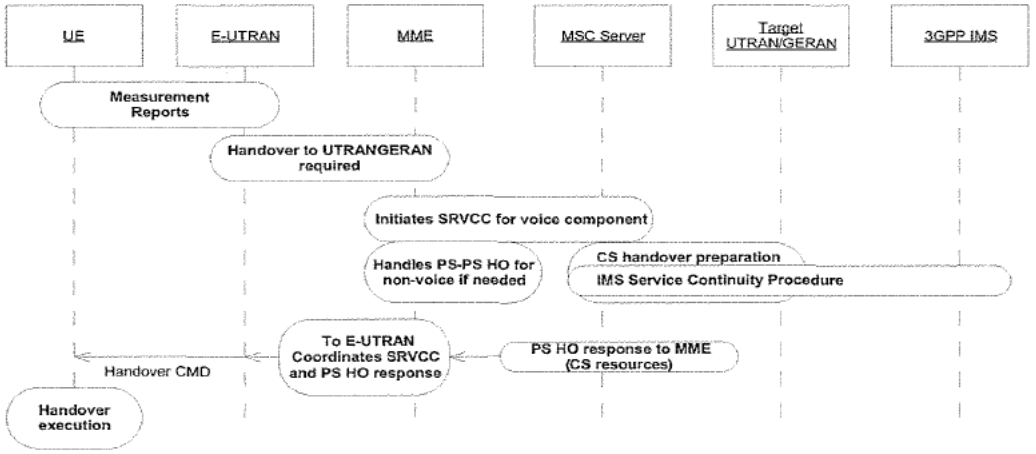


Figure 2.12: The main procedures involved in an SRVCC handover of a PS VoIP call from LTE to CS voice call in UMTS/GERAN.

2.5.1 Protocol Structure over S1

The protocol structure over S1 is based on a full IP transport stack with no dependency on legacy SS7[10] network configuration as used in GSM or UMTS networks. This simplification provides one area of potential savings on operational expenditure with LTE networks.

2.5.1.1 Control Plane

Figure 2.13 shows the protocol structure of the S1 control plane which is based on the Stream Control Transmission Protocol / IP (SCTP/IP) stack.

The SCTP protocol is well known for its advanced features inherited from TCP which ensure the required reliable delivery of the signalling messages. In addition, it makes it possible to benefit from improved features such as the handling of multistreams to implement transport network redundancy easily and avoid head-of-line blocking or multihoming (see 'IETF RFC4960' [9]).

A further simplification in LTE (compared to the UMTS Iu interface, for example) is the direct mapping of the S1-AP (S1 Application Protocol) on top of SCTP which results in a simplified protocol stack with no intermediate connection management protocol. The individual connections are directly handled at the application layer. Multiplexing takes place between S1-AP and SCTP whereby each stream of an SCTP association is multiplexed with the signalling traffic of multiple individual connections.

One further area of flexibility that comes with LTE lies in the lower layer protocols for which full optionality has been left regarding the choice of the IP version and the choice

[10]Signalling System #7 (SS7) is a communications protocol defined by the International Telecommunication Union (ITU) Telecommunication Standardization Sector (ITU-T) with a main purpose of setting up and tearing down telephone calls. Other uses include Short Message Service (SMS), number translation, prepaid billing mechanisms, and many other services.

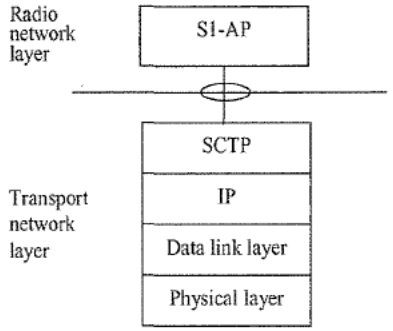


Figure 2.13: S1-MME control plane protocol stack. Reproduced by permission of © 3GPP.

of the data link layer. For example, this enables the operator to start deployment using IP version 4 with the data link tailored to the network deployment scenario.

2.5.1.2 User Plane

Figure 2.14 shows the protocol structure of the S1 user plane, which is based on the GTP/ User Datagram Protocol (UDP) IP stack which is already well known from UMTS networks.

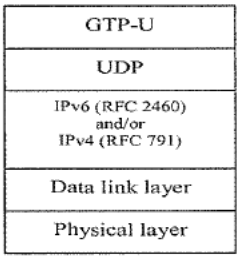


Figure 2.14: S1-U user plane protocol stack. Reproduced by permission of © 3GPP.

One of the advantages of using GTP-User plane (GTP-U) is its inherent facility to identify tunnels and also to facilitate intra-3GPP mobility.

The IP version number and the data link layer have been left fully optional, as for the control plane stack.

A transport bearer is identified by the GTP tunnel endpoints and the IP address (source Tunnelling End ID (TEID), destination TEID, source IP address, destination IP address).

The S-GW sends downlink packets of a given bearer to the eNodeB IP address (received in S1-AP) associated to that particular bearer. Similarly, the eNodeB sends upstream packets of a given bearer to the EPC IP address (received in S1-AP) associated to that particular bearer.

Vendor-specific traffic categories (e.g. real-time traffic) can be mapped onto Differentiated Services (Diffserv) code points (e.g. expedited forwarding) by network O&M (Operation and Maintenance) configuration to manage QoS differentiation between the bearers.

2.5.2 Initiation over S1

The initialization of the S1-MME control plane interface starts with the identification of the MMEs to which the eNodeB must connect, followed by the setting up of the Transport Network Layer (TNL).

With the support of the S1-flex function in LTE, an eNodeB must initiate an S1 interface towards each MME node of the pool area to which it belongs. This list of MME nodes of the pool together with an initial corresponding remote IP address can be directly configured in the eNodeB at deployment (although other means may also be used). The eNodeB then initiates the TNL establishment with that IP address. Only one SCTP association is established between one eNodeB and one MME.

During the establishment of the SCTP association, the two nodes negotiate the maximum number of streams which will be used over that association. However, multiple pairs of streams[11] are typically used in order to avoid the head-of-line blocking issue mentioned above. Among these pairs of streams, one particular pair must be reserved by the two nodes for the signalling of the common procedures (i.e. those which are not specific to one UE). The other streams are used for the sole purpose of the dedicated procedures (i.e. those which are specific to one UE).

Once the TNL has been established, some basic application-level configuration data for the system operation is automatically exchanged between the eNodeB and the MME through an 'S1 SETUP' procedure initiated by the eNodeB. This procedure is one case of a Self-Optimizing Network process and is explained in detail in Section 25.3.1.

Once the S1 SETUP procedure has been completed, the S1 interface is operational.

2.5.3 Context Management over S1

Within each pool area, a UE is associated to one particular MME for all its communications during its stay in this pool area. This creates a context in this MME for the UE. This particular MME is selected by the NAS Node Selection Function (NNSF) in the first eNodeB from which the UE entered the pool.

Whenever the UE becomes active (i.e. makes a transition from idle to active mode) under the coverage of a particular eNodeB in the pool area, the MME provides the UE context information to this eNodeB using the 'INITIAL CONTEXT SETUP REQUEST' message (see Figure 2.15). This enables the eNodeB in turn to create a context and manage the UE while it is in active mode.

Even though the setup of bearers is otherwise relevant to a dedicated 'Bearer Management' procedure described below, the creation of the eNodeB context by the INITIAL CONTEXT SETUP procedure also includes the creation of one or several bearers including the default bearers.

At the next transition back to idle mode following a 'UE CONTEXT RELEASE' message sent from the MME, the eNodeB context is erased and only the MME context remains.

[11] Note that a stream is unidirectional and therefore pairs must be used.

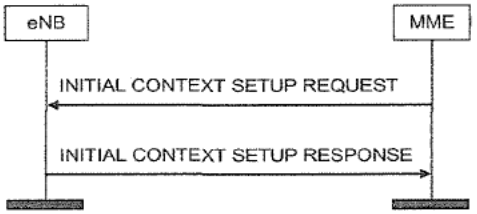


Figure 2.15: Initial context setup procedure. Reproduced by permission of © 3GPP.

2.5.4 Bearer Management over S1

LTE uses independent dedicated procedures respectively covering the setup, modification and release of bearers. For each bearer requested to be set up, the transport layer address and the tunnel endpoint are provided to the eNodeB in the 'BEARER SETUP REQUEST' message to indicate the termination of the bearer in the S-GW where uplink user plane data must be sent. Conversely, the eNodeB indicates in the 'BEARER SETUP RESPONSE' message the termination of the bearer in the eNodeB where the downlink user plane data must be sent.

For each bearer, the QoS parameters (see Section 2.4) requested for the bearer are also indicated. Independently of the standardized QCI values, it is also still possible to use extra proprietary labels for the fast introduction of new services if vendors and operators agree upon them.

2.5.5 Paging over S1

As mentioned in Section 2.5.3, in order to re-establish a connection towards a UE in idle mode, the MME distributes a 'PAGING REQUEST' message to the relevant eNodeBs based on the TAs where the UE is expected to be located. When receiving the paging request, the eNodeB sends a page over the radio interface in the cells which are contained within one of the TAs provided in that message.

The UE is normally paged using its S-TMSI. The 'PAGING REQUEST' message also contains a UE identity index value in order for the eNodeB to calculate the paging occasions at which the UE will switch on its receiver to listen for paging messages (see Section 3.4).

In Release 10, paging differentiation is introduced over the S1 interface to handle Multimedia Priority Service (MPS)[12] users. In case of MME or RAN overload, it is necessary to page a UE with higher priority during the establishment of a mobile-terminated MPS call. In case of MME overload, the MME can itself discriminate between the paging messages and discard the lower priority ones. In case of RAN overload in some cells, the eNodeB can perform this discrimination based on a new Paging Priority Indicator sent by the MME. The MME can signal up to eight such priority values to the eNodeB. In case of an IMS MPS call, the terminating UE will further set up an RRC connection with the same eNodeB that will also get automatically prioritized. In case of a CS fallback call, the eNodeB will instead

[12]MPS allows the delivery of calls or complete sessions of a high priority nature, in case for example of public safety or national security purposes, from mobile to mobile, mobile to fixed, and fixed to mobile networks during network congestion conditions.

Samsung Ex. 1008
59 of 126

signal to the UE that it must set the cause value 'high priority terminating call' when trying to establish the UMTS RRC Connection.

2.5.6 Mobility over S1

LTE/SAE supports mobility within LTE/SAE, and also to other systems using both 3GPP and non-3GPP technologies. The mobility procedures over the radio interface are defined in Section 3.2. These mobility procedures also involve the network interfaces. The sections below discuss the procedures over S1 to support mobility. The mobility performance requirements from the UE point of view are outlined in Chapter 22.

2.5.6.1 Intra-LTE Mobility

There are two types of handover procedure in LTE for UEs in active mode: the S1-handover procedure and the X2-handover procedure.

For intra-LTE mobility, the X2-handover procedure is normally used for the inter-eNodeB handover (described in Section 2.6.3). However, when there is no X2 interface between the two eNodeBs, or if the source eNodeB has been configured to initiate handover towards a particular target eNodeB via the S1 interface, then an S1-handover will be triggered.

The S1-handover procedure has been designed in a very similar way to the UMTS Serving Radio Network Subsystem (SRNS) relocation procedure and is shown in Figure 2.16: it consists of a preparation phase involving the core network, where the resources are first prepared at the target side (steps 2 to 8), followed by an execution phase (steps 8 to 12) and a completion phase (after step 13).

Compared to UMTS, the main difference is the introduction of the 'STATUS TRANSFER' message sent by the source eNodeB (steps 10 and 11). This message has been added in order to carry some PDCP status information that is needed at the target eNodeB in cases when PDCP status preservation applies for the S1-handover (see Section 4.2.4); this is in alignment with the information which is sent within the X2 'STATUS TRANSFER' message used for the X2-handover (see below). As a result of this alignment, the handling of the handover by the target eNodeB as seen from the UE is exactly the same, regardless of the type of handover (S1 or X2) the network had decided to use; indeed, the UE is unaware of which type of handover is used by the network.

The 'Status Transfer' procedure is assumed to be triggered in parallel with the start of data forwarding after the source eNodeB has received the 'HANDOVER COMMAND' message from the source MME. This data forwarding can be either direct or indirect, depending on the availability of a direct path for the user plane data between the source eNodeB and the target eNodeB.

The 'HANDOVER NOTIFY' message (step 13), which is sent later by the target eNodeB when the arrival of the UE at the target side is confirmed, is forwarded by the MME to trigger the update of the path switch in the S-GW towards the target eNodeB. In contrast to the X2-handover, the message is not acknowledged and the resources at the source side are released later upon reception of a 'RELEASE RESOURCE' message directly triggered from the source MME (step 17 in Figure 2.16).

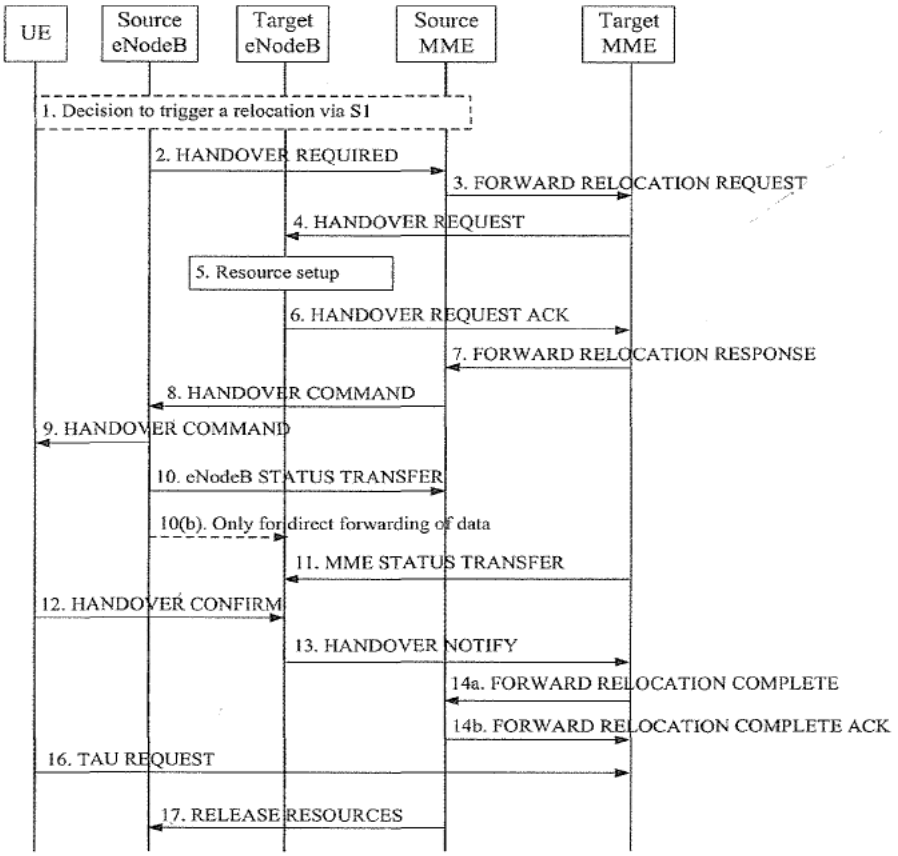


Figure 2.16: S1-based handover procedure. Reproduced by permission of © 3GPP.

2.5.6.2 Inter-RAT Mobility

One key element of the design of LTE is the need to co-exist with other Radio Access Technologies (RATs).

For mobility from LTE towards UMTS, the handover process can reuse the S1-handover procedures described above, with the exception of the 'STATUS TRANSFERŠ message which is not needed at steps 10 and 11 since no PDCP context is continued.

For mobility towards CDMA2000, dedicated uplink and downlink procedures have been introduced in LTE. They essentially aim at tunnelling the CDMA2000 signalling between the UE and the CDMA2000 system over the S1 interface, without being interpreted by the eNodeB on the way. An 'UPLINK S1 CDMA2000 TUNNELLING' message is sent from the eNodeB to the MME; this also includes the RAT type in order to identify which CDMA2000

Samsung Ex. 1008
61 of 126

RAT the tunnelled CDMA2000 message is associated with in order for the message to be routed to the correct node within the CDMA2000 system.

2.5.6.3 Mobility towards Home eNodeBs

Mobility towards HeNBs involves additional functions from the source LTE RAN node and the MME. In addition to the E-UTRAN Cell Global Identifier (ECGI), the source RAN node should include the Closed Subscriber Group Identity (CSG ID) and the access mode of the target HeNB in the 'HANDOVER REQUIRED' message to the MME so that the MME can perform the access control to that HeNB. If the target HeNB operates in closed access mode (see Chapter 24) and the MME fails the access control, the MME will reject the handover by sending back a 'HANDOVER PREPARATION FAILURE' message. Otherwise the MME will accept and continue the handover while indicating to the target HeNB whether the UE is a 'CSG member' if the HeNB is operating in hybrid mode. A detailed description of mobility towards the HeNB and the associated call flow is provided in Chapter 24.

2.5.7 Load Management over S1

Three types of load management procedures apply over S1: a normal 'load balancing' procedure to distribute the traffic, an 'overload' procedure to overcome a sudden peak in the loading and a 'load rebalancing' procedure to partially/fully offload an MME.

The MME load balancing procedure aims to distribute the traffic to the MMEs in the pool evenly according to their respective capacities. To achieve that goal, the procedure relies on the normal NNSF present in each eNodeB as part of the S1-flex function. Provided that suitable weight factors corresponding to the capacity of each MME node are available in the eNodeBs beforehand, a weighted NNSF done by every eNodeB in the network normally achieves a statistically balanced distribution of load among the MME nodes without further action. However, specific actions are still required for some particular scenarios:

- If a new MME node is introduced (or removed), it may be necessary temporarily to increase (or decrease) the weight factor normally corresponding to the capacity of this node in order to make it catch more (or less) traffic at the beginning until it reaches an adequate level of load.

- In case of an unexpected peak in the loading, an 'OVERLOAD' message can be sent over the S1 interface by the overloaded MME. When received by an eNodeB, this message calls for a temporary restriction of a certain type of traffic. An MME can adjust the reduction of traffic it desires by defining the number of eNodeBs to which it sends the 'OVERLOAD' message and by defining the types of traffic subject to restriction.Two new rejection types are introduced in Release 10 to combat CN Overload:

  - 'reject low priority access', which can be used by the MME to reduce access of some low-priority devices or applications such as Machine-Type Communication (MTC) devices (see Section 31.4);
  - 'permit high priority sessions', to allow access only to high-priority users and mobile-terminated services.

- Finally, if the MME wants to force rapidly the offload of part or all of its UEs, it will use the rebalancing function. This function forces the UEs to reattach to another MME by using a specific 'cause value' in the 'UE Release Command S1' message. In a first step it applies to idle mode UEs and in a second step it may also apply to UEs in connected mode (if the full MME offload is desired, e.g. for maintenance reasons).

2.5.8 Trace Function

In order to trace the activity of a UE in connected mode, two types of trace session can be started in the eNodeB:

- Signalling-Based Trace. This is triggered by the MME and is uniquely identified by a trace identity. Only one trace session can be activated at a time for one UE. The MME indicates to the eNodeB the interfaces to trace (e.g. S1, X2, Uu) and the associated trace depth. The trace depth represents the granularity of the signalling to be traced from the high-level messages down to the detailed ASN.1[13] and is comprised of three levels: minimum, medium and maximum. The MME also indicates the IP address of a Trace Collection Entity where the eNodeB must send the resulting trace record file. If an X2 handover preparation has started at the time when the eNodeB receives the order to trace, the eNodeB will signal back a TRACE FAILURE INDICATION message to the MME, and it is then up to the MME to take appropriate action based on the indicated failure reason. Signalling-based traces are propagated at X2 and S1 handover.

- Management-Based Trace. This is triggered in the eNodeB when the conditions required for tracing set by O&M are met. The eNodeB then allocates a trace identity that it sends to the MME in a CELL TRAFFIC TRACE message over S1, together with the Trace Collection Entity identity that shall be used by the MME for the trace record file (in order to assemble the trace correctly in the Trace Collection Entity). Management-based traces are propagated at X2 and S1 handover.

In Release 10, the trace function supports the Minimization of Drive Tests (MDT) feature, which is explained in Section 31.3.

2.5.9 Delivery of Warning Messages

Two types of warning message may need to be delivered with the utmost urgency over a cellular system, namely Earthquake and Tsunami Warning System (ETWS)) messages and Commercial Mobile Alert System (CMAS) messages (see Section 13.7). The delivery of ETWS messages is already supported since Release 8 via the S1 Write-Replace Warning procedure which makes it possible to carry either primary or secondary notifications over S1 for the eNodeB to broadcast over the radio. The Write-Replace Warning procedure also includes a Warning Area List where the warning message needs to be broadcast. It can be a list of cells, tracking areas or emergency area identities. The procedure also contains information on how the broadcast is to be performed (for example, the number of broadcasts requested).

[13] Abstract Syntax Notation One

Samsung Ex. 1008
63 of 126

In contrast to ETWS, the delivery of CMAS messages is only supported from Release 9 onwards. One difference between the two public warning systems is that in ETWS the eNodeB can only broadcast one message at a time, whereas CMAS allows the broadcast of multiple concurrent warning messages over the radio. Therefore an ongoing ETWS broadcast needs to be overwritten if a new ETWS warning has to be delivered immediately in the same cell. With CMAS, a new Kill procedure has also been added to allow easy cancellation of an ongoing broadcast when needed. This Kill procedure includes the identity of the message to be stopped and the Warning Area where it is to be stopped.

2.6 The E-UTRAN Network Interfaces: X2 Interface

The X2 interface is used to inter-connect eNodeBs. The protocol structure for the X2 interface and the functionality provided over X2 are discussed below.

2.6.1 Protocol Structure over X2

The control plane and user plane protocol stacks over the X2 interface are the same as over the S1 interface, as shown in Figures 2.17 and 2.18 respectively (with the exception that in Figure 2.17 the X2-AP (X2 Application Protocol) is substituted for the S1-AP). This also means again that the choice of the IP version and the data link layer are fully optional. The use of the same protocol structure over both interfaces provides advantages such as simplifying the data forwarding operation.

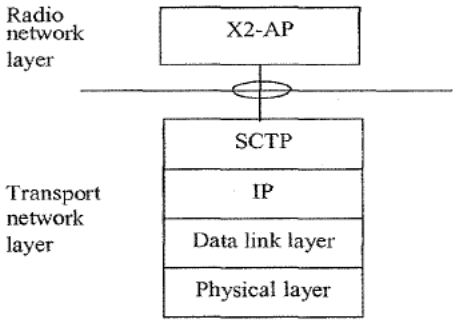


Figure 2.17: X2 signalling bearer protocol stack. Reproduced by permission of © 3GPP.

2.6.2 Initiation over X2

The X2 interface may be established between one eNodeB and some of its neighbour eNodeBs in order to exchange signalling information when needed. However, a full mesh is not mandated in an E-UTRAN network. Two types of information may typically need to be exchanged over X2 to drive the establishment of an X2 interface between two eNodeBs:

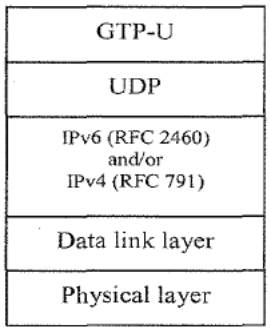


Figure 2.18: Transport network layer for data streams over X2. Reproduced by permission of © 3GPP.

load or interference related information (see Section 2.6.4) and handover related information (see mobility in Section 2.6.3).

Because these two types of information are fully independent of one another, it is possible that an X2 interface may be present between two eNodeBs for the purpose of exchanging load or interference information, even though the X2-handover procedure is not used to hand over UEs between those eNodeBs.[14]

The initialization of the X2 interface starts with the identification of a suitable neighbour followed by the setting up of the TNL.

The identification of a suitable neighbour may be done by configuration, or alternatively by a self-optimizing process known as the Automatic Neighbour Relation Function (ANRF).[15] This is described in more detail in Section 25.2.

Once a suitable neighbour has been identified, the initiating eNodeB can further set up the TNL using the transport layer address of this neighbour – either as retrieved from the network or locally configured. The automatic retrieval of the X2 IP address(es) via the network and the eNodeB Configuration Transfer procedure are described in details in Section 25.3.2.

Once the TNL has been set up, the initiating eNodeB must trigger the X2 setup procedure. This procedure enables an automatic exchange of application level configuration data relevant to the X2 interface, similar to the S1 setup procedure already described in Section 2.5.2. For example, each eNodeB reports within the 'X2 SETUP REQUESTŠ message to a neighbour eNodeB information about each cell it manages, such as the cell's physical identity, the frequency band, the tracking area identity and/or the associated PLMNs.

This automatic exchange of application-level configuration data within the X2 setup procedure is also the core of two additional SON features: automatic self-configuration of the Physical Cell Identities (PCIs) and RACH self-optimization. These features both aim to avoid conflicts between cells controlled by neighbouring eNodeBs; they are explained in detail in Sections 25.4 and 25.7 respectively.

Once the X2 setup procedure has been completed, the X2 interface is operational.

[14] In such a case, the S1-handover procedure is used instead.

[15] Under this function the UEs are requested to detect neighbour eNodeBs by reading the Cell Global Identity (CGI) contained in the broadcast information.

Samsung Ex. 1008
65 of 126

2.6.3 Mobility over X2

Handover via the X2 interface is triggered by default unless there is no X2 interface established or the source eNodeB is configured to use the S1-handover instead.

The X2-handover procedure is illustrated in Figure 2.19. Like the S1-handover, it is also composed of a preparation phase (steps 4 to 6), an execution phase (steps 7 to 9) and a completion phase (after step 9).

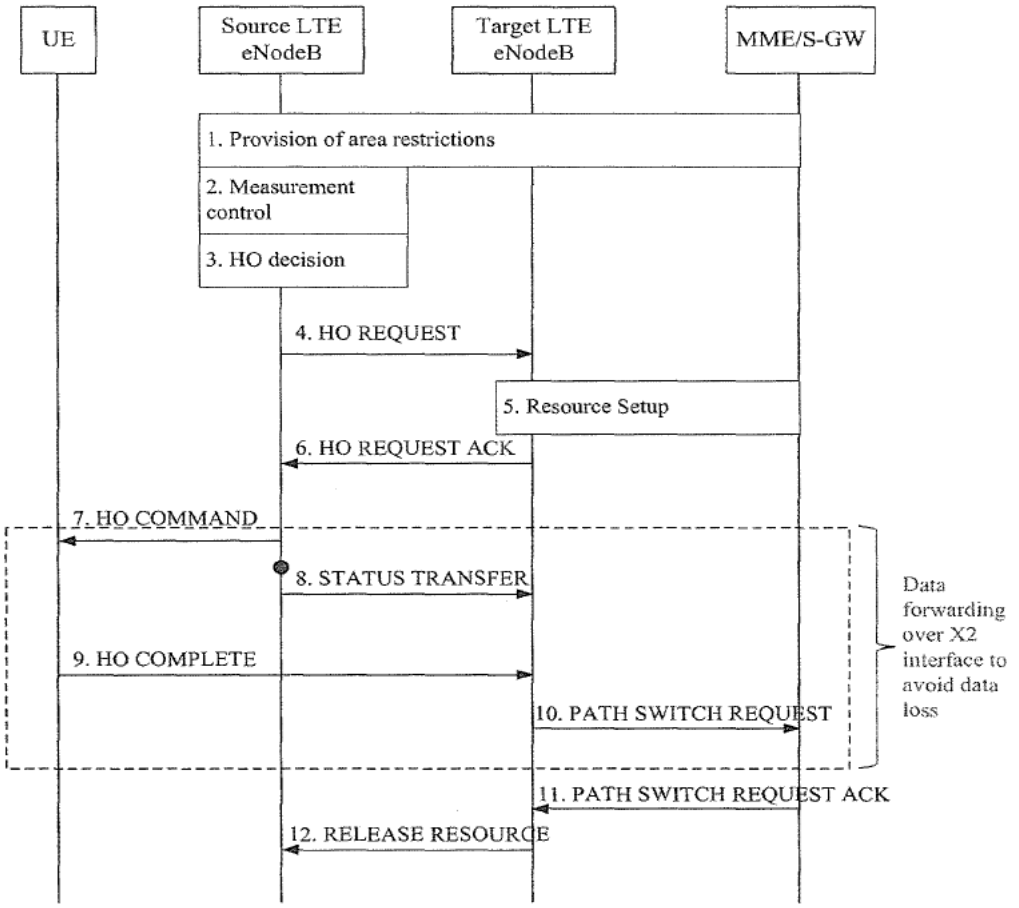


Figure 2.19: X2-based handover procedure.

The key features of the X2-handover for intra-LTE handover are:

- The handover is directly performed between two eNodeBs. This makes the preparation phase quick.

- Data forwarding may be operated per bearer in order to minimize data loss.
- The MME is only informed at the end of the handover procedure once the handover is successful, in order to trigger the path switch.
- The release of resources at the source side is directly triggered from the target eNodeB.

For those bearers for which in-sequence delivery of packets is required, the 'STATUS TRANSFER' message (step 8) provides the Sequence Number (SN) and the Hyper Frame Number (HFN) which the target eNodeB should assign to the first packet with no sequence number yet assigned that it must deliver. This first packet can either be one received over the target S1 path or one received over X2 if data forwarding over X2 is used (see below). When it sends the 'STATUS TRANSFER' message, the source eNodeB freezes its transmitter/receiver status – i.e. it stops assigning PDCP SNs to downlink packets and stops delivering uplink packets to the EPC.

Mobility over X2 can be categorized according to its resilience to packet loss: the handover can be said 'seamless' if it minimizes the interruption time during the move of the UE, or 'lossless' if it tolerates no loss of packets at all. These two modes use data forwarding of user plane downlink packets. The source eNodeB may decide to operate one of these two modes on a per-EPS-bearer basis, based on the QoS received over S1 for this bearer (see Section 2.5.4) and the service in question. These two modes are described in more detail below.

2.6.3.1 Seamless Handover

If, for a given bearer, the source eNodeB selects the seamless handover mode, it proposes to the target eNodeB in the 'HANDOVER REQUEST' message to establish a GTP tunnel to operate the downlink data forwarding. If the target eNodeB accepts, it indicates in the 'HANDOVER REQUEST ACK' message the tunnel endpoint where the forwarded data is expected to be received. This tunnel endpoint may be different from the one set up as the termination point of the new bearer established over the target S1.

Upon reception of the 'HANDOVER REQUEST ACK' message, the source eNodeB can start forwarding the data freshly arriving over the source S1 path towards the indicated tunnel endpoint in parallel with sending the handover trigger to the UE over the radio interface. This forwarded data is thus available at the target eNodeB to be delivered to the UE as early as possible.

When forwarding is in operation and in-sequence delivery of packets is required, the target eNodeB is assumed to deliver first the packets forwarded over X2 before delivering the first ones received over the target S1 path once the S1 path switch has been performed. The end of the forwarding is signalled over X2 to the target eNodeB by the reception of some 'special GTP packets' which the S-GW has inserted over the source S1 path just before switching this S1 path; these are then forwarded by the source eNodeB over X2 like any other regular packets.

2.6.3.2 Lossless Handover

If the source eNodeB selects the lossless mode for a given bearer, it will additionally forward over X2 those user plane downlink packets which it has PDCP processed but are still buffered locally because they have not yet been delivered and acknowledged by the UE. These packets

Samsung Ex. 1008
67 of 126

are forwarded together with their assigned PDCP SN included in a GTP extension header field. They are sent over X2 prior to the freshly arriving packets from the source S1 path. The same mechanisms described above for the seamless handover are used for the GTP tunnel establishment. The end of forwarding is also handled in the same way, since in-sequence packet delivery applies to lossless handovers. In addition, the target eNodeB must ensure that all the packets – including the ones received with sequence number over X2 – are delivered in sequence at the target side. Further details of seamless and lossless handover are described in Section 4.2.

Selective retransmission. A new feature in LTE compared to previous systems is the optimization of the radio interface usage by selective retransmission. When lossless handover is operated, the target eNodeB may, however, not deliver over the radio interface some of the forwarded downlink packets received over X2 if it is informed by the UE that those packets have already been received at the source side (see Section 4.2.6). This is called downlink selective retransmission.

Similarly in the uplink, the target eNodeB may desire that the UE does not retransmit packets already received earlier at the source side by the source eNodeB, for example to avoid wasting radio resources. To operate this uplink selective retransmission scheme for one bearer, it is necessary that the source eNodeB forwards to the target eNodeB, over another new GTP tunnel, those user plane uplink packets which it has received out of sequence. The target eNodeB must first request the source eNodeB to establish this new forwarding tunnel by including in the 'HANDOVER REQUEST ACK' message a GTP tunnel endpoint where it expects the forwarded uplink packets to be received. The source eNodeB must, if possible, then indicate in the 'STATUS TRANSFER' message for this bearer the list of SNs corresponding to the forwarded packets which are to be expected. This list helps the target eNodeB to inform the UE earlier of the packets not to be retransmitted, making the overall uplink selective retransmission scheme faster (see also Section 4.2.6).

2.6.3.3 Multiple Preparation

'Multiple preparation' is another new feature of the LTE handover procedure. This feature enables the source eNodeB to trigger the handover preparation procedure towards multiple candidate target eNodeBs. Even though only one of the candidates is indicated as target to the UE, this makes recovery faster in case the UE fails on this target and connects to one of the other prepared candidate eNodeBs. The source eNodeB receives only one 'RELEASE RESOURCE' message from the final selected eNodeB.

Regardless of whether multiple or single preparation is used, the handover can be cancelled during or after the preparation phase. If the multiple preparation feature is operated, it is recommended that upon reception of the 'RELEASE RESOURCE' message the source eNodeB triggers a 'cancel' procedure towards each of the non-selected prepared eNodeBs.

2.6.3.4 Mobility Robustness Handling

In order to detect and report the cases where the mobility is unsuccessful and results in connection failures, specific messages are available over the X2 interface from Release 9 onwards to report handovers that are triggered too late or too early or to an inappropriate cell. These scenarios are explained in detail in Section 25.6.

Samsung Ex. 1008
68 of 126

2.6.3.5 Mobility towards Home eNodeBs via X2

In Release 10, in order to save backhaul bandwidth reduce delays, mobility between two HeNBs does not necessarily need to use S1 handover and transit via the MME but can directly use the X2 handover. This optimization is described in detail in Section 24.2.3.

2.6.4 Load and Interference Management Over X2

The exchange of load information between eNodeBs is of key importance in the flat architecture used in LTE, as there is no central Radio Resource Management (RRM) node as was the case, for example, in UMTS with the Radio Network Controller (RNC).

The exchange of load information falls into two categories depending on the purpose it serves:

- **Load balancing.** If the exchange of load information is for the purpose of load balancing, the frequency of exchange is rather low (in the order of seconds). The objective of load balancing is to counteract local traffic load imbalance between neighbouring cells with the aim of improving the overall system capacity. The mechanisms for this are explained in detail in Section 25.5.

  In Release 10, partial reporting is allowed per cell and per measurement. Therefore, if a serving eNodeB does not support some measurements, it will still report the other measurements that it does support. For each unsupported measurement, the serving eNodeB can indicate if the lack of support is permanent or temporary.

- **Interference coordination.** If the exchange of load information is to optimize RRM processes such as interference coordination, the frequency of exchange is rather high (in the order of tens of milliseconds). A special X2 'LOAD INDICATION' message is provided over the X2 interface for the exchange of load information related to interference management. For uplink interference management, two indicators can be provided within the 'LOAD INDICATION' message: a 'High Interference Indicator' and an 'Overload Indicator'. The usage of these indicators is explained in detail in Section 12.5.

The Load Indication procedure allows an eNodeB to signal to its neighbour eNodeBs new interference coordination intentions when applicable. This can either be frequency-domain interference management, as explained in Sections 12.5.1 and 12.5.2, or time-domain interference management, as explained in Section 31.2.3.

2.6.5 UE Historical Information Over X2

The provision of UE historical information is part of the X2-handover procedure and is designed to support self-optimization of the network.

Generally, the UE historical information consists of some RRM information which is passed from the source eNodeB to the target eNodeB within the 'HANDOVER REQUEST' message to assist the RRM management of a UE. The information can be partitioned into two types:

- UE RRM-related information, passed over X2 within the RRC transparent container;

- Cell RRM-related information, passed over X2 directly as an information element of the 'X2 AP HANDOVER REQUEST' message itself.

An example of such UE historical information is the list of the last few cells visited by the UE, together with the time spent in each one. This information is propagated from one eNodeB to another and can be used to determine the occurrence of ping-pong between two or three cells for instance. The length of the history information can be configured for more flexibility.

2.7 Summary

The EPS provides UEs with IP connectivity to the packet data network. In this chapter we have seen an overview of the EPS network architecture, including the functionalities provided by the E-UTRAN access network and the evolved packet core network..

It can be seen that the concept of EPS bearers, together with their associated quality of service attributes, provide a powerful tool for the provision of a variety of simultaneous services to the end user. Depending on the nature of the application, the EPS can supply the UE with multiple data flows with different QoSs. A UE can thus be engaged in a VoIP call which requires guaranteed delay and bit rate at the same time as browsing the web with a best effort QoS.

From the perspective of the network operator, the LTE system breaks new ground in terms of its degree of support for self-optimization and self-configuration of the network via the X2, S1 and Uu interfaces; these aspects are described in more detail in Chapter 25.

References[16]

[1] 3GPP Technical Specification 24.301, 'Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS); Stage 3', www.3gpp.org.

[2] 3GPP Technical Specification 33.401, 'System Architecture Evolution (SAE): Security Architecture', www.3gpp.org.

[3] 3GPP Technical Specification 23.402, 'Architecture enhancements for non-3GPP accesses', www.3gpp.org.

[4] 3GPP Technical Specification 29.060, 'General Packet Radio Service (GPRS); GPRS Tunnelling Protocol (GTP) across the Gn and Gp interface', www.3gpp.org.

[5] 3GPP Technical Specification 23.203, 'Policy and charging control architecture', www.3gpp.org.

[6] 3GPP Technical Specification 36.300, 'Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2', www.3gpp.org.

[7] 3GPP Technical Specification 23.272, ' Circuit Switched (CS) fallback in Evolved Packet System (EPS); Stage 2', www.3gpp.org.

[8] 3GPP Technical Specification 23.272, 'Single Radio Voice Call Continuity (SRVCC); Stage 2', www.3gpp.org.

[9] Request for Comments 4960 The Internet Engineering Task Force (IETF), Network Working Group, 'Stream Control Transmission Protocol', http://www.ietf.org.

[16] All web sites confirmed 1st March 2011.

3

Control Plane Protocols

Himke van der Velde

3.1 Introduction

As introduced in Section 2.2.2, the Control Plane of the Access Stratum (AS) handles radio-specific functionalities. The AS interacts with the Non-Access Stratum (NAS), also referred to as the 'upper layers'. Among other functions, the NAS control protocols handle Public Land Mobile Network[1] (PLMN) selection, tracking area update, paging, authentication and Evolved Packet System (EPS) bearer establishment, modification and release.

The applicable AS-related procedures largely depend on the Radio Resource Control (RRC) state of the User Equipment (UE), which can be either RRC_IDLE or RRC_CONNECTED.

A UE in RRC_IDLE performs cell selection and reselection – in other words, it decides on which cell to camp. The cell (re)selection process takes into account the priority of each applicable frequency of each applicable Radio Access Technology (RAT), the radio link quality and the cell status (i.e. whether a cell is barred or reserved). An RRC_IDLE UE monitors a paging channel to detect incoming calls, and also acquires system information. The System Information (SI) mainly consists of parameters by which the network (E-UTRAN) can control the cell (re)selection process.

In RRC_CONNECTED, the E-UTRAN allocates radio resources to the UE to facilitate the transfer of (unicast) data via shared data channels.[2] To support this operation, the UE monitors an associated control channel[3] used to indicate the dynamic allocation of the shared transmission resources in time and frequency. The UE provides the network with reports of its

[1]The network of one operator in one country.

[2]The Physical Downlink Shared CHannel (PDSCH) and Physical Uplink Shared CHannel (PUSCH)– see Sections 9.2.2 and 16.2 respectively.

[3]The Physical Downlink Control CHannel (PDCCH) – see Section 9.3.5.

LTE – The UMTS Long Term Evolution: From Theory to Practice, Second Edition.
Stefania Sesia, Issam Toufik and Matthew Baker.
© 2011 John Wiley & Sons, Ltd. Published 2011 by John Wiley & Sons, Ltd.

buffer status and of the downlink channel quality, as well as neighbouring cell measurement information to enable E-UTRAN to select the most appropriate cell for the UE. These measurement reports include cells using other frequencies or RATs. The UE also receives SI, consisting mainly of information required to use the transmission channels. To extend its battery lifetime, a UE in RRC_CONNECTED may be configured with a Discontinuous Reception (DRX) cycle.

RRC, as specified in [1], is the protocol by which the E-UTRAN controls the UE behaviour in RRC_CONNECTED. RRC also includes the control signalling applicable for a UE in RRC_IDLE, namely paging and SI. The UE behaviour in RRC_IDLE is specified in [2].

Chapter 22 gives some further details of the UE measurements which support the mobility procedures.

Functionality related to Multimedia Broadcast/Multicast Services (MBMSs) is covered separately in Chapter 13.

3.2 Radio Resource Control (RRC)

3.2.1 Introduction

The RRC protocol supports the transfer of *common* NAS information (i.e. NAS information which is applicable to all UEs) as well as *dedicated* NAS information (which is applicable only to a specific UE). In addition, for UEs in RRC_IDLE, RRC supports notification of incoming calls (via paging).

The RRC protocol covers a number of functional areas:

- **System information** handles the broadcasting of SI, which includes NAS common information. Some of the system information is applicable only for UEs in RRC_IDLE while other SI is also applicable for UEs in RRC_CONNECTED.
- **RRC connection control** covers all procedures related to the establishment, modification and release of an RRC connection, including paging, initial security activation, establishment of Signalling Radio Bearers (SRBs) and of radio bearers carrying user data (Data Radio Bearers, DRBs), handover within LTE (including transfer of UE RRC context information[4]), configuration of the lower protocol layers,[5] access class barring and radio link failure.
- **Network controlled inter-RAT mobility** includes handover, cell change orders and redirection upon connection release, security activation and transfer of UE RRC context information.
- **Measurement configuration and reporting** for intra-frequency, inter-frequency and inter-RAT mobility, includes configuration and activation of measurement gaps.
- **Miscellaneous functions** including, for example, transfer of dedicated NAS information and transfer of UE radio access capability information.

[4]This UE context information includes the radio resource configuration including local settings not configured across the radio interface, UE capabilities and radio resource management information.

[5]Packet Data Convergence Protocol (PDCP), Radio Link Control (RLC), Medium Access Control (MAC), all of which are explained in detail in Chapter 4, and the physical layer which is explained in Chapters 5–11 and 14–18.

Samsung Ex. 1008
72 of 126

Dedicated RRC messages are transferred across SRBs, which are mapped via the PDCP and RLC layers onto logical channels – either the Common Control CHannel (CCCH) during connection establishment or a Dedicated Control CHannel (DCCH) in RRC_CONNECTED. System Information and Paging messages are mapped directly to logical channels – the Broadcast Control CHannel (BCCH) and Paging Control CHannel (PCCH) respectively. The various logical channels are described in more detail in Section 4.4.1.2.

SRB0 is used for RRC messages which use the CCCH, SRB1 is for RRC messages using DCCH, and SRB2 is for the (lower-priority) RRC messages using DCCH which only include NAS dedicated information.[6] All RRC messages using DCCH are integrity-protected and ciphered by the PDCP layer (after security activation) and use Automatic Repeat reQuest (ARQ) protocols for reliable delivery through the RLC layer. The RRC messages using CCCH are not integrity-protected and do not use ARQ in the RLC layer.

It should also be noted that the NAS independently applies integrity protection and ciphering.

Figure 3.1 illustrates the overall radio protocol architecture as well as the use of radio bearers, logical channels, transport channels and physical channels.

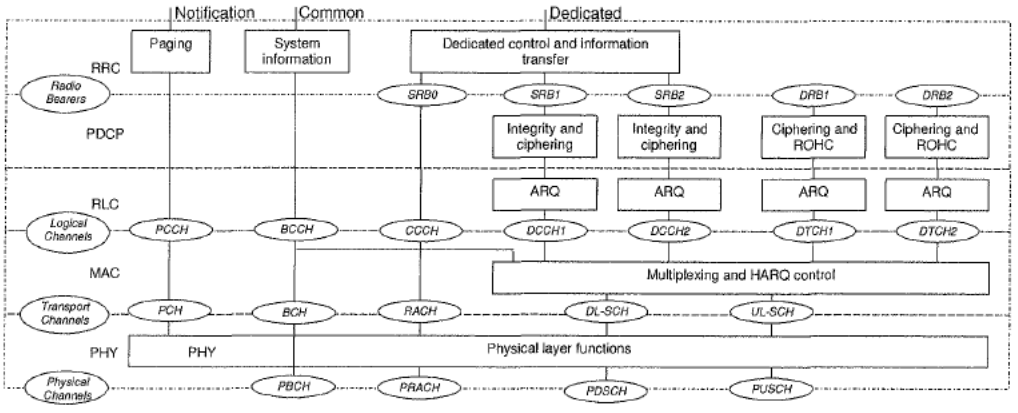


Figure 3.1: Radio architecture.

For control information for which low transfer delay is more important than reliable transfer (i.e. for which the use of ARQ is inappropriate due to the additional delay it incurs), MAC signalling is used provided that there are no security concerns (integrity protection and ciphering are not applicable for MAC signalling).

3.2.2 System Information

System information is structured by means of System Information Blocks (SIBs), each of which contains a set of functionally-related parameters. The SIB types that have been defined include:

[6]Prior to SRB2 establishment, SRB1 is also used for RRC messages which only include NAS dedicated information. In addition, SRB1 is used for higher priority RRC messages which only include NAS dedicated information.

- **The Master Information Block** (MIB), which includes a limited number of the most frequently transmitted parameters which are essential for a UE's initial access to the network.
- **System Information Block Type 1** (SIB1), which contains parameters needed to determine if a cell is suitable for cell selection, as well as information about the time-domain scheduling of the other SIBs.
- **System Information Block Type 2** (SIB2), which includes common and shared channel information.
- **SIB3–SIB8**, which include parameters used to control intra-frequency, inter-frequency and inter-RAT cell reselection.
- **SIB9**, which is used to signal the name of a Home eNodeB (HeNBs).
- **SIB10–SIB12**, which include the Earthquake and Tsunami Warning Service (ETWS) notifications and Commercial Mobile Alert System (CMAS) warning messages (See Section 13.7).
- **SIB13**, which includes MBMS related control information (See Section 13.6.3.2.

Three types of RRC message are used to transfer system information: the MIB message, the SIB1 message and SI messages. An SI message, of which there may be several, includes one or more SIBs which have the same scheduling requirements (i.e. the same transmission periodicity). Table 3.1 provides an example of a possible system information scheduling configuration, also showing which SIBs the UE has to acquire in the idle and connected states. The physical channels used for carrying the SI are explained in Section 9.2.1.

Table 3.1: Example of SI scheduling configuration.

Message	Content	Period (ms)	Applicability
MIB	Most essential parameters	40	Idle and connected
SIB1	Cell access related parameters, scheduling information	80	Idle and connected
1st SI	SIB2: Common and shared channel configuration	160	Idle and connected
2nd SI	SIB3: Common cell reselection information and intra-frequency cell reselection parameters other than the neighbouring cell information SIB4: Intra-frequency neighbouring cell information	320	Idle only
3rd SI	SIB5: Inter-frequency cell reselection information	640	Idle only
4th SI	SIB6: UTRA cell reselection information SIB7: GERAN cell reselection information	640	Idle only, depending on UE support of UMTS or GERAN

Samsung Ex. 1008
74 of 126

3.2.2.1 Time-Domain Scheduling of System Information

The time-domain scheduling of the MIB and SIB1 messages is fixed with a periodicities of 40 ms and 80 ms respectively, as explained in Sections 9.2.1 and 9.2.2.2.

The time-domain scheduling of the SI messages is dynamically flexible: each SI message is transmitted in a defined periodically-occurring time-domain window, while physical layer control signalling[7] indicates in which subframes[8] within this window the SI is actually scheduled. The scheduling windows of the different SI messages (referred to as SI-windows) are consecutive (i.e. there are neither overlaps nor gaps between them) and have a common length that is configurable. SI-windows can include subframes in which it is not possible to transmit SI messages, such as subframes used for SIB1, and subframes used for the uplink in TDD.

Figure 3.2 illustrates an example of the time-domain scheduling of SI, showing the subframes used to transfer the MIB, SIB1 and four SI messages. The example uses an SI-window of length 10 subframes, and shows a higher number of 'blind' Hybrid ARQ (HARQ) transmissions[9] being used for the larger SI messages.

SI messages may have different periodicities. Consequently, in some clusters of SI-windows all the SI messages are scheduled, while in other clusters only the SIs with shorter repetition periods are transmitted. For the example of Table 3.1, the cluster of SI-windows beginning at System Frame Number (SFN) 0 contains all the SI messages, the cluster starting at SFN160 contains only the first SI message, that beginning at SFN320 contains the first and second SI messages, and the one starting at SFN480 contains only the first SI message.

Note that Figure 3.2 shows a cluster of SI-windows where all the SI messages are transmitted. At occasions where a given SI is not transmitted (due to a longer repetition period), its corresponding SI-window is not used.

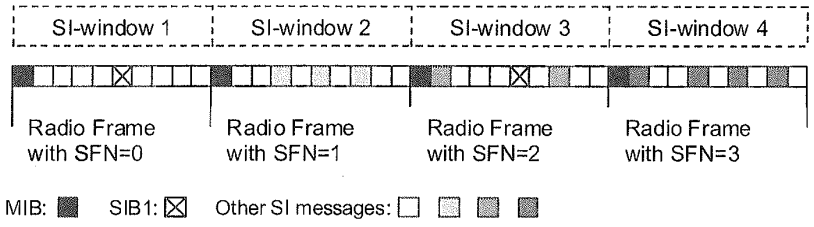


Figure 3.2: SI scheduling example.

3.2.2.2 Validity and Change Notification

SI normally changes only at specific radio frames whose System Frame Number is given by SFN mod $N = 0$, where N is configurable and defines the period between two radio frames at which a change may occur, known as the *modification period*. Prior to performing a change

[7]The Physical Downlink Control Channel – PDCCH; see Section 9.3.5.

[8]A subframe in LTE has a duration of 1 ms; see Section 6.2

[9]With blind HARQ retransmissions, there is no feedback to indicate whether the reception has been successful.

of the system information, the E-UTRAN notifies the UEs by means of a *Paging* message including a *SystemInfoModification* flag. Figure 3.3 illustrates the change of SI, with different shading indicating different content.

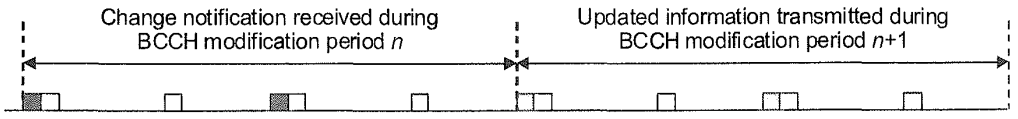


Figure 3.3: SI modification periods. Reproduced by permission of © 3GPP.

LTE provides two mechanisms for indicating that SI has changed:

1. A paging message including a flag indicating whether or not SI has changed.
2. A value tag in SIB1 which is incremented every time one or more SI message changes.

UEs in RRC_IDLE use the first mechanism, while UEs in RRC_CONNECTED can use either mechanism; the second being useful, for example, in cases when a UE was unable to receive the paging messages.

UEs in RRC_IDLE are only required to receive the paging message at their normal paging occasions – i.e. no additional wake-ups are expected to detect changes of SI. In order to ensure reliability of reception, the change notification paging message is normally repeated a number of times during the BCCH modification period preceding that in which the new system information is first transmitted. Correspondingly, the modification period is expressed as a multiple of the cell-specific default paging cycle.

UEs in RRC_CONNECTED are expected to try receiving a paging message the same number of times per modification period as UEs in RRC_IDLE using the default paging cycle. The exact times at which UEs in RRC_CONNECTED which are using this method have to try to receive a paging message are not specified; the UE may perform these tries at convenient times, such as upon wake-up from DRX, using any of the subframes which are configured for paging during the modification period. Since the eNodeB anyway has to notify all the UEs in RRC_IDLE, it has to send a paging message in all subframes which are configured for paging (up to a maximum of four subframes per radio frame) during an entire modification period. Connected mode UEs can utilize any of these subframes. The overhead of transmitting paging messages to notify UEs of a change of SI is considered marginal, since such changes are expected to be infrequent – at most once every few hours.

If the UE receives a notification of a change of SI, it starts acquiring SI from the start of the next modification period. Until the UE has successfully acquired the updated SI, it continues to use the existing parameters. If a critical parameter changes, the communication may be seriously affected, but any service interruption that may result is considered acceptable since it is short and infrequent.

If the UE returns to a cell, it is allowed to assume that the SI previously acquired from the cell remains valid if it was received less than 3 hours previously and the value tag matches.

3.2.3 Connection Control within LTE

Connection control involves:

- Security activation;
- Connection establishment, modification and release;
- DRB establishment, modification and release;
- Mobility within LTE.

3.2.3.1 Security Key Management

Security is a very important feature of all 3GPP RATs. LTE provides security in a similar way to its predecessors UMTS and GSM.

Two functions are provided for the maintenance of security: *ciphering* of both control plane (RRC) data (i.e. SRBs 1 and 2) and user plane data (i.e. all DRBs), and *integrity protection* which is used for control plane (RRC) data only. Ciphering is used in order to protect the data streams from being received by a third party, while integrity protection allows the receiver to detect packet insertion or replacement. RRC always activates both functions together, either following connection establishment or as part of the handover to LTE.

The hierarchy of keys by which the AS security keys are generated is illustrated in Figure 3.4. The process is based on a common secret key K_{ASME} (Access Security Management Entity) which is available only in the Authentication Centre in the Home Subscriber Server (HSS) (see Section 2.2.1) and in a secure part of the Universal Subscriber Identity Module (USIM) in the UE. A set of keys and checksums are generated at the Authentication Centre using this secret key and a random number. The generated keys, checksums and random number are transferred to the Mobility Management Entity (MME) (see Section 2.2.1), which passes one of the generated checksums and the random number to the UE. The USIM in the UE then computes the same set of keys using the random number and the secret key. Mutual authentication is performed by verifying the computed checksums in the UE and network using NAS protocols.

Upon connection establishment, the AS derives an *AS base-key* K_{eNB} (eNodeB-specific) and Next Hop (NH), from K_{ASME}.

K_{eNB} is used to generate three further security keys known as the *AS derived-keys*: one, called $K_{RRC\,int}$, is used for integrity protection of the RRC signalling (SRBs), one for ciphering of the RRC signalling known as $K_{RRC\,enc}$ and $K_{UP\,enc}$ used for ciphering of user data (i.e. DRBs).

NH is an intermediate key used to implement 'forward security'[10] [3]. It is derived by the UE and MME using K_{ASME} and K_{eNB} when the security context is established or using K_{ASME} and the previous NH otherwise. NH is associated with a counter called Next hop Chaining Counter (NCC) which is initially set to 0 at connection establishment.

In case of handover within E-UTRAN, a new AS base-key and new AS Derived-keys are computed from the AS base-key used in the source cell. An intermediate key, K_{eNB^*} is derived by the UE and the source eNodeB based on the Physical Cell Identity (PCI) of the target cell,

[10]Forward security refers to the property that, for an eNodeB sharing a K_{eNB} with a UE, it shall be computationally infeasible to predict any future K_{eNB}, that will be used between the same UE and another eNodeB

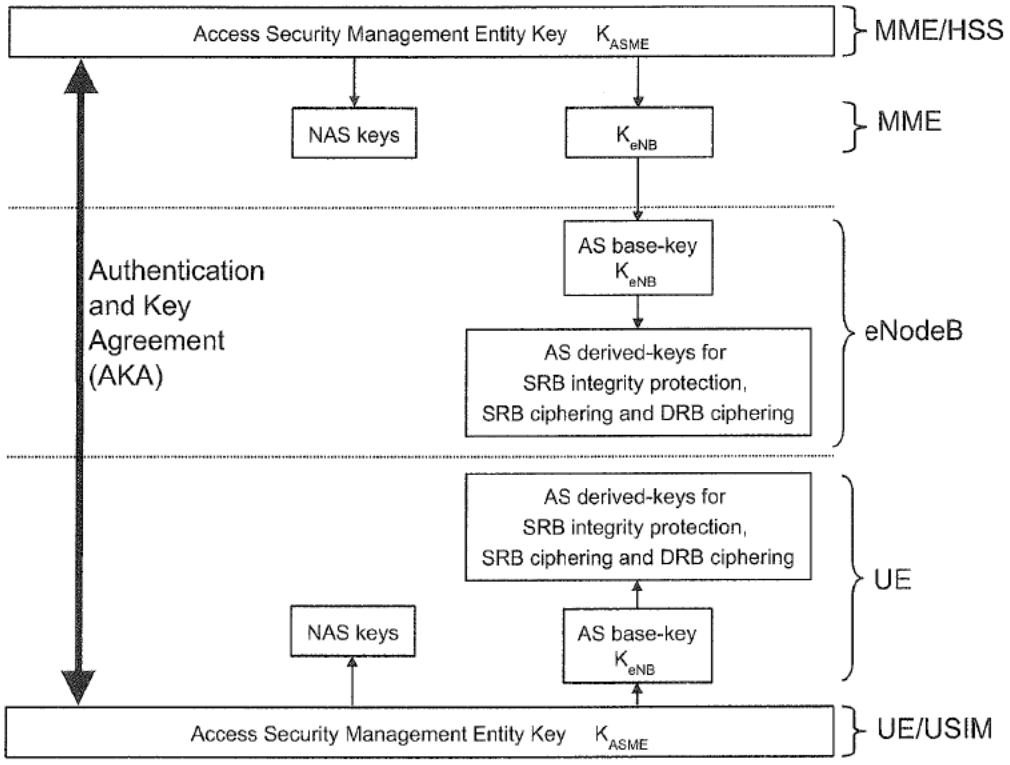


Figure 3.4: Security key derivation.

the target frequency and NH or K_{eNB}. If a fresh NH is available[11], the derivation of K_{eNB^*} is based on NH (referred to as vertical derivation). If no fresh NH is available then the K_{eNB^*} derivation is referred to as horizontal derivation and is based on K_{eNB}. K_{eNB^*} is then used at the target cell as the new K_{eNB} for RRC and data traffic.

For handover to E-UTRAN from UTRAN or GERAN, the AS base-key is derived from integrity and ciphering keys used in the UTRAN or GERAN. Handover within LTE may be used to take a new K_{ASME} into account, i.e. following a re-authentication by NAS.

The use of the security keys for the integrity protection and ciphering functions is handled by the PDCP layer, as described in Section 4.2.3.

The security functions are never deactivated, although it is possible to apply a 'NULL' ciphering algorithm. The 'NULL' algorithm may also be used in certain special cases, such as for making an emergency call without a USIM.

[11] In this case the NCC is incremented and is then larger than that of the currently active K_{eNB^*}

3.2.3.2 Connection Establishment and Release

Two levels of NAS states reflect the state of a UE in respect of connection establishment: the EPS Mobility Management (EMM) state (EMM-DEREGISTERED or EMM-REGISTERED) reflects whether the UE is registered in the MME, and the EPS Connection Management (ECM) state (ECM-IDLE or ECM-CONNECTED) reflects the connectivity of the UE with the Evolved Packet Core (EPC – see Chapter 2).

The NAS states, and their relationship to the AS RRC states, are illustrated in Figure 3.5.

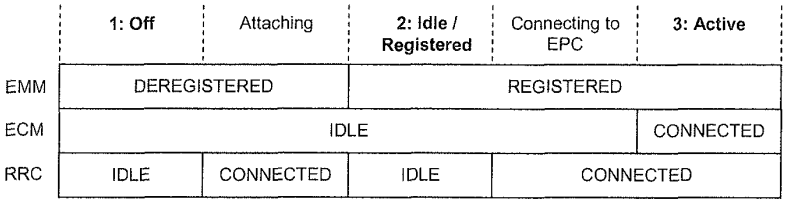


Figure 3.5: Possible combinations of NAS and AS states.

The transition from ECM-IDLE to ECM-CONNECTED not only involves establishment of the RRC connection but also includes establishment of the S1-connection (see Section 2.5). RRC connection establishment is initiated by the NAS and is completed prior to S1-connection establishment, which means that connectivity in RRC_CONNECTED is initially limited to the exchange of control information between UE and E-UTRAN.

UEs are typically moved to ECM-CONNECTED when becoming active. It should be noted, however, that in LTE the transition from ECM-IDLE to ECM-CONNECTED is performed within 100 ms. Hence, UEs engaged in intermittent data transfer need not be kept in ECM-CONNECTED if the ongoing services can tolerate such transfer delays. In any case, an aim in the design of LTE was to support similar battery power consumption levels for UEs in RRC_CONNECTED as for UEs in RRC_IDLE.

RRC connection release is initiated by the eNodeB following release of the S1-connection between the eNodeB and the Core Network (CN).

Connection establishment message sequence. RRC connection establishment involves the establishment of SRB1 and the transfer of the initial uplink NAS message. This NAS message triggers the establishment of the S1-connection, which normally initiates a subsequent step during which E-UTRAN activates AS-security and establishes SRB2 and one or more DRBs (corresponding to the default and optionally dedicated EPS bearers).

Figure 3.6 illustrates the RRC connection establishment procedure, including the subsequent step of initial security activation and radio bearer establishment.

Step 1: Connection establishment

- Upper layers in the UE trigger connection establishment, which may be in response to paging. The UE checks if access is barred (see Section 3.3.4.6). If this is not the case, the lower layers in the UE perform a contention-based random access procedure

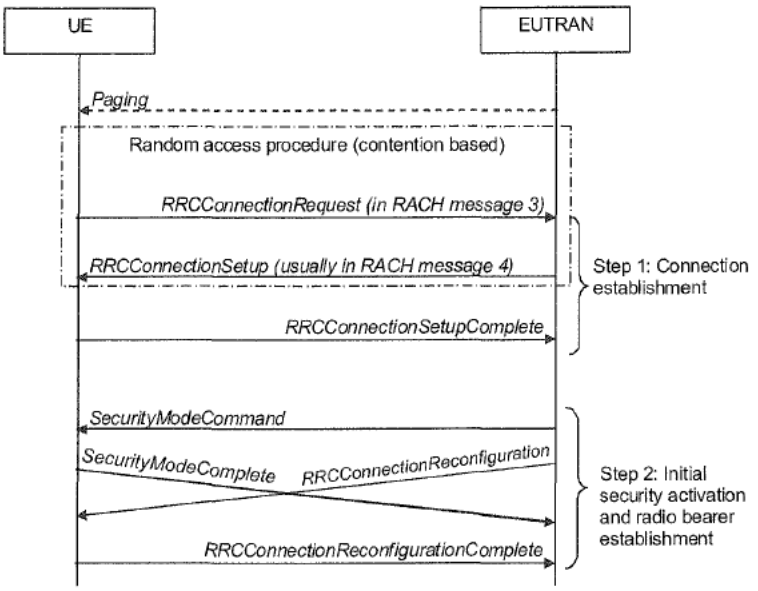


Figure 3.6: Connection establishment (see Section 17.3.1 for details of the contention-based RACH procedure).

as described in Section 17.3, and the UE starts a timer (known as *T300*) and sends the RRCConnectionRequest message. This message includes an initial identity (S-TMSI[12] or a random number) and an establishment cause.

- If E-UTRAN accepts the connection, it returns the RRCConnectionSetup message that includes the initial radio resource configuration including SRB1. Instead of signalling each individual parameter, E-UTRAN may order the UE to apply a default configuration – i.e. a configuration for which the parameter values are specified in the RRC specification [1].
- The UE returns the RRCConnectionSetupComplete message and includes the NAS message, an identifier of the selected PLMN (used to support network sharing) and, if provided by upper layers, an identifier of the registered MME. Based on the last two parameters, the eNodeB decides on the CN node to which it should establish the S1-connection.

Step 2: Initial security activation and radio bearer establishment

- E-UTRAN sends the SecurityModeCommand message to activate integrity protection and ciphering. This message, which is integrity-protected but not ciphered, indicates which algorithms shall be used.
- The UE verifies the integrity protection of the SecurityModeControl message, and, if this succeeds, it configures lower layers to apply integrity protection and ciphering to all subsequent messages (with the exception that ciphering is not applied to the

[12] S-Temporary Mobile Subscriber Identity.

response message, i.e. the SecurityModeComplete (or SecurityModeFailure) message).

- E-UTRAN sends the RRCConnectionReconfiguration message including a radio resource configuration used to establish SRB2 and one or more DRBs. This message may also include other information such as a piggybacked NAS message or a measurement configuration. E-UTRAN may send the RRCConnectionReconfiguration message prior to receiving the SecurityModeComplete message. In this case, E-UTRAN should release the connection when one (or both) procedures fail (because the two procedures result from a single S1-procedure, which does not support partial success).
- The UE finally returns the RRCConnectionReconfigurationComplete message.

A connection establishment may fail for a number of reasons, such as the following:

- Access may be barred (see Section 3.3.4.6).
- In case cell reselection occurs during connection establishment, the UE aborts the procedure and informs upper layers of the failure to establish the connection.
- E-UTRAN may temporarily reject the connection establishment by including a wait timer, in which case the UE rejects any connection establishment request until the wait time has elapsed.
- The NAS may abort an ongoing RRC connection establishment, for example upon NAS timer expiry.

3.2.3.3 DRB Establishment

To establish, modify or release DRBs, E-UTRAN applies the RRC connection reconfiguration procedure as described in Section 3.2.3.2.

When establishing a DRB, E-UTRAN decides how to transfer the packets of an EPS bearer across the radio interface. An EPS bearer is mapped (1-to-1) to a DRB, a DRB is mapped (1-to-1) to a DTCH (Dedicated Traffic CHannel – see Section 4.4.1.2) logical channel, all logical channels are mapped (n-to-1) to the Downlink or Uplink Shared Transport CHannel (DL-SCH or UL-SCH), which are mapped (1-to-1) to the corresponding Physical Downlink or Uplink Shared CHannel (PDSCH or PUSCH). This radio bearer mapping is illustrated in Figure 3.1.

The radio resource configuration covers the configuration of the PDCP, RLC, MAC and physical layers. The main configuration parameters / options include the following:

- For services using small packet sizes (e.g. VoIP), PDCP may be configured to apply indexheader!compressionheader compression to significantly reduce the signalling overhead.
- The RLC Mode is selected from those listed in Section 4.3.1. RLC Acknowledged Mode (AM) is applicable, except for services which require a very low transfer delay and for which reliable transfer is less important.
- E-UTRAN assigns priorities and Prioritized Bit-Rates (PBRs) to control how the UE divides the granted uplink resources between the different radio bearers (see Section 4.4.2.6).

- Unless the transfer delay requirements for any of the ongoing services are very strict, the UE may be configured with a DRX cycle (see Section 4.4.2.5).
- For services involving a semi-static packet rate (e.g. VoIP), semi-persistent scheduling may be configured to reduce the control signalling overhead (see Section 4.4.2.1). Specific resources may also be configured for reporting buffer status and radio link quality.
- Services tolerating higher transfer delays may be configured with a HARQ profile involving a higher average number of HARQ transmissions.

3.2.3.4 Mobility Control in RRC_IDLE and RRC_CONNECTED

Mobility control in RRC_IDLE is UE-controlled (cell-reselection), while in RRC_CONNECTED it is controlled by the E-UTRAN (handover). However, the mechanisms used in the two states need to be consistent so as to avoid ping-pong (i.e. rapid handing back and forth) between cells upon state transitions. The mobility mechanisms are designed to support a wide variety of scenarios including network sharing, country borders, home deployment and varying cell ranges and subscriber densities; an operator may, for example, deploy its own radio access network in populated areas and make use of another operator's network in rural areas.

If a UE were to access a cell which does not have the best radio link quality of the available cells on a given frequency, it may create significant interference to the other cells. Hence, as for most technologies, radio link quality is the primary criterion for selecting a cell on an LTE frequency. When choosing between cells on different frequencies or RATs the interference concern does not apply. Hence, for inter-frequency and inter-RAT cell reselection other criteria may be considered such as UE capability, subscriber type and call type. As an example, UEs with no (or limited) capability for data transmission may be preferably handled on GSM, while home customers or 'premium subscribers' might be given preferential access to the frequency or RAT supporting the highest data rates. Furthermore, in some LTE deployment scenarios, voice services may initially be provided by a legacy RAT only (as a Circuit Switching (CS) application), in which case the UE needs to be moved to the legacy RAT upon establishing a voice call (also referred to as *CS FallBack* (CSFB)).

E-UTRAN provides a list of neighbouring frequencies and cells which the UE should consider for cell reselection and for reporting of measurements. In general, such a list is referred to as a *white-list* if the UE is to consider only the listed frequencies or cells – i.e. other frequencies or cells are not available; conversely, in the case of a *black-list* being provided, a UE may consider any *un*listed frequencies or cells. In LTE, white-listing is used to indicate all the neighbouring frequencies of each RAT that the UE is to consider. On the other hand, E-UTRAN is not required to indicate all the neighbouring cells that the UE shall consider. Which cells the UE is required to detect by itself depends on the UE state as well as on the RAT, as explained below.

Note that for GERAN, typically no information is provided about individual cells. Only in specific cases, such as at country borders, is signalling[13] provided to indicate the group of cells that the UE is to consider – i.e. a white cell list.

[13]The 'NCC-permitted' parameter – see GERAN specifications.

Mobility in idle mode. In RRC_IDLE, cell reselection between frequencies is based on absolute priorities, where each frequency has an associated priority. Cell-specific default values of the priorities are provided via SI. In addition, E-UTRAN may assign UE-specific values upon connection release, taking into account factors such as UE capability or subscriber type. In case equal priorities are assigned to multiple cells, the cells are ranked based on radio link quality. Equal priorities are not applicable between frequencies of different RATs. The UE does not consider frequencies for which it does not have an associated priority; this is useful in situations such as when a neighbouring frequency is applicable only for UEs of one of the sharing networks.

Table 3.2 provides an overview of the SI parameters which E-UTRAN may use to control cell reselection . Other than the cell reselection priority of a frequency, no idle mode mobility-related parameters may be assigned via dedicated signalling. Further details of the parameters listed are provided in Section 3.3.

Mobility in connected mode. In RRC_CONNECTED, the E-UTRAN decides to which cell a UE should hand over in order to maintain the radio link. As with RRC_IDLE, E-UTRAN may take into account not only the radio link quality but also factors such as UE capability, subscriber type and access restrictions. Although E-UTRAN may trigger handover without measurement information (blind handover), normally it configures the UE to report measurements of the candidate target cells – see Section 22.3. Table 3.3 provides an overview of the frequency- and cell-specific parameters which E-UTRAN can configure for mobility-related measurement reporting.

In LTE the UE always connects to a single cell only – in other words, the switching of a UE's connection from a source cell to a target cell is a *hard* handover. The hard handover process is normally a 'backward' one, whereby the eNodeB which controls the source cell requests the target eNodeB to prepare for the handover. The target eNodeB subsequently generates the RRC message to order the UE to perform the handover, and the message is transparently forwarded by the source eNodeB to the UE. LTE also supports a kind of 'forward' handover, in which the UE by itself decides to connect to the target cell, where it then requests that the connection be continued. The UE applies this connection re-establishment procedure only after loss of the connection to the source cell; the procedure only succeeds if the target cell has been prepared in advance for the handover.

Besides the handover procedure, LTE also provides for a UE to be redirected to another frequency or RAT upon connection release. Redirection during connection establishment is not supported, since at that time the E-UTRAN may not yet be in possession of all the relevant information such as the capabilities of the UE and the type of subscriber (as may be reflected, for example, by the SPID, the Subscriber Profile ID for RAT/Frequency Priority). However, the redirection may be performed while AS-security has not (yet) been activated. When redirecting the UE to UTRAN or GERAN, E-UTRAN may provide SI for one or more cells on the relevant frequency. If the UE selects one of the cells for which SI is provided, it does not need to acquire it.

Message sequence for handover within LTE. In RRC_CONNECTED, the E-UTRAN controls mobility by ordering the UE to perform handover to another cell, which may be

Table 3.2: List of SI parameters which may be used to control cell reselection.

Parameter	Intra-Freq.	Inter-Freq.	UTRA	GERAN	CDMA2000
Common	(SIB3)	(SIB5)	(SIB6)	(SIB7)	(SIB8)
Reselection info	Q-Hyst MobilityStatePars Q-HystSF S-Search[a]		T-Reselect T-ReselectSF	T-Reselect T-ReselectSF	T-Reselect T-ReselectSF
Frequency list	(SIB3)	(SIB5)	(SIB6)	(SIB7)	(SIB8)
White frequency list	n/a	+	+	+	+
Frequency specific reselection info[b]	Priority $Thresh_{Serving\text{-}Low}$ $Thresh_{Serving\text{-}LowQ}$ T-Reselect T-ReselectSF $Thresh_{X\text{-}LowQ}$ T-Reselect T-ReselectSF	Priority Qoffset, $Thresh_{X\text{-}High}$, $Thresh_{X\text{-}Low}$ $Thresh_{X\text{-}HighQ}$,	Priority $Thresh_{X\text{-}High}$, $Thresh_{X\text{-}Low}$ $Thresh_{X\text{-}HighQ}$, $Thresh_{X\text{-}LowQ}$	Priority $Thresh_{X\text{-}High}$, $Thresh_{X\text{-}Low}$	Priority $Thresh_{X\text{-}High}$, $Thresh_{X\text{-}Low}$
Frequency specific suitability info[c]	Q-RxLevMin MaxTxPower Q-QualMin	Q-RxLevMin MaxTxPower Q-QualMin	Q-RxLevMin, MaxTxPower, Q-QualMin	Q-RxLevMin MaxTxPower	
Cell list	(SIB4)	(SIB5)	(SIB6)	(SIB7)	(SIB8)
White cell list	–	–	–	NCC permitted[d]	–
Black cell list	+	+	–	–	–
List of cells with specific info[c]	Qoffset	Qoffset	–	–	–

[a] Separate parameters for intra/ inter-frequency, both for RSRP and RSRQ.
[b] See Section 3.3.4.2.
[c] See Section 3.3.3.
[d] See GERAN specifications.
[e] See Section 3.3.4.3.

on the same frequency ('intra-frequency') or a different frequency ('inter-frequency'). Inter-frequency measurements may require the configuration of measurement gaps, depending on the capabilities of the UE (e.g. whether it has a dual receiver) – see Section 22.3.

The E-UTRAN may also use the handover procedures for completely different purposes, such as to change the security keys to a new set (see Section 3.2.3.1), or to perform a 'synchronized reconfiguration' in which the E-UTRAN and the UE apply the new configuration simultaneously.

The message sequence for the procedure for handover within LTE is shown in Figure 3.7. The sequence is as follows:

1. The UE may send a MeasurementReport message (see Section 3.2.5).

Table 3.3: Frequency- and cell-specific information which can be configured in connected mode.

Parameter	Intra-Freq.	Inter-Freq.	UTRA	GERAN	CDMA2000
Frequency list					
White frequency list	n/a	+	+	+	+
Frequency specific info	Qoffset	Qoffset	Qoffset	Qoffset	Qoffset
Cell list					
White cell list	–	–	+	NCC permitted	+
Black cell list	+	+	–	–	–
List of cells with specific info.	Qoffset	Qoffset	–	–	–

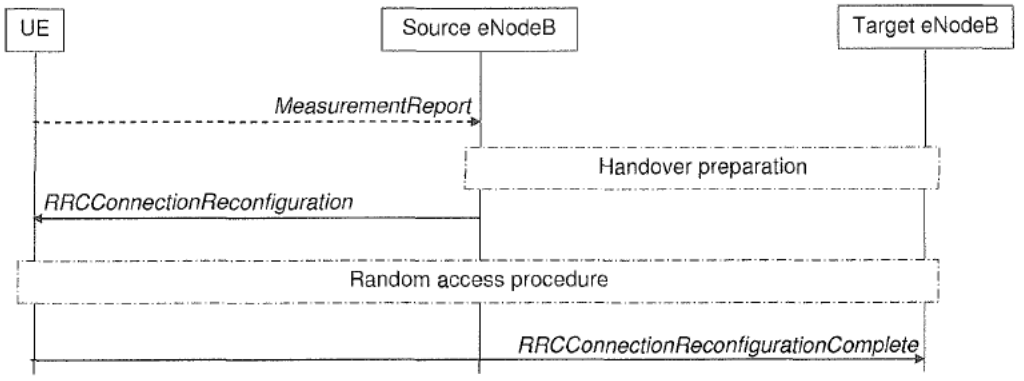


Figure 3.7: Handover within LTE.

2. Before sending the handover command to the UE, the source eNodeB requests one or more target cells to prepare for the handover. As part of this 'handover preparation request', the source eNodeB provides UE RRC context information[14] about the UE capabilities, the current AS-configuration and UE-specific Radio Resource Management (RRM) information. In response, the eNodeB controlling the target cell generates the 'handover command'. The source eNodeB will forward this command to the UE in the RRCConnectionReconfiguration message. This is done transparently (apart from performing integrity protection and ciphering) – i.e. the source eNodeB does not add or modify the protocol information contained in the message.

3. The source eNodeB sends the RRCConnectionReconfiguration message which to the UE orders it to perform handover. It includes mobility control information (namely the identity, and optionally the frequency, of the target cell) and the radio resource

[14]This UE context information includes the radio resource configuration including local settings not configured across the radio interface, UE capabilities and radio resource management information.

Samsung Ex. 1008
85 of 126

configuration information which is common to all UEs in the target cell (e.g. information required to perform random access. The message also includes the dedicated radio resource configuration, the security configuration and the C-RNTI[15] to be used in the target cell. Although the message may optionally include the measurement configuration, the E-UTRAN is likely to use another reconfiguration procedure for re-activating measurements, in order to avoid the RRCConnectionReconfiguration message becoming excessively large. If no measurement configuration information is included in the message used to perform inter-frequency handover, the UE stops any inter-frequency and inter-RAT measurements and deactivates the measurement gap configuration.

4. If the UE is able to comply with the configuration included in the received RRC-ConnectionReconfiguration message, the UE starts a timer, known as *T304*, and initiates a random access procedure (see Section 17.3), using the received Random Access CHannel (RACH) configuration, to the target cell at the first available occasion.[16] It is important to note that the UE does not need to acquire system information from the target cell prior to initiating random access and resuming data communication. However, the UE may be unable to use some parts of the physical layer configuration from the very start (e.g. semi-persistent scheduling (see Section 4.4.2.1), the PUCCH (see Section 16.3) and the Sounding Reference Signal (SRS) (see Section 15.6)). The UE derives new security keys and applies the received configuration in the target cell.

5. Upon successful completion of the random access procedure, the UE stops the timer T304. The AS informs the upper layers in the UE about any uplink NAS messages for which transmission may not have completed successfully, so that the NAS can take appropriate action.

For handover to cells broadcasting a Closed Subscriber Group (CSG) identity, normal measurement and mobility procedures are used to support handover. In addition, E-UTRAN may configure the UE to report that it is entering or leaving the proximity of cell(s) included in its CSG whitelist. Furthermore, E-UTRAN may request the UE to provide additional information broadcast by the handover candidate cell, for example the cell global identity, CSG identity or CSG membership status. E-UTRAN may use a indexproximity report'proximity report' to configure measurements and to decide whether or not to request the UE to provide additional information broadcast by the handover candidate cell. The additional information is used to verify whether or not the UE is authorized to access the target cell and may also be needed to identify handover candidate cells.[17] Further details of the mobility procedures for HeNBs can be found in Section 24.2.3.

[15]The Cell Radio Network Temporary Identifier is the RNTI to be used by a given UE while it is in a particular cell.

[16]The target cell does not specify when the UE is to initiate random access in that cell. Hence, the handover process is sometimes described as *asynchronous*.

[17]This may be the case if PCI confusion occurs, i.e. when the PCI that is included in the measurement report does not uniquely identify the cell.

3.2.3.5 Connection Re-Establishment Procedure

In a number of failure cases (e.g. radio link failure, handover failure, RLC unrecoverable error, reconfiguration compliance failure), the UE initiates the RRC connection re-establishment procedure, provided that security is active. If security is not active when one of the indicated failures occurs, the UE moves to RRC_IDLE instead.

To attempt RRC connection re-establishment, the UE starts a timer known as *T311* and performs cell selection. The UE should prioritize searching on LTE frequencies. However, no requirements are specified regarding for how long the UE shall refrain from searching for other RATs. Upon finding a suitable cell on an LTE frequency, the UE stops the timer T311, starts the timer T301 and initiates a contention based random access procedure to enable the RRCConnectionReestablishmentRequest message to be sent. In the RRCConnectionReestablishmentRequest message, the UE includes the identity used in the cell in which the failure occurred, the identity of that cell, a short Message Authentication Code and a cause.

The E-UTRAN uses the re-establishment procedure to continue SRB1 and to re-activate security without changing algorithms. A subsequent RRC connection reconfiguration procedure is used to resume operation on radio bearers other than SRB1 and to re-activate measurements. If the cell in which the UE initiates the re-establishment is not prepared (i.e. does not have a context for that UE), the E-UTRAN will reject the procedure, causing the UE to move to RRC_IDLE.

3.2.4 Connected Mode Inter-RAT Mobility

The overall procedure for the control of mobility is explained in this section; some further details can be found in Chapter 22.

3.2.4.1 Handover to LTE

The procedure for handover to LTE is largely the same as the procedure for handover within LTE, so it is not necessary to repeat the details here. The main difference is that upon handover to LTE the entire AS-configuration needs to be signalled, whereas within LTE it is possible to use 'delta signalling', whereby only the changes to the configuration are signalled.

If ciphering had not yet been activated in the previous RAT, the E-UTRAN activates ciphering, possibly using the NULL algorithm, as part of the handover procedure. The E-UTRAN also establishes SRB1, SRB2 and one or more DRBs (i.e. at least the DRB associated with the default EPS bearer).

3.2.4.2 Mobility from LTE

Generally, the procedure for mobility from LTE to another RAT supports both handover and Cell Change Order (CCO), possibly with Network Assistance (NACC – Network Assisted Cell Change). The CCO/NACC procedure is applicable only for mobility to GERAN. Mobility from LTE is performed only after security has been activated. When used for enhanced CSFB[18] to CDMA2000, the procedure includes support for parallel handover (i.e.

[18] See Section 2.4.2.1.

to both 1XRTT and HRPD), for handover to 1XRTT in combination with redirection to HRPD, and for redirection to HRPD only.

The procedure is illustrated in Figure 3.8.

1. The UE may send a MeasurementReport message (see Section 3.2.5 for further details).
2. In case of handover (as opposed to CCO), the source eNodeB requests the target Radio Access Network (RAN) node to prepare for the handover. As part of the 'handover preparation request' the source eNodeB provides information about the applicable inter-RAT UE capabilities as well as information about the currently-established bearers. In response, the target RAN generates the 'handover command' and returns this to the source eNodeB.
3. The source eNodeB sends a MobilityFromEUTRACommand message to the UE, which includes either the inter-RAT message received from the target (in case of handover), or the target cell/frequency and a few inter-RAT parameters (in case of CCO).
4. Upon receiving the MobilityFromEUTRACommand message, the UE starts the timer T304 and connects to the target node, either by using the received radio configuration (handover) or by initiating connection establishment (CCO) in accordance with the applicable specifications of the target RAT.

Upper layers in the UE are informed, by the AS of the target RAT, which bearers are established. From this, the UE can derive if some of the established bearers were not admitted by the target RAN node.

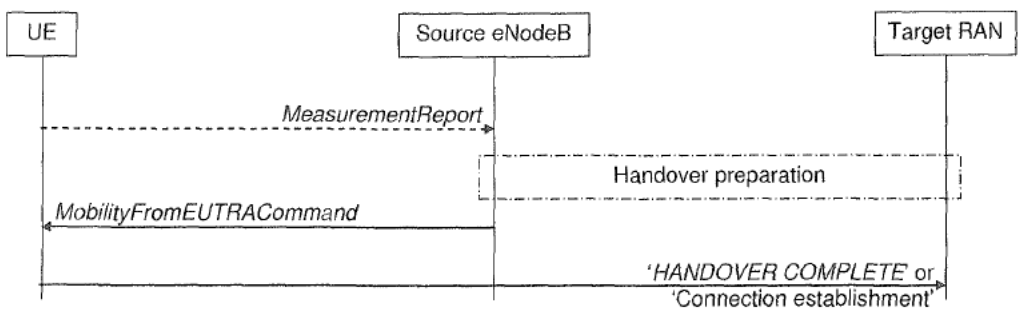


Figure 3.8: Mobility from LTE.

3.2.4.3 CDMA2000

For CDMA2000, additional procedures have been defined to support the transfer of dedicated information from the CDMA2000 upper layers, which are used to register the UE's presence in the target core network prior to performing the handover (referred to as preregistration). These procedures use SRB1.

Samsung Ex. 1008
88 of 126

3.2.5 Measurements

3.2.5.1 Measurement Configuration

The E-UTRAN can configure the UE to report measurement information to support the control of UE mobility. The following measurement configuration elements can be signalled via the RRCConnectionReconfiguration message.

1. **Measurement objects.** A measurement object defines on what the UE should perform the measurements – such as a carrier frequency. The measurement object may include a list of cells to be considered (white-list or black-list) as well as associated parameters, e.g. frequency- or cell-specific offsets.

2. **Reporting configurations.** A reporting configuration consists of the (periodic or event-triggered) criteria which cause the UE to send a measurement report, as well as the details of what information the UE is expected to report (e.g. the quantities, such as Received Signal Code Power (RSCP) (see Section 22.3.2.1) for UMTS or Reference Signal Received Power (RSRP) (see Section 22.3.1.1) for LTE, and the number of cells).

3. **Measurement identities.** These identify a measurement and define the applicable measurement object and reporting configuration.

4. **Quantity configurations**. The quantity configuration defines the filtering to be used on each measurement.

5. **Measurement gaps.** Measurement gaps define time periods when no uplink or downlink transmissions will be scheduled, so that the UE may perform the measurements. The measurement gaps are common for all gap-assisted measurements. Further details of the measurement gaps are discussed in Section 22.2.1.2.

The details of the above parameters depend on whether the measurement relates to an LTE, UMTS, GERAN or CDMA2000 frequency. Further details of the measurements performed by the UE are explained in Section 22.3. The E-UTRAN configures only a single measurement object for a given frequency, but more than one measurement identity may use the same measurement object. The identifiers used for the measurement object and reporting configuration are unique across all measurement types. An example of a set of measurement objects and their corresponding reporting configurations is shown in Figure 3.9.

In LTE it is possible to configure the quantity which triggers the report (RSCP or RSRP) for each reporting configuration. The UE may be configured to report either the trigger quantity or both quantities.

The RRC measurement reporting procedures include some extensions specifically to support Self-Optimizing Network (SON) functions such as the determination of Automatic Neighbour Relations (ANR) – see Section 25.2. The RRC measurement procedures also support UE positioning[19] by means of the enhanced cell identity method – see Section 19.4.

[19] See Chapter 19.

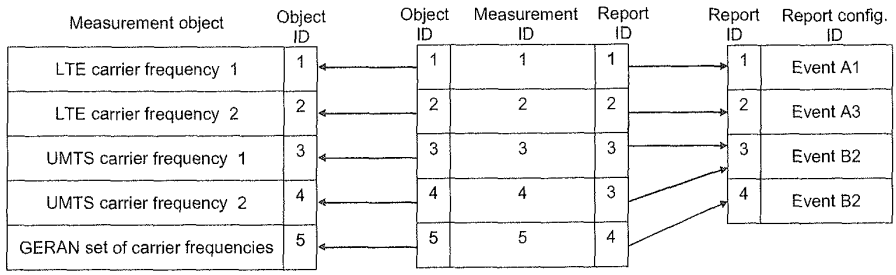


Figure 3.9: Example measurement configuration.

3.2.5.2 Measurement Report Triggering

Depending on the measurement type, the UE may measure and report any of the following:

- The serving cell;
- Listed cells (i.e. cells indicated as part of the measurement object);
- Detected cells on a listed frequency (i.e. cells which are not listed cells but are detected by the UE).

For some RATs, the UE measures and reports listed cells only (i.e. the list is a white-list), while for other RATs the UE also reports detected cells. For further details, see Table 3.3. Additionally, E-UTRAN can configure UTRAN PCI ranges for which the UE is allowed to send a measurement reports (mainly for the support of handover to UTRAN cells broadcasting a CSG identity).

For LTE, the following event-triggered reporting criteria are specified:

- **Event A1.** Serving cell becomes better than absolute threshold.
- **Event A2.** Serving cell becomes worse than absolute threshold.
- **Event A3.** Neighbour cell becomes better than an offset relative to the serving cell.
- **Event A4.** Neighbour cell becomes better than absolute threshold.
- **Event A5.** Serving cell becomes worse than one absolute threshold and neighbour cell becomes better than another absolute threshold.

For inter-RAT mobility, the following event-triggered reporting criteria are specified:

- **Event B1.** Neighbour cell becomes better than absolute threshold.
- **Event B2.** Serving cell becomes worse than one absolute threshold and neighbour cell becomes better than another absolute threshold.

The UE triggers an event when one or more cells meets a specified 'entry condition'. The E-UTRAN can influence the entry condition by setting the value of some configurable parameters used in these conditions – for example, one or more thresholds, an offset, and/or a hysteresis. The entry condition must be met for at least a duration corresponding to a 'timeToTrigger' parameter configured by the E-UTRAN in order for the event to be triggered.

Samsung Ex. 1008
90 of 126

The UE scales the timeToTrigger parameter depending on its speed (see Section 3.3 for further detail).

Figure 3.10 illustrates the triggering of event A3 when a timeToTrigger and an offset are configured.

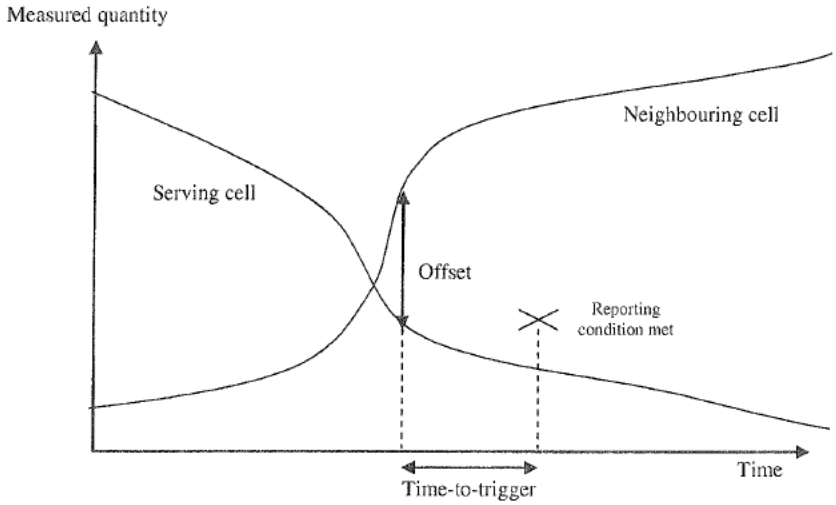


Figure 3.10: Event triggered report condition (Event A3).

The UE may be configured to provide a number of periodic reports after having triggered an event. This 'event-triggered periodic reporting' is configured by means of parameters 'reportAmount' and 'reportInterval', which specify respectively the number of periodic reports and the time period between them. If event-triggered periodic reporting is configured, the UE's count of the number of reports sent is reset to zero whenever a new cell meets the entry condition. The same cell cannot then trigger a new set of periodic reports unless it first meets a specified 'leaving condition'.

In addition to event-triggered reporting, the UE may be configured to perform periodic measurement reporting. In this case, the same parameters may be configured as for event-triggered reporting, except that the UE starts reporting immediately rather than only after the occurrence of an event.

3.2.5.3 Measurement Reporting

In a MeasurementReport message, the UE only includes measurement results related to a single measurement – in other words, measurements are not combined for reporting purposes. If multiple cells triggered the report, the UE includes the cells in order of decreasing value of the reporting quantity – i.e. the best cell is reported first. The number of cells the UE includes in a MeasurementReport may be limited by a parameter 'indexmaxReportCellsmaxReportCells'.

3.2.6 Other RRC Signalling Aspects

3.2.6.1 UE Capability Transfer

In order to avoid signalling of the UE radio access capabilities across the radio interface upon each transition from RRC_IDLE to RRC_CONNECTED, the core network stores the AS capabilities (both the E-UTRA and GERAN capabilities) while the UE is in RRC_IDLE/EMM-REGISTERED. Upon S1 connection establishment, the core network provides the capabilities to the E-UTRAN. If the E-UTRAN does not receive the (required) capabilities from the core network (e.g. due to the UE being in EMM-DEREGISTERED), it requests the UE to provide its capabilities using the UE capability transfer procedure. The E-UTRAN can indicate for each RAT (LTE, UMTS, GERAN) whether it wants to receive the associated capabilities. The UE provides the requested capabilities using a separate container for each RAT. Dynamic change of UE capabilities is not supported, except for change of the GERAN capabilities in RRC_IDLE which is supported by the tracking area update procedure.

3.2.6.2 Uplink/Downlink Information Transfer

The uplink/downlink information transfer procedures are used to transfer only upper layer information (i.e. no RRC control information is included). The procedure supports the transfer of 3GPP NAS dedicated information as well as CDMA2000 dedicated information.

In order to reduce latency, NAS information may also be included in the RRCConnectionSetupComplete and RRCConnectionReconfiguration messages. For the latter message, NAS information is only included if the AS and NAS procedures are dependent (i.e. they jointly succeed or fail). This applies for EPS bearer establishment, modification and release.

As noted earlier, some additional NAS information transfer procedures have also been defined for CDMA2000 for preregistration.

3.2.6.3 UE Information Transfer

The UE information transfer procedure was introduced in Release 9 to support SON (see Chapter 25). The procedure supports network optimization for mobility robustness by the reporting, at a later point in time, of measurement information available when a radio link failure occurs (see Section 25.6). E-UTRAN may also use the UE information transfer procedure to retrieve information regarding the last successful random access, which it may use for RACH optimization – see Section 25.7.

3.3 PLMN and Cell Selection

3.3.1 Introduction

After a UE has selected a PLMN, it performs *cell selection* – in other words, it searches for a suitable cell on which to camp (see Chapter 7). While camping on the chosen cell, the UE acquires the SI that is broadcast (see Section 9.2.1). Subsequently, the UE registers its presence in the tracking area, after which it can receive paging information which is used

to notify UEs of incoming calls. The UE may establish an RRC connection, for example to establish a call or to perform a tracking area update.

When camped on a cell, the UE regularly verifies if there is a better cell; this is known as performing *cell reselection*.

LTE cells are classified according to the service level the UE obtains on them: a *suitable cell* is a cell on which the UE obtains normal service. If the UE is unable to find a suitable cell, but manages to camp on a cell belonging to another PLMN, the cell is said to be an *acceptable cell*, and the UE enters a 'limited service' state in which it can only perform emergency calls (and receive public warning messages) – as is also the case when no USIM is present in the UE. Finally, some cells may indicate via their SI that they are barred or reserved; a UE can obtain no service on such a cell.

A category called 'operator service' is also supported in LTE, which provides normal service but is applicable only for UEs with special access rights.

Figure 3.11 provides a high-level overview of the states and the cell (re)selection procedures.

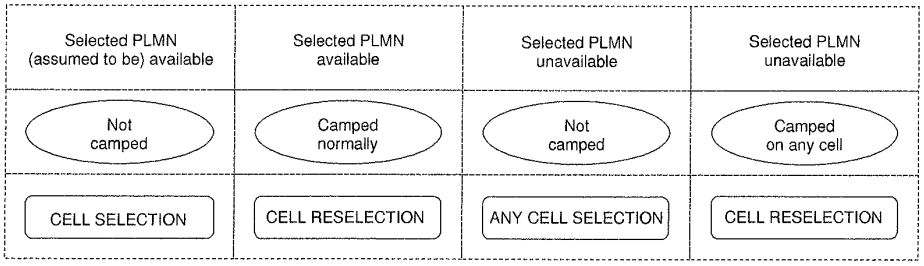


Figure 3.11: Idle mode states and procedures.

3.3.2 PLMN Selection

The NAS handles PLMN selection based on a list of available PLMNs provided by the AS. The NAS indicates the selected PLMN together with a list of equivalent PLMNs, if available. After successful registration, the selected PLMN becomes the *Registered* PLMN (R-PLMN).

The AS may autonomously indicate available PLMNs. In addition, NAS may request the AS to perform a full search for available PLMNs. In the latter case, the UE searches for the strongest cell on each carrier frequency. For these cells, the UE retrieves the PLMN identities from SI. If the quality of a cell satisfies a defined radio criterion, the corresponding PLMNs are marked as *high quality*; otherwise, the PLMNs are reported together with their quality.

3.3.3 Cell Selection

Cell selection consists of the UE searching for the strongest cell on all supported carrier frequencies of each supported RAT until it finds a suitable cell. The main requirement for cell selection is that it should not take too long, which becomes more challenging with the ever increasing number of frequencies and RATs to be searched. The NAS can speed up the

search process by indicating the RATs associated with the selected PLMN. In addition, the UE may use information stored from a previous access.

The cell selection criterion is known as the *S-criterion* and is fulfilled when the cell-selection receive level and the quality level are above a given value: Srxlev > 0 and Squal > 0, where

$$\text{Srxlev} = Q_{\text{rxlevmeas}} - (Q_{\text{rxlevmin}} - Q_{\text{rxlevminoffset}})$$

$$\text{Squal} = Q_{\text{qualmeas}}(Q_{\text{qualmin}} + Q_{\text{qualminoffset}})$$

in which $Q_{\text{rxlevmeas}}$ is the measured cell receive level value, also known as the RSRP (see Section 22.3.1.1), and Q_{rxlevmin} is the minimum required receive level in the cell. Q_{qualmeas} and Q_{qualmin} are the corresponding parameters for the quality level, also known as the RSRQ.

$Q_{\text{rxlevminoffset}}$ and $Q_{\text{qualminoffset}}$ are offsets which may be configured to prevent ping-pong between PLMNs, which may otherwise occur due to fluctuating radio conditions. The offsets are taken into account only when performing a periodic search for a higher priority PLMN while camped on a suitable cell in a visited PLMN.

The cell selection related parameters are broadcast within the SIB1 message.

For some specific cases, additional requirements are defined:

- Upon leaving connected mode, the UE should normally attempt to select the cell to which it was connected. However, the connection release message may include information directing the UE to search for a cell on a particular frequency.
- When performing 'any cell selection', the UE tries to find an acceptable cell of any PLMN by searching all supported frequencies on all supported RATs. The UE may stop searching upon finding a cell that meets the 'high quality' criterion applicable for that RAT.

Note that the UE only verifies the suitability of the strongest cell on a given frequency. In order to avoid the UE needing to acquire SI from a candidate cell that does not meet the S-criterion, suitability information is provided for inter-RAT neighbouring cells.

3.3.4 Cell Reselection

Once the UE camps on a suitable cell, it starts cell reselection. This process aims to move the UE to the 'best' cell of the selected PLMN and of its equivalent PLMNs, if any. As described in Section 3.2.3.4, cell reselection between frequencies and RATs is primarily based on absolute priorities. Hence, the UE first evaluates the frequencies of all RATs based on their priorities. Secondly, the UE compares the cells on the relevant frequencies based on radio link quality, using a ranking criterion. Finally, upon reselecting to the target cell the UE verifies the cell's accessibility. Further rules have also been defined to allow the UE to limit the frequencies to be measured, to speed up the process and save battery power, as discussed in Section 3.3.4.1. Figure 3.12 provides a high-level overview of the cell reselection procedure.

It should be noted that the UE performs cell reselection only after having camped for at least one second on the current serving cell.

Samsung Ex. 1008
94 of 126

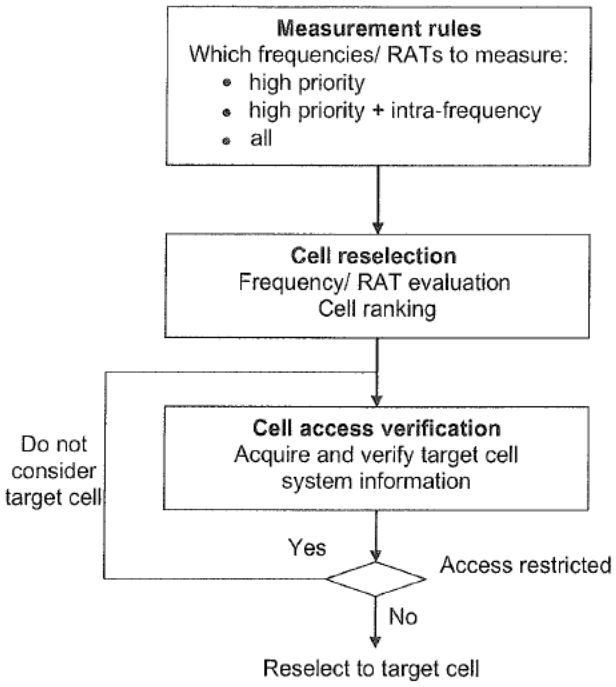


Figure 3.12: Cell reselection.

3.3.4.1 Measurement Rules

To enable the UE to save battery power, rules have been defined which limit the measurements the UE is required to perform. Firstly, the UE is required to perform intra-frequency measurements only when the quality of the serving cell is below or equal to a threshold ('SintraSearch'). Furthermore, the UE is required to measure other frequencies/RATs of lower or equal priority only when the quality of the serving cell is below or equal to another threshold ('SnonintraSearch'). The UE is always required to measure frequencies and RATs of higher priority. For both cases (i.e. intra-frequency and inter-frequency) the UE may refrain from measuring when a receive level and a quality criterion is fulfilled. The required performance (i.e. how often the UE is expected to make the measurements, and to what extent this depends on, for example, the serving cell quality) is specified in [4].

3.3.4.2 Frequency/RAT Evaluation

E-UTRAN configures an absolute priority for all applicable frequencies of each RAT. In addition to the cell-specific priorities which are optionally provided via SI, E-UTRAN can assign UE-specific priorities via dedicated signalling. Of the frequencies that are indicated in the system information, the UE is expected to consider for cell reselection only those for which it has priorities. Equal priorities are not applicable for inter-RAT cell reselection.

The UE reselects to a cell on a higher priority frequency if the S-criterion (see Section 3.3.3) of the concerned target cell exceeds a high threshold ($\text{Thresh}_{\text{X-High}}$) for longer

than a certain duration $T_{reselection}$. The UE reselects to a cell on a lower-priority frequency if the S-criterion of the serving cell is below a low threshold ($Thresh_{Serving\text{-}Low}$) while the S-criterion of the target cell on a lower-priority frequency (possibly on another RAT) exceeds a low threshold ($Thresh_{X\text{-}Low}$) during the time interval $T_{reselection}$, and in the same time no cell on a higher-priority frequency is available. The UE evaluates the thresholds either based on receive level or on quality level, depending on which parameters E-UTRAN configures. Figure 3.13 illustrates the condition(s) to be met for reselecting to a cell on a higher-priority frequency (light grey bar) and to a cell on a lower priority frequency (dark grey bars).

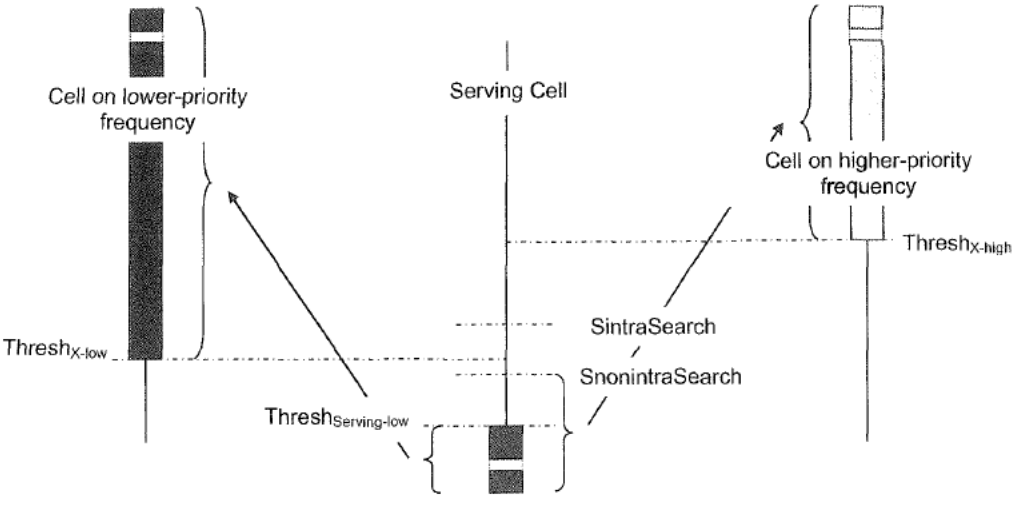


Figure 3.13: Frequency/RAT evaluation.

When reselecting to a frequency, possibly on another RAT, which has a different priority, the UE reselects to the highest-ranked cell on the concerned frequency (see Section 3.3.4.3).

Note that, as indicated in Section 3.2.3.4, thresholds and priorities are configured per frequency, while $T_{reselection}$ is configured per RAT.

From Release 8 onwards, UMTS and GERAN support the same priority-based cell reselection as provided in LTE, with a priority per frequency. Release 8 RANs will continue to handle legacy UEs by means of offset-based ranking. Likewise, Release 8 UEs should apply the ranking based on radio link quality (with offsets) unless UMTS or GERAN indicate support for priority-based reselection.

3.3.4.3 Cell Ranking

The UE ranks the intra-frequency cells and the cells on other frequencies having equal priority which fulfil the S-criterion using a criterion known as the *R-criterion*. The R-criterion generates rankings R_s and R_n for the serving cell and neighbour cells respectively:

$$\text{For the serving cell:} \quad R_s = Q_{meas,s} + Q_{hyst,s}$$

$$\text{For neighbour cells:} \quad R_n = Q_{meas,n} + Q_{offs,n}$$

where Q_{meas} is the measured cell received quality (RSRP) (see Section 22.3.1.1), $Q_{hyst,s}$ is a parameter controlling the degree of hysteresis for the ranking, and $Q_{off\,s,n}$ is an offset applicable between serving and neighbouring cells on frequencies of equal priority (the sum of the cell-specific and frequency-specific offsets).

The UE reselects to the highest-ranked candidate cell provided that it is better ranked than the serving cell for at least the duration of $T_{reselection}$. The UE scales the parameters $T_{reselection}$ and Q_{hyst}, depending on the UE speed (see Section 3.3.4.5 below).

3.3.4.4 Accessibility Verification

If the best cell on an LTE frequency is barred or reserved, the UE is required to exclude this cell from the list of cell reselection candidates. In this case, the UE may consider other cells on the same frequency unless the barred cell indicates (by means of field 'intraFreqReselection' within SIB1) that intra-frequency reselection is not allowed for a certain duration, unless the barred cell is an inaccessible Closed Subscriber Group (CSG) cell. If, however, the best cell is unsuitable for some other specific reason (e.g. because it belongs to a forbidden tracking area or to another non-equivalent PLMN), the UE is not permitted to consider any cell on the concerned frequency as a cell reselection candidate for a maximum of 300 s.

3.3.4.5 Speed Dependent Scaling

The UE scales the cell reselection parameters depending on its speed. This applies both in idle mode ($T_{reselection}$ and Q_{hyst}) and in connected mode (timeToTrigger). The UE speed is categorized by a mobility state (high, normal or low), which the UE determines based on the number of cell reselections/handovers which occur within a defined period, excluding consecutive reselections/handovers between the same two cells. The state is determined by comparing the count with thresholds for medium and high state, while applying some hysteresis. For idle and connected modes, separate sets of control parameters are used, signalled in SIB3 and within the measurement configuration respectively.

3.3.4.6 Cell Access Restrictions

The UE performs an access barring check during connection establishment (see Section 3.2.3.2). This function provides a means to control the load introduced by UE-originated traffic. There are separate means for controlling Mobile Originated (MO) calls and MO signalling. On top of the regular access class barring, Service Specific Access Control (SSAC) may be applied. SSAC facilitates separate control for MultiMedia TELephony (MMTEL) voice and video calls. Most of the SSAC functionality is handled by upper layers. In addition, separate access control exists to protect against E-UTRAN overload due to UEs accessing E-UTRAN merely to perform CSFB to CDMA2000.

Each UE belongs to an Access Class (AC) in the range 0–9. In addition, some UEs may belong to one or more high-priority ACs in the range 11–15, which are reserved for specific uses (e.g. security services, public utilities, emergency services, PLMN staff). AC10 is used for emergency access. Further details, for example regarding in which PLMN the high priority ACs apply, are provided in [5]. The UE considers access to be barred if access is barred for all its applicable ACs.

Samsung Ex. 1008
97 of 126

SIB2 may include a set of AC barring parameters for MO calls and/or MO signalling. This set of parameters comprises a probability factor and a barring timer for AC0–9 and a list of barring bits for AC11–15. For AC0–9, if the UE initiates a MO call and the relevant AC barring parameters are included, the UE draws a random number. If this number exceeds the probability factor, access is not barred. Otherwise access is barred for a duration which is randomly selected centred on the broadcast barring timer value. For AC11–15, if the UE initiates a MO call and the relevant AC barring parameters are included, access is barred whenever the bit corresponding to all of the UE's ACs is set. The behaviour is similar in the case of UE-initiated MO signalling.

For cell (re)selection, the UE is expected to consider cells which are neither barred nor reserved for operator or future use. In addition, a UE with an access class in the range 11–15 shall consider a cell that is (only) reserved for operator use and part of its home PLMN (or an equivalent) as a candidate for cell reselection. The UE is never allowed to (re)select a cell that is not a reselection candidate even for emergency access.

3.3.4.7 Any Cell Selection

When the UE is unable to find a suitable cell of the selected PLMN, it performs 'any cell selection'. In this case, the UE performs normal idle mode operation: monitoring paging, acquiring SI, performing cell reselection. In addition, the UE regularly attempts to find a suitable cell on other frequencies or RATs (i.e. not listed in SI). If a UE supporting voice services is unable to find a suitable cell, it should attempt to find an acceptable cell on any supported RAT regardless of the cell reselection priorities that are broadcast. The UE is not allowed to receive MBMS in this state.

3.3.4.8 Closed Subscriber Group

LTE supports the existence of cells which are accessible only for a limited set of UEs – a Closed Subscriber Group (CSG). In order to prevent UEs from attempting to register on a CSG cell on which they do not have access, the UE maintains a CSG white list, i.e. a list of CSG identities for which access has been granted to the UE. The CSG white list can be transferred to the UE by upper layers, or updated upon successful access to a CSG cell. To facilitate the latter, UEs support 'manual selection' of CSG cells which are not in the CSG white list. The manual selection may be requested by the upper layers, based on a text string broadcast by the cell. LTE also supports hybrid cells. Like CSG cells, hybrid cells broadcast a CSG identity; they are accessible as CSG cells by UEs whose CSG white lists include the CSG identity, and as normal cells by all other UEs (see Section 24.2.2).

3.4 Paging

To receive paging messages from E-UTRAN, UEs in idle mode monitor the PDCCH channel for an RNTI value used to indicate paging: the P-RNTI (see Section 9.2.2.2). The UE only needs to monitor the PDCCH channel at certain UE-specific occasions (i.e. at specific subframes within specific radio frames – see Section 6.2 for an introduction to the LTE radio frame structure.). At other times, the UE may apply DRX, meaning that it can switch off its receiver to preserve battery power.

Samsung Ex. 1008
98 of 126

The E-UTRAN configures which of the radio frames and subframes are used for paging. Each cell broadcasts a default paging cycle. In addition, upper layers may use dedicated signalling to configure a UE-specific paging cycle. If both are configured, the UE applies the lowest value. The UE calculates the radio frame (the Paging Frame (PF)) and the subframe within that PF (the Paging Occasion (PO)), which E-UTRAN applies to page the UE as follows:

$$\begin{aligned} \text{SFN mod } T &= (T/N) \times (\text{UE_ID mod } N) \\ \text{i_s} &= \lfloor \text{UE_ID}/N \rfloor \text{ mod Ns} \\ T &= \text{UE DRX cycle (i.e. paging cycle)} = \min(T_c, T_{ue}) \\ N &= \min(T, \text{nB}) \\ \text{Ns} &= \max(1, \text{nB}/T) \end{aligned} \tag{3.1}$$

where:

T_c is the cell-specific default paging cycle {32, 64, 128, 256} radio frames,

T_{ue} is the UE-specific paging cycle {32, 64, 128, 256} radio frames,

N is the number of paging frames within the paging cycle of the UE,

UE_ID is the IMSI[20] mod 1024, with IMSI being the decimal rather than the binary number,

i_s is an index pointing to a pre-defined table defining the corresponding subframe,

nB is the number of 'paging subframes' per paging cycle (across all UEs in the cell),

Ns is the number of 'paging subframes' in a radio frame that is used for paging.

Table 3.4 includes a number of examples to illustrate the calculation of the paging radio frames (PF) and subframes (PO).

Table 3.4: Examples for calculation of paging frames and subframes.

Case	UE_ID	T_c	T_{ue}	T	nB	N	Ns	PF	i_s	PO
A	147	256	256	256	64	64	1	76	0	9
B	147	256	128	128	32	32	1	76	0	9
C	147	256	128	128	256	128	2	19	1	4

In cases A and B in Table 3.4, one out of every four radio frames is used for paging, using one subframe in each of those radio frames. For case B, there are 32 paging frames within the UE's paging cycle, across which the UEs are distributed based on the UE-identity. In case C, two subframes in each radio frame are used for paging, i.e. Ns = 2. In this case, there are 128 paging frames within the UE's paging cycle and the UEs are also distributed across the two subframes within the paging frame. The LTE specifications include a table that indicates the subframe applicable for each combination of Ns and i_s, which is the index that follows from Equation (3.1). Figure 3.14 illustrates cases B and C. All the shaded subframes can be used for paging; the darker ones are applicable for the UE with the indicated identity.

[20]International Mobile Subscriber Identity.

Samsung Ex. 1008
99 of 126

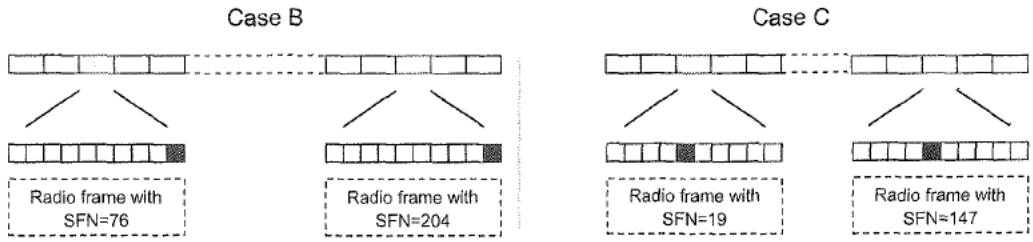


Figure 3.14: Paging frame and paging occasion examples.

3.5 Summary

The main aspects of the Control Plane protocols in LTE can be broken down into the Cell Selection and Reselection Procedures when the UE is in Idle Mode, and the RRC protocol when the UE is in Connected Mode.

The roles of these protocols include supporting security, mobility both between different LTE cells and between LTE and other radio systems, and establishment and reconfiguration of the radio bearers which carry control information and user data.

References[21]

[1] 3GPP Technical Specification 36.331, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification', www.3gpp.org.

[2] 3GPP Technical Specification 36.304, 'Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) procedures in idle mode (Release 9)', www.3gpp.org.

[3] 3GPP Technical Specification 33.401, '3GPP System Architecture Evolution; Security Architecture', www.3gpp.org.

[4] 3GPP Technical Specification 36.133, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Requirements for support of radio resource management', www.3gpp.org.

[5] 3GPP Technical Specification 22.011, 'Service accessibility', www.3gpp.org.

[21] All web sites confirmed 1st March 2011.

Samsung Ex. 1008
100 of 126

9

Downlink Physical Data and Control Channels

Matthew Baker and Tim Moulsley

9.1 Introduction

Chapters 7 and 8 have described the signals which enable User Equipment (UEs) to synchronize with the network and estimate the downlink radio channel in order to be able to demodulate data. This chapter first reviews the downlink physical channels which transport the data and then explains the control-signalling channels; the latter support the data channels by indicating the particular time-frequency transmission resources to which the data is mapped and the format in which the data itself is transmitted.

9.2 Downlink Data-Transporting Channels

9.2.1 Physical Broadcast Channel (PBCH)

In cellular systems, the basic System Information (SI) which allows the other channels in the cell to be configured and operated is usually carried by a Broadcast CHannel (BCH). Therefore the achievable coverage for reception of the BCH is crucial to the successful operation of such cellular communication systems; LTE is no exception. As already noted in Chapter 3, the broadcast (SI) is divided into two categories:

LTE – The UMTS Long Term Evolution: From Theory to Practice, Second Edition.
Stefania Sesia, Issam Toufik and Matthew Baker.
© 2011 John Wiley & Sons, Ltd. Published 2011 by John Wiley & Sons, Ltd.

- The 'Master Information Block' (MIB), which consists of a limited number of the most frequently transmitted parameters essential for initial access to the cell,[1] and is carried on the Physical Broadcast CHannel (PBCH).
- The other System Information Blocks (SIBs) which, at the physical layer, are multiplexed with unicast data transmitted on the Physical Downlink Shared CHannel (PDSCH) as discussed in Section 9.2.2.2.

This section focuses in particular on the PBCH, the design of which reflects some specific requirements:

- Detectability without prior knowledge of the system bandwidth;
- Low system overhead;
- Reliable reception right to the edge of the LTE cells;
- Decodability with low latency and low impact on UE battery life.

The resulting overall PBCH structure is shown in Figure 9.1.

Detectability without the UE having prior knowledge of the system bandwidth is achieved by mapping the PBCH only to the central 72 subcarriers of the OFDM[2] signal (which corresponds to the minimum possible LTE system bandwidth of 6 Resource Blocks (RBs)), regardless of the actual system bandwidth. The UE will have first identified the system centre-frequency from the synchronization signals as described in Chapter 7.

Low system overhead for the PBCH is achieved by deliberately keeping the amount of information carried on the PBCH to a minimum, since achieving stringent coverage requirements for a large quantity of data would result in a high system overhead. The size of the MIB is therefore just 14 bits, and, since it is repeated every 40 ms, this corresponds to a data rate on the PBCH of just 350 bps.

The main mechanisms employed to facilitate reliable reception of the PBCH in LTE are time diversity, Forward Error Correction (FEC) coding and antenna diversity.

Time diversity is exploited by spreading out the transmission of each MIB on the PBCH over a period of 40 ms. This significantly reduces the likelihood of a whole MIB being lost in a fade in the radio propagation channel, even when the mobile terminal is moving at pedestrian speeds.

The FEC coding for the PBCH uses a convolutional coder, as the number of information bits to be coded is small; the details of the convolutional coder are explained in Section 10.3.3. The basic code rate is 1/3, after which a high degree of repetition of the systematic (i.e. information) bits and parity bits is used, such that each MIB is coded at a very low code rate (1/48 over a 40 ms period) to give strong error protection.

Antenna diversity may be utilized at both the eNodeB and the UE. The UE performance requirements specified for LTE assume that all UEs can achieve a level of decoding performance commensurate with dual-antenna receive diversity (although it is recognized that in low-frequency deployments, such as below 1 GHz, the advantage obtained from receive antenna diversity is reduced due to the correspondingly higher correlation between

[1]The MIB information consists of the downlink system bandwidth, the PHICH size (Physical Hybrid ARQ Indicator CHannel, see Section 9.3.4), and the most-significant eight bits of the System Frame Number (SFN) – the remaining two bits being gleaned from the 40 ms periodicity of the PBCH.

[2]Orthogonal Frequency Division Multiplexing.

Samsung Ex. 1008
102 of 126

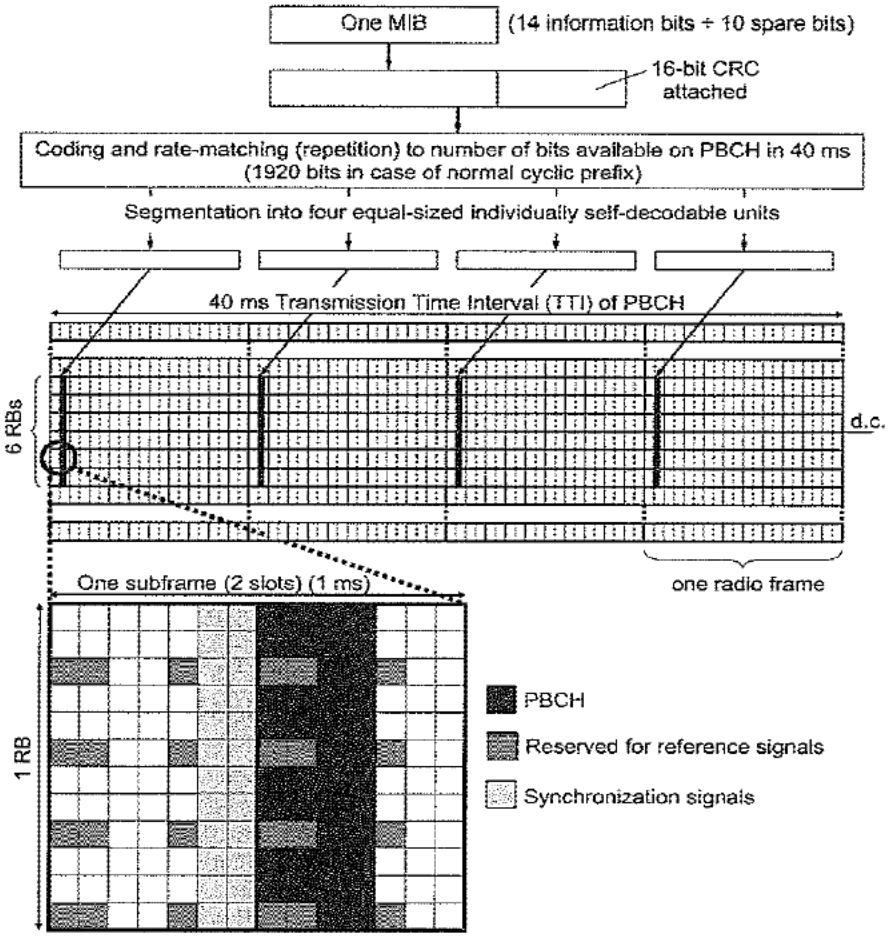


Figure 9.1: PBCH structure.

the antennas); this enables LTE system planners to rely on this level of performance being common to all UEs, thereby enabling wider cell coverage to be achieved with fewer cell sites than would otherwise be possible. Transmit antenna diversity may be also employed at the eNodeB to further improve coverage, depending on the capability of the eNodeB; eNodeBs with two or four transmit antenna ports transmit the PBCH using a Space-Frequency Block Code (SFBC), details of which are explained in Section 11.2.2.1.

The precise set of Resource Elements (REs) used by the PBCH is independent of the number of transmit antenna ports used by the eNodeB; any REs which may be used for Reference Signal (RS) transmission are avoided by the PBCH, irrespective of the actual number of transmit antenna ports deployed at the eNodeB. The number of transmit antenna ports used by the eNodeB must be ascertained blindly by the UE, by performing the decoding

for each SFBC scheme corresponding to the different possible numbers of transmit antenna ports (namely one, two or four). This discovery of the number of transmit antenna ports is further facilitated by the fact that the Cyclic Redundancy Check (CRC) on each MIB is masked with a codeword representing the number of transmit antenna ports.

Finally, achieving low latency and a low impact on UE battery life is also facilitated by the design of the coding outlined above: the low code rate with repetition enables the full set of coded bits to be divided into four subsets, each of which is self-decodable in its own right. Each of these subsets of the coded bits is then transmitted in a different one of the four radio frames during the 40 ms transmission period, as shown in Figure 9.1. This means that if the Signal to Interference Ratio (SIR) of the radio channel is sufficiently good to allow the UE to decode the MIB correctly from the transmission in less than four radio frames, then the UE does not need to receive the other parts of the PBCH transmission in the remainder of the 40 ms period; on the other hand, if the SIR is low, the UE can receive further parts of the MIB transmission, soft-combining each part with those received already, until successful decoding is achieved.

The timing of the 40 ms transmission interval for each MIB on the PBCH is not indicated explicitly to the UE; it is ascertained implicitly from the scrambling and bit positions, which are re-initialized every 40 ms. The UE can therefore initially determine the 40 ms timing by performing four separate decodings of the PBCH using each of the four possible phases of the PBCH scrambling code, checking the CRC for each decoding.

When a UE initially attempts to access a cell by reading the PBCH, a variety of approaches may be taken to carry out the necessary blind decodings. A simple approach is always to perform the decoding using a soft combination of the PBCH over four radio frames, advancing a 40 ms sliding window one radio frame at a time until the window aligns with the 40 ms period of the PBCH and the decoding succeeds. However, this would result in a 40–70 ms delay before the PBCH can be decoded. A faster approach would be to attempt to decode the PBCH from the first single radio frame, which should be possible provided the SIR is sufficiently high; if the decoding fails for all four possible scrambling code phases, the PBCH from the first frame could be soft-combined with the PBCH bits received in the next frame – there is a 3-in-4 chance that the two frames contain data from the same transport block. If decoding still fails, a third radio frame could be combined, and failing that a fourth. It is evident that the latter approach may be much faster (potentially taking only 10 ms), but on the other hand requires slightly more complex logic.

9.2.2 Physical Downlink Shared CHannel (PDSCH)

The Physical Downlink Shared CHannel (PDSCH) is the main data-bearing downlink channel in LTE. It is used for all user data, as well as for broadcast system information which is not carried on the PBCH, and for paging messages – there is no specific physical layer paging channel in LTE. The use of the PDSCH for user data is explained in Section 9.2.2.1; the use of the PDSCH for system information and paging is covered in Section 9.2.2.2.

Data is transmitted on the PDSCH in units known as *Transport Blocks* (TBs), each of which corresponds to a Medium Access Control (MAC) layer Protocol Data Unit (PDU) as described in Section 4.4. Transport blocks may be passed down from the MAC layer to the physical layer once per Transmission Time Interval (TTI), where a TTI is 1 ms, corresponding to the subframe duration.

Samsung Ex. 1008
104 of 126

9.2.2.1 General Use of the PDSCH

When employed for user data, one or, at most, two TBs can be transmitted per UE per subframe, depending on the transmission mode selected for the PDSCH for each UE. The transmission mode configures the multi-antenna transmission scheme usually applied:[3]

Transmission Mode 1: Transmission from a single eNodeB antenna port;

Transmission Mode 2: Transmit diversity (see Section 11.2.2.1);

Transmission Mode 3: Open-loop spatial multiplexing (see Section 11.2.2.2);

Transmission Mode 4: Closed-loop spatial multiplexing (see Section 11.2.2.2);

Transmission Mode 5: Multi-User Multiple-Input Multiple-Output (MU-MIMO) (see Section 11.2.3);

Transmission Mode 6: Closed-loop rank-1 precoding (see Section 11.2.2.2);

Transmission Mode 7: Transmission using UE-specific RSs with a single spatial layer (see Sections 8.2 and 11.2.2.3);

Transmission Mode 8: Introduced in Release 9, transmission using UE-specific RSs with up to two spatial layers (see Sections 8.2.3 and 11.2.2.3);

Transmission Mode 9: Introduced in Release 10, transmission using UE-specific RSs with up to eight spatial layers (see Sections 29.1 and 29.3).

With the exception of transmission modes 7, 8 and 9, the phase reference for demodulating the PDSCH is given by the cell-specific Reference Signals (RSs) described in Section 8.2.1, and the number of eNodeB antenna ports used for transmission of the PDSCH is the same as the number of antenna ports used in the cell for the PBCH. In transmission modes 7, 8 and 9, UE-specific RSs (see Sections 8.2.2, 8.2.3 and 29.1.1 respectively) provide the phase reference for the PDSCH. The configured transmission mode also controls the formats of the associated downlink control signalling messages, as described in Section 9.3.5.1, and the modes of channel quality feedback from the UE (see Section 10.2.1).

After channel coding (see Section 10.3.2) and mapping to spatial layers according to the selected transmission mode, the coded PDSCH data bits are mapped to modulation symbols depending on the modulation scheme selected for the current radio channel conditions and required data rate.

The modulation order may be selected between two bits per symbol (using QPSK (Quadrature Phase Shift Keying)), four bits per symbol (using 16QAM (Quadrature Amplitude Modulation)) and six bits per symbol (using 64QAM); constellation diagrams for these modulation schemes are illustrated in Figure 9.2. Support for reception of 64QAM modulation is mandatory for all classes of LTE UE.

The REs used for the PDSCH can be any which are not reserved for other purposes (i.e. RSs, synchronization signals, PBCH and control signalling). Thus when the control

[3] In addition to the transmission schemes listed here for each mode, transmission modes 3 to 9 also support the use of transmit diversity as a 'fallback' technique; this is useful, for example, when radio conditions are temporarily inappropriate for the usual scheme, or to ensure that a common scheme is available during reconfiguration of the transmission mode.

Samsung Ex. 1008
105 of 126

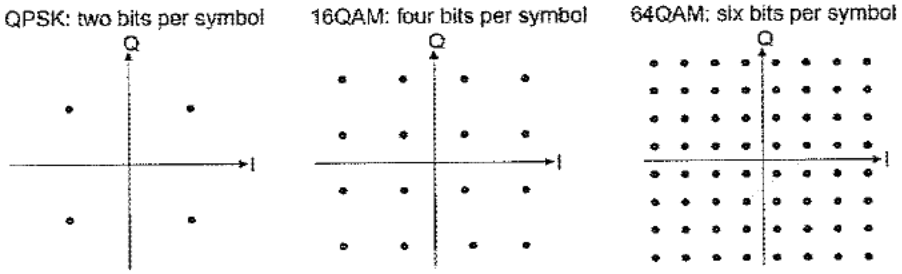


Figure 9.2: Constellations of modulation schemes applicable to PDSCH transmission.

signalling informs a UE that a particular pair of RBs[4] in a subframe are allocated to that UE, it is only the *available* REs within those RBs which actually carry PDSCH data.

Normally the allocation of pairs of RBs to PDSCH transmission for a particular UE is signalled to the UE by means of dynamic control signalling transmitted at the start of the relevant subframe using the Physical Downlink Control Channel (PDCCH), as described in Section 9.3.

The mapping of data to physical RBs can be carried out in one of two ways: *localized mapping* and *distributed mapping*.[5]

Localized resource mapping entails allocating all the available REs in a pair of RBs to the same UE. This is suitable for most scenarios, including the use of dynamic channel-dependent scheduling according to frequency-specific channel quality information reported by the UE (see Sections 10.2.1 and 12.4).

Distributed resource mapping entails separating in frequency the two physical RBs comprising each pair, with a frequency-hop occurring at the slot boundary in the middle of the subframe, as shown in Figure 9.3. This is a useful means of obtaining frequency diversity for small amounts of data which would otherwise be constrained to a narrow part of the downlink bandwidth and would therefore be more susceptible to narrow-band fading. An example of a typical use for this transmission mode could be a Voice-over-IP (VoIP) service, where, in order to minimize overhead, certain frequency resources may be 'semi-persistently scheduled' (see Section 4.4.2.1) – in other words, certain RBs in the frequency domain are allocated on a periodic basis to a specific UE by Radio Resource Control (RRC) signalling rather than by dynamic PDCCH signalling. This means that the transmissions are not able to benefit from dynamic channel-dependent scheduling, and therefore the frequency diversity which is achieved through distributed mapping is a useful tool to improve performance. Moreover, as the amount of data to be transmitted per UE for a VoIP service is small (typically sufficient to occupy only one or two pairs of RBs in a given subframe), the degree of frequency diversity obtainable via localized scheduling is very limited.

[4]The term 'pair of RBs' here means a pair of resource blocks which occupy the same set of 12 subcarriers and are contiguous in time, thus having a duration of one subframe.

[5]Distributed mapping is not supported in conjunction with UE-specific RSs in transmission modes 8 and 9.

Samsung Ex. 1008
106 of 126

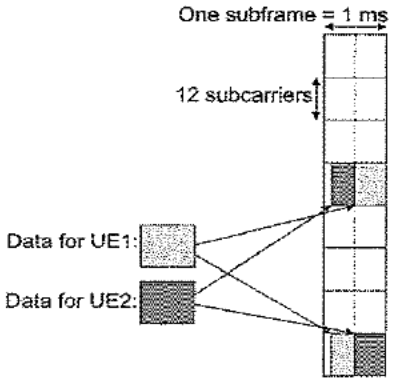
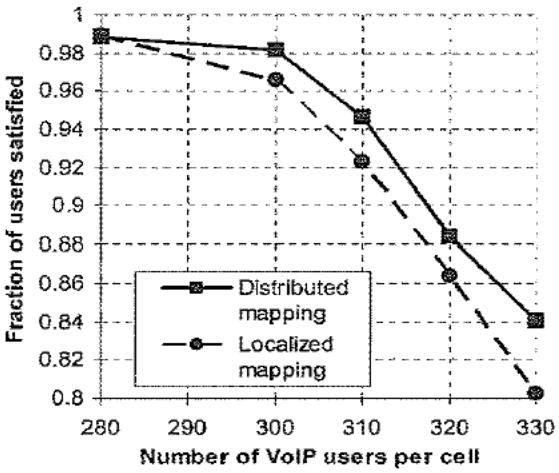


Figure 9.3: Frequency-distributed data mapping in LTE downlink.

The potential increase in the number of VoIP users which can be accommodated in a cell as a result of using distributed resource mapping as opposed to localized resource mapping is illustrated by way of example in Figure 9.4.



Parameter	Value
Carrier frequency	2 GHz
Bandwidth	5 MHz
Channel model	Urban micro, 3 km/h
eNB transmit power	43 dBm
VoIP codec	12.2 kbps
Voice activity factor	0.43
Modulation scheme	QPSK
Code rate	2/3
Satisfaction criterion	98% packets within 50 ms

Figure 9.4: Example of increase in VoIP capacity arising from frequency-distributed resource mapping.

9.2.2.2 Special Uses of the PDSCH

As noted above, the PDSCH is used for some special purposes in addition to normal user data transmission.

One such use is for the broadcast system information (i.e. SIBs) that is not carried on the PBCH. The RBs used for broadcast data of this sort are indicated by signalling messages on the PDCCH in the same way as for other PDSCH data, except that the identity indicated on the PDCCH is not the identity of a specific UE but is, rather, a designated broadcast identity known as the System Information Radio Network Temporary Identifier (SI-RNTI), which is fixed in the specifications (see Section 7.1 of [1]) and therefore known a priori to all UEs. Some constraints exist as to which subframes may be used for SI messages on the PDSCH; these are explained in Section 3.2.2.

Another special use of the PDSCH is paging, as no separate physical channel is provided in LTE for this purpose. In previous systems such as WCDMA,[6] a special 'Paging Indicator Channel' was provided, which was specially designed to enable the UE to wake up its receiver periodically for a very short period of time, in order to minimize the impact on battery life; on detecting a paging indicator (typically for a group of UEs), the UE would then keep its receiver switched on to receive a longer message indicating the exact identity of the UE being paged. By contrast, in LTE the PDCCH signalling is already very short in duration, and therefore the impact on UE battery life of monitoring the PDCCH from time to time is low. Therefore the normal PDCCH signalling can be used to carry the equivalent of a paging indicator, with the detailed paging information being carried on the PDSCH in RBs indicated by the PDCCH. In a similar way to broadcast data, paging indicators on the PDCCH use a single fixed identifier called the Paging RNTI (P-RNTI). Rather than providing different paging identifiers for different groups of UEs, different UEs monitor different subframes for their paging messages, as described in Section 3.4. Paging messages may be received in subframes 0, 4, 5 or 9 in each radio frame.

9.2.3 Physical Multicast Channel (PMCH)

In LTE Release 9, the use of a third data-transporting channel became available, namely the Physical Multicast CHannel (PMCH), designed to carry data for Multimedia Broadcast and Multicast Services (MBMS). The PMCH can only be transmitted in certain specific subframes known as MBSFN (Multimedia Broadcast Single Frequency Network) subframes, indicated in the system information carried on the PDSCH. A Release 8 UE must be aware of the possible existence of MBSFN subframes, but is not required to decode the PMCH. The details of the PMCH are explained in Section 13.4.1.

9.3 Downlink Control Channels

9.3.1 Requirements for Control Channel Design

The control channels in LTE are provided to support efficient data transmission. In common with other wireless systems, the control channels convey physical layer messages which cannot be carried sufficiently efficiently, quickly or conveniently by higher layer protocols. The design of the control channels in the LTE downlink aims to balance a number of somewhat conflicting requirements, the most important of which are discussed below.

[6] Wideband Code Division Multiple Access.

Samsung Ex. 1008
108 of 126

9.3.1.1 Physical Layer Signalling to Support the MAC Layer

The general requirement to support MAC operation is very similar to that in WCDMA, but there are a number of differences of detail, mainly arising from the frequency domain resource allocation supported in the LTE multiple access schemes.

The use of the uplink transmission resources on the Physical Uplink Shared Channel (PUSCH) is determined dynamically by an uplink scheduling process in the eNodeB, and therefore physical layer signalling must be provided to indicate to UEs which time/frequency resources they have been granted permission to use.

The eNodeB also schedules downlink transmissions on the PDSCH, and therefore similar physical layer messages from the eNodeB are needed to indicate which resources in the frequency domain contain the downlink data transmissions intended for particular UEs, together with parameters such as the modulation scheme and code rate used for the data. Explicit signalling of this kind avoids the considerable additional complexity which would arise if UEs had to search for their data among all the possible combinations of data packet size, format and resource allocation.

In order to facilitate efficient operation of HARQ,[7] further physical layer signals are needed to convey acknowledgements of uplink data packets received by the eNodeB, and power control commands are needed to ensure that uplink transmissions are made at appropriate power levels (as explained in Section 18.3).

9.3.1.2 Flexibility, Overhead and Complexity

The LTE physical layer specification is intended to allow operation in any system bandwidth from six resource blocks (1.08 MHz) to 110 resource blocks (19.8 MHz). It is also designed to support a range of scenarios including, for example, just a few users in a cell each demanding high data rates, or very many users with low data rates. Considering the possibility that both uplink resource grants and downlink resource allocations could be required for every UE in each subframe, the number of control channel messages carrying resource information could be as many as a couple of hundred if every resource allocation were as small as one resource block. Since every additional control channel message implies additional overhead which consumes downlink resources, it is desirable that the control channel is designed to minimize unnecessary overhead for any given signalling load, whatever the system bandwidth.

Similar considerations apply to the signalling of HARQ acknowledgements for each uplink packet transmission.

Furthermore, as in any mobile communication system, the complexity and power consumption of the terminals are important considerations for LTE. Therefore, the control signalling must be designed so that the necessary scalability and flexibility is achieved without undue decoding complexity.

9.3.1.3 Coverage and Robustness

In order to achieve good coverage it must be possible to configure the system so that the control channels can be received with sufficient reliability over a substantial part of every cell. As an example, if a message indicating resource allocation is not received correctly, then the corresponding data transmission will also fail, with a direct and proportionate impact

[7] Hybrid Automatic Repeat reQuest – see Sections 4.4 and 10.3.2.5.

Samsung Ex. 1008
109 of 126

on throughput efficiency. Techniques such as channel coding and frequency diversity can be used to make the control channels more robust. However, in order to make good use of system resources, it is desirable to be able to adapt the transmission parameters of the control signalling for different UEs or groups of UEs, so that lower code rates and higher power levels are only applied for those UEs for which it is necessary (e.g. near the cell border, where signal levels are likely to be low and interference from other cells high).

Also, it is desirable to avoid unintended reception of control channels from other cells, by applying cell-specific randomization.

9.3.1.4 System-Related Design Aspects

Since the different parts of LTE are intended to provide a complete system, some aspects of control channel design cannot be considered in isolation.

A basic design decision in LTE is that a control channel message is intended to be transmitted to a particular UE (or, in some cases, a group of UEs). Therefore, in order to reach multiple UEs in a cell within a subframe, it must be possible to transmit multiple control channels within the duration of a single subframe. However, in cases where the control channel messages are intended for reception by more than one UE (for example, when relating to the transmission of a SIB on the PDSCH), it is more efficient to arrange for all the UEs to receive a single transmission rather than to transmit the same information to each UE individually. This requires that both common and dedicated control channel messages be supported.

Finally, some scenarios may be characterized by the data arriving at regular intervals, as is typical for VoIP traffic. It is then possible to predict in advance when resources will need to be assigned in the downlink or granted in the uplink, and the number of control channel messages which need to be sent can be reduced by means of 'Semi-Persistent Scheduling' (SPS) as discussed in Section 4.4.2.1.

9.3.2 Control Channel Structure

Three downlink physical control channels are provided in LTE: the Physical Control Format Indicator CHannel (PCFICH), the Physical HARQ Indicator CHannel (PHICH) and the Physical Downlink Control CHannel (PDCCH). In general, the downlink control channels can be configured to occupy the first 1, 2 or 3 OFDM symbols in a subframe, extending over the entire system bandwidth as shown in Figure 9.5.

This flexibility allows the control channel overhead to be adjusted according to the particular system configuration, traffic scenario and channel conditions. There are two special cases: in subframes containing MBSFN transmissions (see Sections 9.2.3 and 13.4.1) there may be 0, 1 or 2 OFDM symbols for control signalling, while for narrow system bandwidths (less than 10 RBs) the number of control symbols is increased and may be 2, 3 or 4 to ensure sufficient coverage at the cell border.

9.3.3 Physical Control Format Indicator CHannel (PCFICH)

The PCFICH carries a Control Format Indicator (CFI) which indicates the number of OFDM symbols (i.e. normally 1, 2 or 3) used for transmission of control channel information in each subframe. In principle, the UE could deduce the value of the CFI without a channel such as

Samsung Ex. 1008
110 of 126

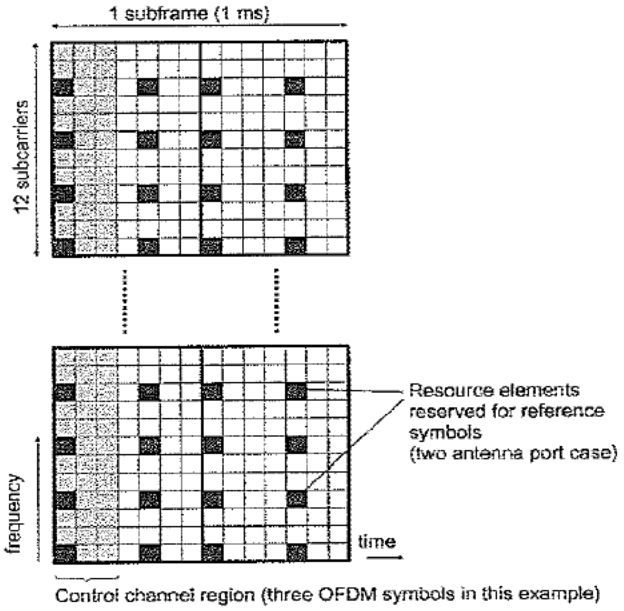


Figure 9.5: The time-frequency region used for downlink control signalling.

the PCFICH, for example by multiple attempts to decode the control channels assuming each possible number of symbols, but this would result in significant additional processing load. Three different CFI values are used in LTE. In order to make the CFI sufficiently robust, each codeword is 32 bits in length, mapped to 16 REs using QPSK modulation. These 16 REs are arranged in groups of 4, known as Resource Element Groups (REGs). The REs occupied by RSs are not included within the REGs, which means that the total number of REGs in a given OFDM symbol depends on whether or not cell-specific RSs are present. The concept of REGs (i.e. mapping in groups of four REs) is also used for the other downlink control channels (the PHICH and PDCCH).

The PCFICH is transmitted on the same set of antenna ports as the PBCH, with transmit diversity being applied if more than one antenna port is used.

In order to achieve frequency diversity, the 4 REGs carrying the PCFICH are distributed across the frequency domain. This is done according to a predefined pattern in the first OFDM symbol in each downlink subframe (see Figure 9.6), so that the UEs can always locate the PCFICH information, which is a prerequisite to being able to decode the rest of the control signalling.

To minimize the possibility of confusion with PCFICH information from a neighbouring cell, a cell-specific frequency offset is applied to the positions of the PCFICH REs; this offset depends on the Physical Cell ID (PCI), which is deduced from the Primary and Secondary Synchronization Signals (PSS and SSS) as explained in Section 7.2. In addition, a cell-specific scrambling sequence (again a function of the PCI) is applied to the CFI codewords, so that the UE can preferentially receive the PCFICH from the desired cell.

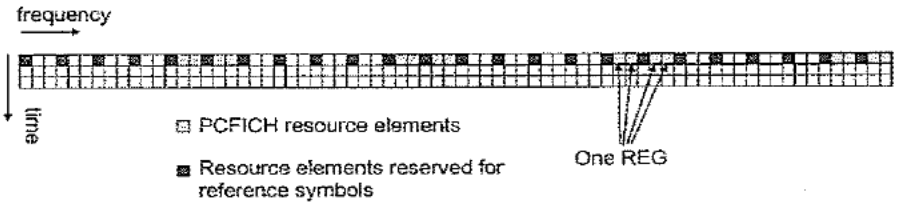


Figure 9.6: PCFICH mapping to Resource Element Groups (REGs).

9.3.4 Physical Hybrid ARQ Indicator Channel (PHICH)

The PHICH carries the HARQ ACK/NACK, which indicates whether the eNodeB has correctly received a transmission on the PUSCH. The HARQ indicator is set to 0 for a positive ACKnowledgement (ACK) and 1 for a Negative ACKnowledgement (NACK). This information is repeated in each of three BPSK[8] symbols.

Multiple PHICHs are mapped to the same set of REs. These constitute a PHICH group, where different PHICHs within the same PHICH group are separated through different complex orthogonal Walsh sequences. Each PHICH is uniquely identified by a PHICH index, which indicates both the group and the sequence. The sequence length is four for the normal cyclic prefix (or two in the case of the extended cyclic prefix). As the sequences are complex, the number of PHICHs in a group (i.e. the number of UEs receiving their acknowledgements on the same set of downlink REs) can be up to twice the sequence length. A cell-specific scrambling sequence is applied.

Factor-3 repetition coding is applied for robustness, resulting in three instances of the orthogonal Walsh code being transmitted for each ACK or NACK. The error rate on the PHICH is intended to be of the order of 10^{-2} for ACKs and as low as 10^{-4} for NACKs. The resulting PHICH construction, including repetition and orthogonal spreading, is shown in Figure 9.7.

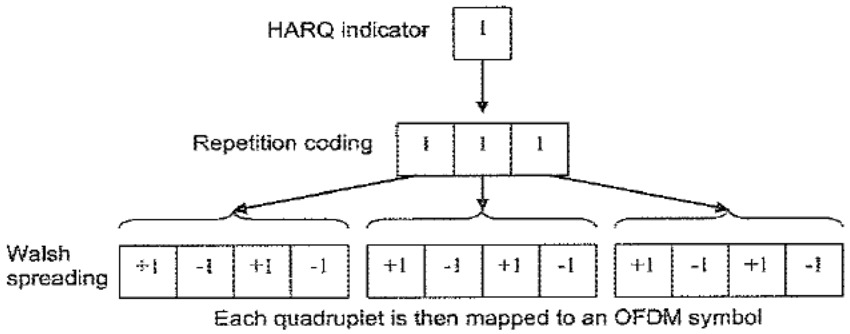


Figure 9.7: An example of PHICH signal construction.

[8]Binary Phase Shift Keying.

The PHICH duration, in terms of the number of OFDM symbols used in the time domain, is configurable (by an indication transmitted on the PBCH), normally to either one or three OFDM symbols.[9] As the PHICH cannot extend into the PDSCH transmission region, the duration configured for the PHICH puts a lower limit on the size of the control channel region at the start of each subframe (as signalled by the PCFICH).

Finally, each of the three instances of the orthogonal code of a PHICH transmission is mapped to an REG on one of the first three OFDM symbols of each subframe,[10] in such a way that each PHICH is partly transmitted on each of the available OFDM symbols. This mapping is illustrated in Figure 9.8 for each possible PHICH duration.

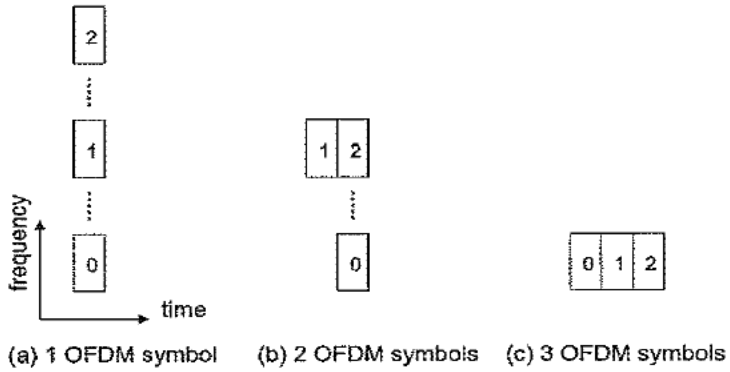


Figure 9.8: Examples of the mapping of the three instances of a PHICH orthogonal code to OFDM symbols, depending on the configured PHICH duration.

The PBCH also signals the number of PHICH groups configured in the cell, which enables the UEs to deduce to which remaining REs in the control region the PDCCHs are mapped.[11]

In order to obviate the need for additional signalling to indicate which PHICH carries the ACK/NACK response for each PUSCH transmission, the PHICH index is implicitly associated with the index of the lowest uplink RB used for the corresponding PUSCH transmission. This relationship is such that adjacent PUSCH RBs are associated with PHICHs in different PHICH groups, to enable some degree of load balancing. However, this mechanism alone is not sufficient to enable multiple UEs to be allocated the same RBs for a PUSCH transmission, as occurs in the case of uplink multi-user MIMO (see Section 16.6); in this case, different cyclic shifts of the uplink demodulation RSs are configured for the different UEs which are allocated the same time-frequency PUSCH resources, and the same

[9] In some special cases, the three-OFDM-symbol duration is reduced to two OFDM symbols; these cases are (i) MBSFN subframes on mixed carriers supporting MBSFN and unicast data, and (ii) the second and seventh subframes in case of frame structure type 2 for Time Division Duplex (TDD) operation.

[10] The mapping avoids REs used for reference symbols or PCFICH.

[11] For Frequency Division Duplex (FDD) operation with Frame Structure Type 1 (see Section 6.2), the configured number of PHICH groups is the same in all subframes; for TDD operation with Frame Structure Type 2, the number of PHICH groups is 0, 1 or 2 times the number signalled by the PBCH, according to the correspondence with uplink subframes.

cyclic shift index is then used to shift the PHICH allocations in the downlink so that each UE receives its ACK or NACK on a different PHICH. This mapping of the PHICH allocations is illustrated in Figure 9.9.

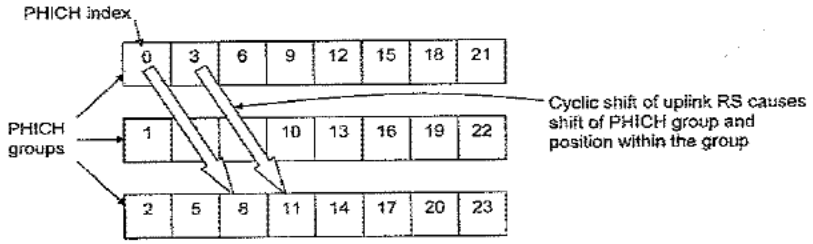


Figure 9.9: Indexing of PHICHs within PHICH groups, and shifting in the case of cyclic shifting of the uplink demodulation reference signals.

The PHICH indexing for the case of uplink MIMO in Release 10 is explained in Section 29.4.1. The use of the PHICH in the case of aggregation of multiple carriers in Release 10 is explained in Section 28.3.1.3.

The PHICHs are transmitted on the same set of antenna ports as the PBCH, and transmit diversity is applied if more than one antenna port is used.

9.3.5 Physical Downlink Control CHannel (PDCCH)

Each PDCCH carries a message known as Downlink Control Information (DCI), which includes resource assignments and other control information for a UE or group of UEs. In general, several PDCCHs can be transmitted in a subframe.

Each PDCCH is transmitted using one or more *Control Channel Elements* (CCEs), where each CCE corresponds to nine REGs. Four QPSK symbols are mapped to each REG.

Four PDCCH formats are supported, as listed in Table 9.1.

Table 9.1: PDCCH formats.

PDCCH format	Number of CCEs (n)	Number of REGs	Number of PDCCH bits
0	1	9	72
1	2	18	144
2	4	36	288
3	8	72	576

CCEs are numbered and used consecutively, and, to simplify the decoding process, a PDCCH with a format consisting of n CCEs may only start with a CCE with a number equal to a multiple of n.

The number of CCEs aggregated for transmission of a particular PDCCH is known as the 'aggregation level' and is determined by the eNodeB according to the channel conditions. For example, if the PDCCH is intended for a UE with a good downlink channel (e.g. close to the eNodeB), then one CCE is likely to be sufficient. However, for a UE with a poor channel (e.g. near the cell border) then eight CCEs may be required in order to achieve sufficient robustness. In addition, the power level of a PDCCH may be adjusted to match the channel conditions.

9.3.5.1 Formats for Downlink Control Information (DCI)

The required content of the control channel messages depends on the system deployment and UE configuration. For example, if the infrastructure does not support MIMO, or if a UE is configured in a transmission mode which does not involve MIMO, there is no need to signal the parameters that are only required for MIMO transmissions. In order to minimize the signalling overhead, it is therefore desirable that several different message formats are available, each containing the minimum payload required for a particular scenario. On the other hand, to avoid too much complexity in implementation and testing, it is desirable not to specify too many formats. The set of DCI message formats in Table 9.2 is specified in LTE; Format 2B was added in Release 9, and Formats 2C and 4 were added in Release 10. Additional formats may be defined in future.

Table 9.2: Supported DCI formats.

DCI format	Purpose	Applicable PDSCH transmission mode(s)
0	PUSCH grants	All
1	PDSCH assignments with a single codeword	1,2,7
1A	PDSCH assignments using a compact format	All
1B	PDSCH assignments for rank-1 transmission	6
1C	PDSCH assignments using a very compact format	n/a
1D	PDSCH assignments for multi-user MIMO	5
2	PDSCH assignments for closed-loop MIMO operation	4
2A	PDSCH assignments for open-loop MIMO operation	3
2B	PDSCH assignments for dual-layer beamforming	8
2C	PDSCH assignments for up to 8-layer spatial multiplexing	9
3	Transmit Power Control (TPC) commands for multiple users for PUCCH and PUSCH with 2-bit power adjustments	n/a
3A	Transmit Power Control (TPC) commands for multiple users for PUCCH and PUSCH with 1-bit power adjustments	n/a
4	PUSCH grants for up to 4-layer spatial multiplexing	All (if configured for PUSCH transmission mode 2)

The information content of the different DCI message formats is listed below for Frequency Division Duplex (FDD) operation. Some small differences exist for Time Division Duplex (TDD), and these are outlined afterwards.

Format 0. DCI Format 0 is used for the transmission of resource grants for the PUSCH. The following information is transmitted:

- Flag to differentiate between Format 0 and Format 1A;
- Resource assignment and frequency hopping flag;
- Modulation and Coding Scheme (MCS);
- New Data Indicator (NDI);
- HARQ information and Redundancy Version (RV);
- Power control command for scheduled PUSCH;
- Cyclic shift for uplink Demodulation RS;
- Request for transmission of an aperiodic CQI report (see Sections 10.2.1 and 28.3.2.3).

Format 1. DCI Format 1 is used for the transmission of resource assignments for single codeword PDSCH transmissions (transmission modes 1, 2 and 7 (see Section 9.2.2.1)). The following information is transmitted:

- Resource allocation type (see Section 9.3.5.4);
- RB assignment;
- MCS;
- HARQ information and RV;
- Power control command for Physical Uplink Control CHannel (PUCCH).

Format 1A. DCI Format 1A is used for compact signalling of resource assignments for single codeword PDSCH transmissions for any PDSCH transmission mode. It is also used to allocate a dedicated preamble signature to a UE to trigger contention-free random access (see Section 17.3.2); in this case the PDCCH message is known as a *PDCCH order*. The following information is transmitted:

- Flag to differentiate between Format 0 and Format 1A;
- Flag to indicate that the distributed mapping mode (see Section 9.2.2.1) is used for the PDSCH transmission (otherwise the allocation is a contiguous set of physical RBs);
- RB assignment;
- MCS;
- HARQ information and RV;
- Power control command for PUCCH.

Format 1B. DCI Format 1B is used for compact signalling of resource assignments for PDSCH transmissions using closed-loop precoding with rank-1 transmission (transmission mode 6). The information transmitted is the same as in Format 1A, but with the addition of an indicator of the precoding vector applied for the PDSCH transmission.

Samsung Ex. 1008
116 of 126

Format 1C. DCI Format 1C is used for very compact transmission of PDSCH assignments. When format 1C is used, the PDSCH transmission is constrained to using QPSK modulation. This is used, for example, for signalling paging messages and some broadcast system information messages (see Section 9.2.2.2), and for notifying UEs of a change of MBMS control information on the Multicast Control Channel (MCCH – see Section 13.6.3.2). The following information is transmitted:

- RB assignment;
- Coding scheme.

The RV is not signalled explicitly, but is deduced from the SFN (see [1, Section 5.3.1]).

Format 1D. DCI Format 1D is used for compact signalling of resource assignments for PDSCH transmissions using multi-user MIMO (transmission mode 5). The information transmitted is the same as in Format 1B, but, instead of one of the bits of the precoding vector indicators, there is a single bit to indicate whether a power offset is applied to the data symbols. This is needed to show whether the transmission power is shared between two UEs.

Format 2. DCI Format 2 is used for the transmission of resource assignments for PDSCH for closed-loop MIMO operation (transmission mode 4). The following information is transmitted:

- Resource allocation type (see Section 9.3.5.4);
- RB assignment;
- Power control command for PUCCH;
- HARQ information and RV for each codeword;
- MCS for each codeword;
- A flag to indicate if the mapping from transport blocks to codewords is reversed;
- Number of spatial layers;
- Precoding information and indication of whether one or two codewords are transmitted on the PDSCH.

Format 2A. DCI Format 2A is used for the transmission of resource assignments for PDSCH for open-loop MIMO operation (transmission mode 3). The information transmitted is the same as for Format 2, except that if the eNodeB has two transmit antenna ports, there is no precoding information, and, for four antenna ports, two bits are used to indicate the transmission rank.

Format 2B. DCI Format 2B is introduced in Release 9 and is used for the transmission of resource assignments for PDSCH for dual-layer beamforming (transmission mode 8). The information transmitted is similar to Format 2A, except that no precoding information is included and the bit in Format 2A for indicating reversal of the transport block to codeword mapping is replaced in Format 2B by a bit indicating the scrambling code applied to the UE-specific RSs for the corresponding PDSCH transmission (see Section 8.2.3).

Format 2C. DCI Format 2C is introduced in Release 10 and is used for the transmission of resource assignments for PDSCH for closed-loop single-user or multi-user MIMO operation with up to 8 layers (transmission mode 9). The information transmitted is similar to Format 2B; full details are given in Section 29.3.2.

Formats 3 and 3A. DCI Formats 3 and 3A are used for the transmission of power control commands for PUCCH and PUSCH, with 2-bit or 1-bit power adjustments respectively. These DCI formats contain individual power control commands for a group of UEs.

Format 4. DCI Format 4 is introduced in Release 10 and is used for the transmission of resource grants for the PUSCH when the UE is configured in PUSCH transmission mode 2 for uplink single-user MIMO. The information transmitted is similar to Format 0, with the addition of MCS and NDI information for a second transport block, and precoding information; full details are given in Section 29.4.

DCI Formats for TDD. In TDD operation, the DCI formats contain the same information as for FDD, but with some additions (see Section 23.4.3 for an explanation of the usage of these additions):

- Uplink index (in DCI Formats 0 and 4, uplink-downlink configuration 0 only);
- Downlink Assignment Index (DAI) (in DCI Formats 0, 1, 1A, 1B, 1D, 2, 2A 2B, 2C and 4, uplink-downlink configurations 1–6 only); see Section 23.4.3 for details of DAI usage.

DCI Format modifications in Release 10. In the case of aggregation of multiple carriers in Release 10, DCI Formats 0, 1, 1A, 1B, 1D, 2, 2A, 2B, 2C and 4 can be configured to include a carrier indicator for cross-carrier scheduling; this is explained in detail in Section 28.3.1.1. In DCI Formats 0 and 4, additional fields are included to request transmission of an aperiodic Sounding Reference Signal (SRS) (see Section 29.2.2) and to indicate whether the uplink PRB allocation is contiguous or multi-clustered (see Section 28.3.6.2 for details). In TDD operation, DCI Formats 2B and 2C may also be configured to include an additional field to request transmission of an aperiodic SRS.

9.3.5.2 PDCCH CRC Attachment

In order that the UE can identify whether it has received a PDCCH transmission correctly, error detection is provided by means of a 16-bit CRC appended to each PDCCH. Furthermore, it is necessary that the UE can identify which PDCCH(s) are intended for it. This could in theory be achieved by adding an identifier to the PDCCH payload; however, it turns out to be more efficient to scramble the CRC with the 'UE identity', which saves the additional payload but at the cost of a small increase in the probability of falsely detecting a PDCCH intended for another UE.

In addition, for UEs which support antenna selection for uplink transmissions (see Section 16.6), the requested antenna may be indicated using Format 0 by applying an antenna-specific mask to the CRC. This has the advantage that the same size of DCI message can be used, irrespective of whether antenna selection is used.

Samsung Ex. 1008
118 of 126

9.3.5.3 PDCCH Construction

In general, the number of bits required for resource assignment depends on the system bandwidth, and therefore the message sizes also vary with the system bandwidth. The numbers of payload bits for each DCI format (including information bits and CRC) are summarized in Table 9.3, for each of the supported values of system bandwidth. In addition, padding bits are added if necessary in the following cases:

- To ensure that Formats 0 and 1A are the same size, even in the case of different uplink and downlink bandwidths, in order to avoid additional complexity at the UE receiver;
- To ensure that Formats 3 and 3A are the same size as Formats 0 and 1A, likewise to avoid additional complexity at the UE receiver;
- To avoid potential ambiguity in identifying the correct PDCCH location as described in Section 9.3.5.5;
- To ensure that Format 1 has a different size from Formats 0/1A, so that these formats can be easily distinguished at the UE receiver;
- To ensure that Format 4 has a different size from Formats 1/2/2A/2B/2C, again so that this format can be easily distinguished.

In Release 10, some optional additional information bits may be configured in some of the DCI formats; these are not included in Table 9.3, but are explained below the table. Because their presence may affect the number of padding bits, any such additional bits do not necessarily increase the transmitted size of the DCI format by the same amount.

In order to provide robustness against transmission errors, the PDCCH information bits are coded as described in Section 10.3.3. The set of coded and rate-matched bits for each PDCCH are then scrambled with a cell-specific scrambling sequence; this reduces the possibility of confusion with PDCCH transmissions from neighbouring cells. The scrambled bits are mapped to blocks of four QPSK symbols (REGs). Interleaving is applied to these symbol blocks, to provide frequency diversity, followed by mapping to the available physical REs on the set of OFDM symbols indicated by the PCFICH. This mapping process excludes the REs reserved for RSs and the other control channels (PCFICH and PHICH).

The PDCCHs are transmitted on the same set of antenna ports as the PBCH, and transmit diversity is applied if more than one antenna port is used.

9.3.5.4 Resource Allocation

Conveying indications of physical layer resource allocation is one of the major functions of the PDCCHs. While the exact use of the PDCCHs depends on the algorithms implemented in the eNodeB, it is nevertheless possible to outline some general principles of typical operation.

In each subframe, PDCCHs indicate the frequency-domain resource allocations. As discussed in Section 9.2.2.1, resource allocations are normally localized, meaning that a Physical RB (PRB) in the first half of a subframe is paired with the PRB at the same frequency in the second half of the subframe. For simplicity, the explanation here is in terms of the first half subframe only.

The main design challenge for the signalling of frequency-domain resource allocations (in terms of a set of RBs) is to find a good compromise between flexibility and signalling overhead. The most flexible, and arguably the simplest, approach is to send each UE a

Table 9.3: DCI format payload sizes (in bits), without padding, for different FDD system bandwidths.

	Bandwidth (PRBs)					
	6	15	25	50	75	100
Format 0	35	37	39	41	42	43
Format 1	35	38	43	47	49	55
Format 1A	36	38	40	42	43	44
Format 1B/1D (2 transmit antenna ports)	38	40	42	44	45	46
Format 1C	24	26	28	29	30	31
Format 2 (2 transmit antenna ports)	47	50	55	59	61	67
Format 2A (2 transmit antenna ports)	44	47	52	56	58	64
Format 2B (2 or 4 transmit antenna ports)	44	47	52	56	58	64
Format 2C	46	49	54	58	60	66
Format 4 (2 UE transmit antennas)	46	47	50	52	53	54
Format 1B/1D (4 transmit antenna ports)	40	42	44	46	47	48
Format 2 (4 transmit antenna ports)	50	53	58	62	64	70
Format 2A (4 transmit antenna ports)	46	49	54	58	60	66
Format 4 (4 UE transmit antennas)	49	50	53	55	56	57

Note that for Release 10 UEs:

- DCI Format 0 is extended by 1 bit for a multi-cluster resource allocation flag and may be further extended by 1 or 2 bits for aperiodic CQI request (see Section 28.3.2.3) and aperiodic SRS request (see Section 29.2.2), depending on configuration.
- DCI Format 1A may be extended by 1 bit for aperiodic SRS request, depending on configuration.
- In TDD operation, DCI Formats 2B and 2C may be extended by 1 bit for aperiodic SRS request, depending on configuration.
- DCI Format 4 always includes 2 bits for requesting aperiodic SRS (see Section 29.2.2), and 1 bit for aperiodic CQI request, but may be extended by an additional 1 bit for aperiodic CQI request in case of carrier aggregation (see Section 28.3.2.3).
- DCI Formats 0, 1, 1A, 1B, 1D, 2, 2A, 2B, 2C and 4 can be configured to be extended by 3 bits for a carrier indicator field (see Section 28.3.1.1).

bitmap in which each bit indicates a particular PRB. This would work well for small system bandwidths, but for large system bandwidths (i.e. up to 110 PRBs) the bitmap would need 110 bits, which would be a prohibitive overhead – particularly for small packets, where the PDCCH message could be larger than the data packet! One possible solution would be to send a combined resource allocation message to all UEs, but this was rejected on the grounds of the high power needed to reach all UEs reliably, including those at the cell edges. The approaches adopted in LTE Releases 8 and 9 are listed in Table 9.4, and further details are given below.

Resource allocation Type 0. In resource allocations of Type 0, a bitmap indicates the Resource Block Groups (RBGs) which are allocated to the scheduled UE, where an RBG is a set of consecutive PRBs. The RBG size (P) is a function of the system bandwidth as shown in Table 9.5. The total number of RBGs (N_{RBG}) for a downlink system bandwidth of

Table 9.4: Methods for indicating Resource Block (RB) allocation.

Method	UL/DL	Description	Number of bits required (see text for definitions)
Direct bitmap	DL	The bitmap comprises 1 bit per RB. This method is the only one applicable when the bandwidth is less than 10 resource blocks.	N_{RB}^{DL}
Bitmap: 'Type 0'	DL	The bitmap addresses Resource Block Groups (RBGs), where the group size (2, 3 or 4) depends on the system bandwidth.	$\lceil N_{RB}^{DL}/P \rceil$
Bitmap: 'Type 1'	DL	The bitmap addresses individual RBs in a subset of RBGs. The number of subsets (2, 3, or 4) depends on the system bandwidth. The number of bits is arranged to be the same as for Type 0, so the same DCI format can carry either type of allocation.	$\lceil N_{RB}^{DL}/P \rceil$
Contiguous allocations: 'Type 2'	DL or UL	Any possible arrangement of contiguous RB allocations can be signalled in terms of a starting position and number of RBs.	$\lceil \log_2(N_{RB}^{DL}(N_{RB}^{DL}+1)) \rceil$ or $\lceil \log_2(N_{RB}^{UL}(N_{RB}^{UL}+1)) \rceil$
Distributed allocations	DL	In the downlink, a limited set of resource allocations can be signalled where the RBs are scattered across the frequency domain and shared between two UEs. The number of bits is arranged to be the same as for contiguous allocations Type 2, so the same DCI format can carry either type of allocation.	$\lceil \log_2(N_{RB}^{DL}(N_{RB}^{DL}+1)) \rceil$

N_{RB}^{DL} PRBs is given by $N_{RBG} = \lceil N_{RB}^{DL}/P \rceil$. An example for the case of $N_{RB}^{DL} = 25$, $N_{RBG} = 13$ and $P = 2$ is shown in Figure 9.10, where each bit in the bitmap indicates a pair of PRBs (i.e. two PRBs which are adjacent in frequency).

Table 9.5: RBG size for Type 0 resource allocation.

Downlink bandwidth N_{RB}^{DL}	RBG size (P)
$0 \le 10$	1
11–26	2
27–63	3
64–110	4

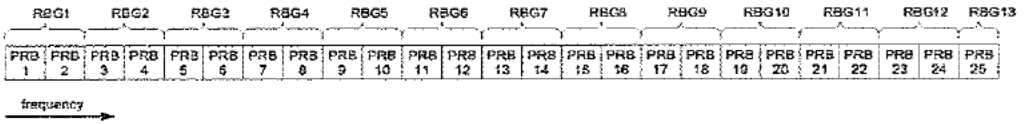


Figure 9.10: PRB addressed by a bitmap Type 0, each bit addressing a complete RBG.

Resource allocation Type 1. In resource allocations of Type 1, individual PRBs can be addressed (but only within a subset of the PRBs available). The bitmap used is slightly smaller than for Type 0, since some bits are used to indicate which subset of the RBG is addressed, and a shift in the position of the bitmap. The total number of bits (including these additional flags) is the same as for Type 0. An example for the case of $N_{\mathrm{RB}}^{\mathrm{DL}} = 25$, $N_{\mathrm{RBG}} = 11$ and $P = 2$ is shown in Figure 9.11.

The motivation for providing this method of resource allocation is flexibility in spreading the resources across the frequency domain to exploit frequency diversity.

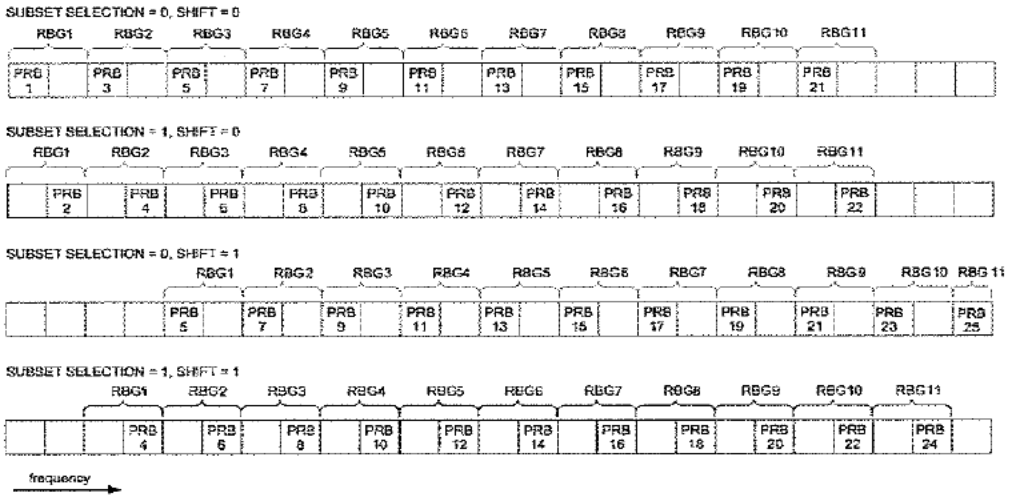


Figure 9.11: PRBs addressed by a bitmap Type 1, each bit addressing a subset of an RBG, depending on a subset selection and shift value.

Resource allocation Type 2. In resource allocations of Type 2, the resource allocation information indicates a contiguous set of PRBs, using either localized or distributed mapping (see Section 9.2.2.1) as indicated by a 1-bit flag in the resource allocation message. PRB allocations may vary from a single PRB up to a maximum number of PRBs spanning the system bandwidth. A Type 2 resource allocation field consists of a Resource Indication

Value (RIV) corresponding to a starting RB (RB_{START}) and a length in terms of contiguously-allocated RBs (L_{CRBs}). The resource indication value is defined by

$$\text{if} \quad (L_{CRBs} - 1) \leq \lfloor N_{RB}^{DL}/2 \rfloor \quad \text{then} \quad RIV = N_{RB}^{DL}(L_{CRBs} - 1) + RB_{START}$$
$$\text{else} \quad RIV = N_{RB}^{DL}(N_{RB}^{DL} - L_{CRBs} + 1) + (N_{RB}^{DL} - 1 - RB_{START})$$

An example of a method for reversing the mapping to derive the resource allocation from the RIV can be found in [2].

Resource allocation in Release 10. In addition to the above methods, Release 10 supports a non-contiguous resource allocation method for the uplink, allowing two separate contiguous sets of PRBs to be assigned, as explained in detail in Section 28.3.6.2.

9.3.5.5 PDCCH Transmission and Blind Decoding

The previous discussion has covered the structure and possible contents of an individual PDCCH message, and transmission by an eNodeB of multiple PDCCHs in a subframe. This section addresses the question of how these transmissions are organized so that a UE can locate the PDCCHs intended for it, while at the same time making efficient use of the resources allocated for PDCCH transmission.

A simple approach, at least for the eNodeB, would be to allow the eNodeB to place any PDCCH anywhere in the PDCCH resources (or CCEs) indicated by the PCFICH. In this case, the UE would need to check all possible PDCCH locations, PDCCH formats and DCI formats, and act on the messages with correct CRCs (taking into account that the CRC is scrambled with a UE identity). Carrying out such a 'blind decoding' of all the possible combinations would require the UE to make many PDCCH decoding attempts in every subframe. For small system bandwidths, the computational load would be reasonable, but for large system bandwidths, with a large number of possible PDCCH locations, it would become a significant burden, leading to excessive power consumption in the UE receiver. For example, blind decoding of 100 possible CCE locations for PDCCH Format 0 would be equivalent to continuously receiving a data rate of around 4 Mbps.

The alternative approach adopted for LTE is to define for each UE a limited set of CCE locations where a PDCCH may be placed. Such a constraint may lead to some limitations as to which UEs can be sent PDCCHs within the same subframe, which thus restricts the UEs to which the eNodeB can grant resources. Therefore it is important for good system performance that the set of possible PDCCH locations available for each UE is not too small.

The set of CCE locations in which the UE may find its PDCCHs can be considered as a 'search space'. In LTE, the search space is a different size for each PDCCH format. Moreover, separate *UE-specific* and *common* search spaces are defined; a UE-specific search space is configured for each UE individually, whereas all UEs are aware of the extent of the common search space. Note that the UE-specific and common search spaces may overlap for a given UE. The sizes of the common and UE-specific search spaces in Releases 8 and 9 are listed in Table 9.6.

With such small search spaces, it is quite possible in a given subframe that the eNodeB cannot find CCE resources to send PDCCHs to all the UEs that it would like to, because having assigned some CCE locations, the remaining ones are not in the search space of a

Table 9.6: Search spaces for PDCCH formats in Releases 8 and 9.

PDCCH format	Number of CCEs (n)	Number of candidates in common search space	Number of candidates in UE-specific search space
0	1	—	6
1	2	—	6
2	4	4	2
3	8	2	2

Note: The search space sizes in Release 10 are discussed in Section 28.3.1.1.

particular UE. To minimize the possibility of such blocking persisting into the next subframe, a UE-specific hopping sequence (derived from the UE identity) is applied to the starting positions of the UE-specific search spaces from subframe to subframe.

In order to keep under control the computational load arising from the total number of blind decoding attempts, the UE is not required to search for all the defined DCI formats simultaneously. Typically, in the UE-specific search space, the UE will always search for Formats 0 and 1A, which are the same size and are distinguished by a flag in the message. In addition, a UE may be required to receive a further format (i.e. 1, 1B, 1D, 2, 2A or 2B, depending on the PDSCH transmission mode configured by the eNodeB).

In the common search space, the UE will typically search for Formats 1A and 1C. In addition, the UE may be configured to search for Formats 3 or 3A, which have the same size as Formats 0 and 1A, and may be distinguished by having the CRC scrambled by a different (common) identity, rather than a UE-specific one.

Considering the above, a Release 8/9 UE is required to carry out a maximum of 44 blind decodings in any subframe (12 in the common search space and 32 in the UE-specific search space). This does not include checking the same message with different CRC values, which requires only a small additional computational complexity. The number of blind decodings required for a Release 10 UE is discussed in Section 28.3.1.1.

Finally, the PDCCH structure is also adapted to avoid cases where a PDCCH CRC 'pass' might occur for multiple positions in the configured search-spaces due to repetition in the channel coding (for example, if a PDCCH was mapped to a high number of CCEs with a low code rate, then the CRC could pass for an overlapping smaller set of CCEs as well, if the channel coding repetition was thus aligned). Such cases are avoided by adding a padding bit to any PDCCH messages having a size which could result in this problem occurring.

9.3.6 PDCCH Scheduling Process

To summarize the arrangement of the PDCCH transmissions in a given subframe, a typical sequence of steps carried out by the eNodeB is depicted in Figure 9.12.

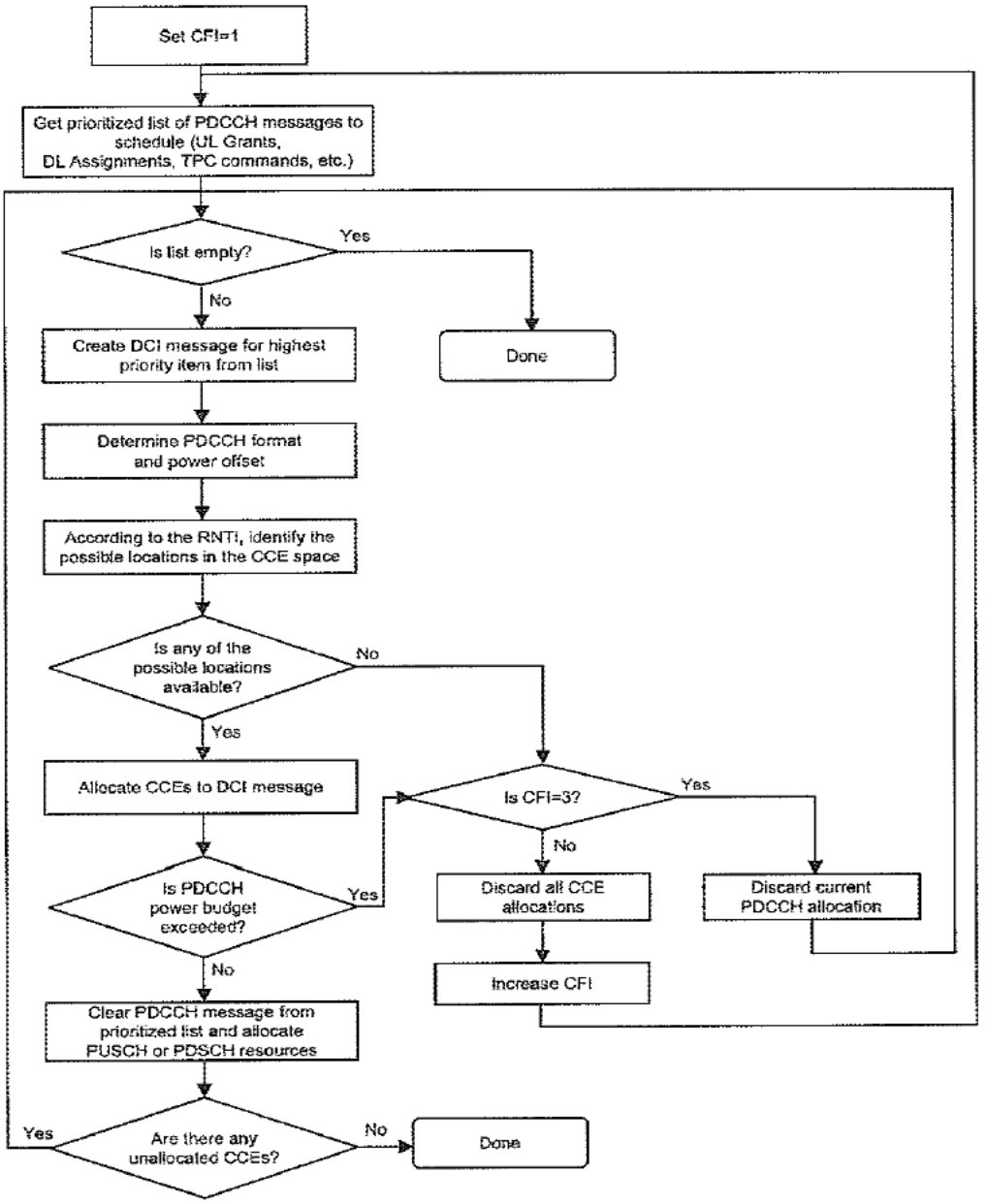


Figure 9.12: A typical sequence of PDCCH scheduling operations in a subframe.

References[12]

[1] 3GPP Technical Specification 36.321, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification', www.3gpp.org.

[2] NEC, 'R1-072119: DL Unicast Resource Allocation Signalling', www.3gpp.org, 3GPP TSG RAN WG1, meeting 49, Kobe, Japan, May 2007.

[12]All web sites confirmed 1st March 2011.

Samsung Ex. 1008
126 of 126