

Martin Sauter

Grundkurs Mobile Kommunikationssysteme

5G New Radio und Kernnetz, LTE-Advanced
Pro, GSM, Wireless LAN und Bluetooth

8. Auflage



Springer Vieweg

Grundkurs Mobile Kommunikationssysteme

Martin Sauter

Grundkurs Mobile Kommunikationssysteme

5G New Radio und Kernnetz, LTE-
Advanced Pro, GSM, Wireless LAN und
Bluetooth

8. Auflage



Springer Vieweg

Martin Sauter
Köln, Deutschland

ISBN 978-3-658-36962-0 ISBN 978-3-658-36963-7 (eBook)
<https://doi.org/10.1007/978-3-658-36963-7>

Die Deutsche Nationalbibliothek verzeichnetet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2004, 2006, 2008, 2011, 2013, 2015, 2018, 2022

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag, noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Planung/Lektorat: Leonardo Milla

Springer Vieweg ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Vorwort zur achten Auflage

Wie auch in den letzten zwei Jahrzehnten ist der Mobilfunk weiterhin einer der sich am schnellsten entwickelnden Zweige der Telekommunikation. Besonders in den letzten Jahren gab es einige gewaltige Umwälzungen, denen diese 8. Auflage mit deutlichen Änderungen zu den vorherigen Auflagen Rechnung trägt.

Seit der letzten Auflage wurde in vielen Teilen der Welt die fünfte Mobilfunkgeneration (5G) eingeführt und ein Großteil der Weiterentwicklung findet nun mit dieser Technologie statt. Aus diesem Grund widmet sich in dieser Auflage nun ein eigenständiges Kapitel diesem neuen System, das zunächst in Kombination mit 4G LTE betrieben wurde, nun aber auch mehr und mehr auf eigenen Beinen steht.

In der Praxis ist jedoch 4G LTE nach wie vor das Rückgrat aller Mobilfunknetzwerke und es gab auch hier zahlreiche Neuerungen, die nun zum Einsatz kommen. Hinzugekommen sind in dieser Auflage Details zu den Themen Downlink und Uplink Carrier Aggregation, Multi-Stream Übertragung (MIMO), sowie Betrachtungen zur Netzkapazität.

Stark an Bedeutung verloren haben in den letzten Jahren die Mobilfunksysteme der zweiten und dritten Generation, GSM und UMTS. Viele Netzbetreiber haben deswegen mittlerweile ihre UMTS Netzwerke abgeschaltet und nutzen das freigewordene Spektrum heute für LTE und 5G. Nach wie vor betreiben jedoch die meisten Mobilfunkbetreiber noch ihr GSM System, da es weiterhin für Teilnehmer mit älteren Endgeräten benötigt wird, die kein Voice over LTE (VoLTE) beherrschen. Auch für das internationale Roaming ist dies relevant, da die meisten Netzbetreiber heute noch kein VoLTE Roaming unterstützen. Und schließlich gibt es noch viele Geräte im industriellen Umfeld, die GSM und GPRS für die Kommunikation verwenden.

Aufgrund dieser zahlreichen Änderungen hat sich auch der Aufbau dieses Buches nun deutlich geändert. Statt zuerst auf die älteren Systeme einzugehen, rückt nun das erweiterte LTE Kapitel an den Anfang, gefolgt von einem neuen Kapitel über 5G New Radio (NR) und das 5G Kernnetz. Darauf folgt dann das Kapitel über den Sprachdienst, der in LTE und 5G Netzwerken verwendet wird (VoLTE, VoNR).

Im Nahfunkbereich sind natürlich auch weiterhin Wireless LAN (Wi-Fi) und Bluetooth von zentraler Bedeutung. Die entsprechenden Kapitel wurden für diese Auflage dem Stand der Technik angepasst. Enthalten sind nun z. B. auch eine Einführung in den

Wi-Fi 6 (802.11ax) Standard, die neue WPA3 Authentifizierung, Protected Management Frames, sowie das Inter-Access Point Roaming. Den Abschluss des Buches bilden dann die Kapitel über GSM und GPRS.

Bleibt mir noch, Ihnen an dieser Stelle viel Freude beim Studium dieses Buches und beim Experimentieren und Nutzen mobiler Kommunikation zu wünschen.

Köln
im Juni 2022

Martin Sauter

Vorwort zur ersten Auflage

Mobile Kommunikationssysteme wie GSM, GPRS, UMTS, Wireless LAN bieten heute eine große Vielfalt von Anwendungsmöglichkeiten. Um einen Einblick in die Technik dieser Systeme zu gewinnen, gibt es eine große Anzahl von Publikationen. In Buchform sind diese jedoch meist sehr umfangreich und für eine Einführung oft zu komplex. Publikationen im Internet hingegen sind meist nur sehr kurz und oberflächlich oder beschäftigen sich nur mit einer speziellen Eigenschaft eines Systems. Aus diesem Grund konnte ich während meiner Vorlesungen zu diesem Thema keine einzelne Publikation empfehlen, die eine Einführung in diese Systeme mit der nötigen Detailtiefe geboten hätte. Mit dem vorliegenden Buch möchte ich dies ändern.

Jedes der fünf Kapitel gibt eine detaillierte Einführung und Überblick über jeweils eines der zu Anfang genannten Systeme. Besonders wichtig ist mir auch, einen Eindruck zu vermitteln, welche Gedanken hinter der Entwicklung der unterschiedlichen Systeme standen. Neben dem „Wie“ ist also auch das „Warum“ zentraler Bestandteil jedes Kapitels. Außerdem wird durch zahlreiche Vergleiche zwischen den unterschiedlichen Technologien deutlich, wo die Anwendungsgebiete der einzelnen Systeme liegen. In manchen Fällen konkurrieren die Systeme miteinander, in vielen Fällen jedoch ergibt erst eine Kombination mehrerer Systeme eine interessante Anwendung. Abgerundet wird jedes Kapitel durch einen Fragen- und Aufgabenkatalog zur Lernzielkontrolle und Wiederholung.

Um einen tieferen Einblick in das eine oder andere System zu gewinnen, sind in den Kapiteln zahlreiche Verweise auf die entsprechenden Standards zu finden. Sie bilden eine ideale Ergänzung für einen tieferen Einblick in die einzelnen Systeme und sollten mit Hilfe der Hintergrundinformationen in diesem Buch auch etwas einfacher zu interpretieren sein.

Den Entschluss, mein Wissen zu diesen Themen als Buch zu veröffentlichen, fasste ich nach vielen theoretischen Gedankenspielen ganz spontan in einer Pariser Buchhandlung. Dort stieß ich zufällig auf ein Buch mit einem ganz anderen Themenschwerpunkt, mit dessen Autor ich jedoch den Umstand gemeinsam habe, dass wir für die gleiche Firma arbeiten. Ich nahm Kontakt mit ihm auf, und er schilderte mir während eines ausgedehnten Mittagessens, wie man von der ersten Idee zu einem fertigen Buch

kommt. An dieser Stelle möchte ich mich deshalb sehr herzlich bei Pierre Lescuyer bedanken, dessen Tipps mir beim Start meines eigenen Buchprojekts sehr weitergeholfen haben.

Außerdem gebührt mein großer Dank auch Berenike, die mir mit Ihrer Liebe und Freundschaft während dieses Projekts immer inspirierend zur Seite stand.

Weiterhin gebührt mein Dank auch Thomas Kempf, Christophe Schmid, Markus Rösch, Thomas Ehrle und ganz besonders Jörg Becker. Mit ihrem Wissen und großen Einsatz ihrer privaten Zeit haben sie mich vor einigen Fehlern bewahrt und in zahlreichen Gesprächen wichtige Anregungen und Verbesserungsvorschläge gegeben.

Nicht zuletzt gilt mein Dank auch Dr. Reinald Klockenbusch, der dieses Buchprojekt von Anfang an begleitet hat und an der Ausrichtung des Buches maßgeblich beteiligt war.

Paris
im Juni 2004

Martin Sauter

Inhaltsverzeichnis

1	Long Term Evolution (LTE) und LTE-Advanced	1
1.1	Einleitung und Überblick.	1
1.2	Netzwerk-Architektur und Schnittstellen	2
1.2.1	LTE-Endgeräte und die LTE Uu-Schnittstelle	3
1.2.2	Der eNodeB und die S1 und X2-Schnittstellen	6
1.2.3	Die Mobility Management Entity (MME).	10
1.2.4	Das Serving-Gateway.	11
1.2.5	Das PDN-Gateway.	12
1.2.6	Das HSS	14
1.2.7	Abrechnungssysteme	16
1.2.8	Qualitäts- und Prioritätssteuerung	16
1.3	Die LTE-Luftschnittstelle und das Radionetzwerk	17
1.3.1	OFDMA für die Datenübertragung im Downlink	18
1.3.2	SC-FDMA für Uplink Übertragungen	20
1.3.3	Quadrature Amplitude Modulation für Subcarrier.	21
1.3.4	Symbole, Slots, Radio Blocks und Frames	23
1.3.5	Referenz- und Synchronisations-Signale.	25
1.3.6	Das LTE-Kanalmodel in Downlink-Richtung	25
1.3.7	Downlink Management-Kanäle	27
1.3.8	System Information Messages (SIBs)	28
1.3.9	Das LTE-Kanalmodel in der Uplink-Richtung	29
1.3.10	Multiple Input Multiple Output Übertragungen.	29
1.3.11	HARQ und ARQ	32
1.3.12	PDCP – Komprimierung und Verschlüsselung	35
1.3.13	Der LTE Protokoll Stack	36
1.4	Scheduling	38
1.4.1	Downlink Scheduling.	38
1.4.2	Uplink Scheduling	42
1.5	Grundsätzliche Prozeduren	43
1.5.1	Netzwerksuche.	43

1.5.2	Attach und Aktivierung des Default Bearers	46
1.5.3	Handover Szenarien	50
1.6	Mobility Management und Leistungsoptimierung.	53
1.6.1	Mobilitätsmanagement im RRC-Connected State	53
1.6.2	Mobility Management im Idle State	56
1.6.3	Mobility Management und Zustandsänderungen in der Praxis	58
1.7	LTE Sicherheitsarchitektur	60
1.8	Zusammenspiel mit UMTS und GSM	61
1.8.1	Cell Reselection zwischen LTE und GSM/UMTS.	62
1.8.2	RRC Connection Release mit Redirect zwischen LTE und GSM/UMTS.	64
1.9	Carrier Aggregation	64
1.9.1	CA Varianten, Bandbreitenklassen und Bandkombinationen	66
1.9.2	CA Konfigurationen, Aktivierung und Deaktivierung	68
1.9.3	Uplink Carrier Aggregation	71
1.10	Von Dipolen zu Aktiven Antennen und Gigabit Backhaul	72
1.11	Self-Organizing Networks	74
1.12	Kapazität eines Standortes und Anzahl gleichzeitiger Nutzer	75
1.13	CS-Fallback für Sprache und SMS mit LTE	77
1.13.1	SMS über SGs	78
1.13.2	CS Fallback	79
1.14	Network Sharing – MOCN und MORAN	81
1.14.1	National Roaming	81
1.14.2	MOCN (Multi-Operator Core Network)	82
1.14.3	MORAN (Mobile Operator Radio Access Network)	84
1.15	IPv6 im Mobilfunk.	84
1.15.1	Das IPv6 Prefix und Interface IDs	85
1.15.2	IPv6 und Roaming	88
1.15.3	IPv6 und Wi-Fi Tethering.	88
1.15.4	IPv6-Only Kommunikation	90
1.16	Network Function Virtualization	91
1.16.1	Virtualisierung auf dem Desktop	91
1.16.2	Nutzung eines Betriebssystems in einer virtuellen Maschine	92
1.16.3	Das gleichzeitige Ausführen mehrerer virtueller Maschinen	93
1.16.4	Snapshots von Virtuellen Maschinen	94
1.16.5	Klonen einer Virtuellen Maschine	94
1.16.6	Virtualisierung in Rechenzentren in der Cloud	95

1.16.7	Administration von Virtuellen Maschinen in der Cloud	96
1.16.8	Network Function Virtualization	97
1.16.9	Router Virtualisierung	98
1.16.10	Software-Defined Networking	99
1.17	Machine Type Communication (MTC) und das Internet der Dinge	100
1.17.1	LTE CAT-1 Endgeräte	101
1.17.2	LTE CAT-0 Endgeräte und PSM	101
1.17.3	LTE CAT-M1 Endgeräte	102
1.17.4	LTE NB1 (NB-IoT) Endgeräte	103
1.17.5	NB-IoT Konfigurationsoptionen	104
1.17.6	Die NB-IoT Luftschnittstelle	104
1.17.7	NB-IoT Control Channels und Scheduling	105
1.17.8	NB-IoT Multi-Carrier Operation	106
1.17.9	NB-IoT Durchsatz und Anzahl der Geräte pro Zelle	107
1.17.10	NB-IoT Stromsparmechanismen	108
1.17.11	NB-IoT – High Latency Kommunikation	108
1.17.12	NB-IoT – Optimierung von IP und nicht-IP basierenden Datenübertragungen	110
1.17.13	NB-IoT Zusammenfassung	112
1.18	Fragen und Aufgaben	112
2	2 5G New Radio (NR) und das 5G Kernnetz	115
2.1	Einführung und Überblick	115
2.1.1	Gründe für den 5G Start als Hybrid Lösung	116
2.1.2	Frequency Range 1 und 2	117
2.1.3	Dynamic Spectrum Sharing in Low- und Mid-Bands	118
2.1.4	Netze in der Praxis und weiterer Aufbau des Kapitels	118
2.2	Die 5G Non-Standalone (NSA) Architektur	119
2.2.1	Netzwerk Architektur und Schnittstellen	119
2.2.2	3GPP 5G NR Option 2 und 3 mit Dynamic Spectrum Sharing	122
2.2.3	Unterschiede der Optionen 3, 3A und 3X	123
2.2.4	Das Fronthaul Interface	125
2.3	Die 5G TDD Luftschnittstelle	126
2.3.1	Flexibles OFDMA für den Downlink	127
2.3.2	Das 5G Resource Grid: Symbole, Slots, Resource Blocks und Frames	130
2.3.3	Synchronisation und Referenz Signale	132
2.3.4	Massive-MIMO für Beamforming und Multi-User Datenübertragung	133
2.3.5	TDD Slot Formate	136
2.3.6	Downlink Control Kanäle	139
2.3.7	Uplink Kanäle	140

2.3.8	Bandwidth Parts	141
2.3.9	Der Downlink Control Channel und das Scheduling	142
2.3.10	Downlink Geschwindigkeit in Theorie und Praxis	145
2.3.11	Downlink Datendurchsatz	147
2.3.12	Das TDD Air Interface in den mmWave Bändern (FR2)	148
2.4	Die 5G FDD Luftschnittstelle	149
2.4.1	Reframing und Dynamic Spectrum Sharing	151
2.5	EN-DC Bearer und Scheduling	156
2.5.1	Split Bearer und Flusskontrolle	157
2.5.2	Zwei Sender für EN-DC	158
2.6	Grundsätzliche Prozeduren und Mobility Management im Non-Standalone Mode	160
2.6.1	Aufbau eines LTE-Only Bearers als 5G Anker aus dem Flugmodus	160
2.6.2	Hinzufügen einer NR Zelle im Non-Standalone Modus	164
2.6.3	5G Anzeige im Display	169
2.6.4	Handover Szenarien	170
2.6.5	EN-DC Signaling Radio Bearers	173
2.6.6	5G Non-Standalone und VoLTE	174
2.7	Netzwerkplanung und Rollout Aspektse	174
2.7.1	Die Reichweite von Band n78	175
2.7.2	Backhaul Betrachtungen	175
2.8	Die 5G NR Standalone (SA) Architektur und grundsätzliche Prozeduren	176
2.8.1	Funktionen des 5G Kernnetz	176
2.8.2	Netzwerkschnittstellen	178
2.8.3	Teilnehmer und Geräte IDs	179
2.8.4	Prozeduren im 5G Kernnetz	180
2.8.5	Connection Management	180
2.8.6	Registration Management Prozeduren	181
2.8.7	Session Management	183
2.8.8	Mobility Management	187
2.8.9	Neue Sicherheitsfunktionen	188
2.8.10	Der 5G Kern und unterschiedliche RAN Optionen	191
2.8.11	Zusammenspiel der 5G und 4G Kernnetzwerke	191
2.8.12	Das 5G Kernnetz und SMS	196
2.8.13	Das Cloud Native 5G Kernnetz	197
2.9	Die 5G Standalone Luftschnittstelle	200
2.9.1	Der RRC Inactive Zustand	201
2.9.2	System Information Nachrichten	202
2.9.3	Messkonfiguration, Events und Handover	203
2.10	Network Slicing	204
2.11	Fragen	206

3	Voice over LTE und NR (VoLTE, VoNR)	209
3.1	Das Session Initiation Protocol (SIP)	210
3.2	Das IP Multimedia Subsystem (IMS) und VoLTE.	215
3.2.1	Architekturüberblick	215
3.2.2	IMS Registrierung	217
3.2.3	Der VOLTE Gesprächsaufbau.	219
3.2.4	LTE Bearer Konfiguration für VoLTE	221
3.2.5	Dedicated Bearer Setup mit Preconditions	224
3.2.6	Header Compression und DRX	226
3.2.7	Sprachcodecs und Aushandlung der Bandbreite	227
3.2.8	Freiton, Ring-Back Melodien und Early-Media	231
3.2.9	Verwendung von Ports	231
3.2.10	Filterung von Nachrichten und Asserted Identities	232
3.2.11	DTMF Töne	233
3.2.12	SMS über IMS	234
3.2.13	Konfiguration der Anrufweiterleitung und XCAP	235
3.2.14	Single Radio Voice Call Continuity	238
3.2.15	Wahl des Radionetzwerkes, T-ADS und VoLTE Interworking mit GSM und UMTS	242
3.2.16	VoLTE Notrufe.	243
3.3	VoLTE Roaming.	245
3.3.1	Option 1: VoLTE Local Breakout.	246
3.3.2	Option 2: VoLTE S8-Home Routing	247
3.4	Voice und 5G NR	249
3.4.1	IMS Signalisierung über 5G SA	250
3.4.2	5G NR EPS Fallback	252
3.4.3	5G Voice over NR (VoNR).	253
3.5	Voice over Wifi (VoWifi)	253
3.5.1	VoWifi Netzwerkarchitektur	254
3.5.2	VoWifi Handover	256
3.5.3	Wifi-Preferred und Cellular-Preferred	257
3.5.4	SMS, MMS und Supplementary Services über Wifi	258
3.5.5	VoWifi Roaming.	259
3.6	VoLTE und Festnetz IMS – Ein Vergleich	260
3.7	Fragen und Aufgaben.	261
4	Wireless LAN IEEE 802.11	265
4.1	Wireless LAN Überblick	265
4.2	Geschwindigkeiten und Standards.	266
4.3	WLAN-Konfigurationen: Von Ad-hoc bis Wireless Bridging	269
4.3.1	Ad-hoc, BSS, ESS und Wireless Bridging.	269
4.3.2	SSID und Frequenzwahl	273

4.4	Management-Operationen	275
4.5	Die MAC-Schicht	281
4.5.1	Zugriffssteuerung auf das Übertragungsmedium.	281
4.5.2	Der MAC Header	284
4.6	Physical Layer und MAC-Erweiterungen	286
4.6.1	IEEE 802.11b mit bis zu 11 Mbit/s	286
4.6.2	IEEE 802.11g mit bis zu 54 Mbit/s	289
4.6.3	IEEE 802.11a mit bis zu 54 Mbit/s	291
4.6.4	IEEE 802.11n mit bis zu 600 Mbit/s	292
4.6.5	IEEE 802.11ac mit bis zu 6,8 Gbit/s	302
4.6.6	IEEE 802.11ax – Wi-Fi 6 – High Efficiency Erweiterungen	308
4.7	Wireless LAN-Sicherheit	313
4.7.1	Wired Equivalent Privacy (WEP) und frühere Sicherheitsverfahren.	313
4.7.2	WPA und WPA-2 Personal Mode-Authentifizierung	313
4.7.3	WPA und WPA-2 Enterprise Mode Authentifizierung – EAP-TLS	315
4.7.4	WPA und WPA-2 Enterprise Mode Authentication – EAP-TTLS	317
4.7.5	WPA und WPA-2 Enterprise Mode Authentication – EAP-PEAP	319
4.7.6	WPA und WPA Enterprise Mode Authentifizierung – EAP-SIM	321
4.7.7	Verschlüsselung mit WPA und WPA-2	323
4.7.8	Wi-Fi Protected Setup (WPS)	324
4.7.9	WPA3 Personal Mode Authentication	326
4.7.10	Protected Management Frames	328
4.8	IEEE 802.11e und WMM – Quality of Service	330
4.9	Fragen und Aufgaben	337
5	Bluetooth	339
5.1	Überblick und Anwendungen	339
5.2	Physikalische Eigenschaften	341
5.3	Piconetze und das Master Slave Konzept	344
5.4	Der Bluetooth Protokoll Stack	347
5.4.1	Der Baseband Layer	347
5.4.2	Der Link Controller	353
5.4.3	Der Link Manager	356
5.4.4	Das HCI Interface	357
5.4.5	Der L2CAP Layer	360
5.4.6	Das Service Discovery Protocol	362

5.4.7	Der RFCOMM Layer	364
5.4.8	Aufbau einer Verbindung im Überblick	366
5.5	Bluetooth Sicherheit	367
5.5.1	Pairing bis Bluetooth 2.0	367
5.5.2	Pairing ab Bluetooth 2.1 (Secure Simple Pairing)	369
5.5.3	Authentifizierung	371
5.5.4	Verschlüsselung	371
5.5.5	Autorisierung	373
5.5.6	Sicherheitsmodi	374
5.6	Bluetooth Profile	375
5.6.1	Grundlegende Profile: GAP, SDP und Serial Profile	376
5.6.2	Object Exchange Profile: FTP, Object Push und Synchronize	377
5.6.3	Headset, Hands-Free und SIM-Access Profile	380
5.6.4	High Quality Audio Streaming	384
5.6.5	Das Human Interface Device (HID) Profile	387
5.7	Fragen und Aufgaben	389
6	GSM	391
6.1	Leitungsvermittelnde Datenübertragung	391
6.1.1	Klassische Leitungsvermittlung	392
6.1.2	Virtuelle Leitungsvermittlung über IP	394
6.2	Standards	395
6.3	Übertragungsgeschwindigkeiten	396
6.4	Das Signalisierungssystem Nr. 7	397
6.4.1	Klassischer SS-7-Protokollstack	398
6.4.2	Spezielle SS-7-Protokolle für GSM	401
6.4.3	IP-basierter SS-7-Protokollstack	402
6.5	Die GSM Subsysteme	403
6.6	Das Network Subsystem	404
6.6.1	Die Mobile Vermittlungsstelle (MSC), Server und Gateway	405
6.6.2	Das Visitor Location Register (VLR)	409
6.6.3	Das Home Location Register (HLR)	410
6.6.4	Das Authentication Center (AC)	413
6.6.5	Das Short Message Service Center (SMSC)	415
6.7	Das Base Station Subsystem (BSS) und Sprachcodierung	417
6.7.1	Frequenzbereiche	417
6.7.2	Base Transceiver Station (BTS)	419
6.7.3	Die GSM-Luftschnittstelle	420
6.7.4	Der Base Station Controller (BSC)	428
6.7.5	Die TRAU für Sprachdatenübertragung	433
6.7.6	Channel Coder und Interleaver in der BTS	438

6.7.7	Verschlüsselung	440
6.7.8	Modulation	442
6.7.9	Voice Activity Detection	442
6.8	Mobility Management und Call Control	444
6.8.1	Cell Reselection und Location Area Update	444
6.8.2	Mobile Terminated Call	446
6.8.3	Handoverszenarien	449
6.9	Mobile Endgeräte	452
6.9.1	Aufbau eines einfachen GSM Telefons	452
6.9.2	Aufbau eines Smartphones	454
6.10	Die SIM-Karte	456
6.11	Das Intelligent Network Subsystem und CAMEL	461
6.12	Fragen und Aufgaben	464
7	GPRS und EDGE	465
7.1	Leitungsvermittelte Datenübertragung	466
7.2	Paketorientierte Datenübertragung	467
7.3	GPRS auf der Luftschnittstelle	470
7.3.1	GPRS Timeslot-Nutzung im Vergleich zu GSM	470
7.3.2	Gleichzeitige Nutzung einer Basisstation von GSM und GPRS	472
7.3.3	Coding Schemes	473
7.3.4	EDGE (EGPRS)	474
7.3.5	Mobile Device Classes	476
7.3.6	Network Operation Mode (NOM)	477
7.3.7	GPRS-Kanalstruktur auf der Luftschnittstelle	479
7.4	GPRS-Zustandsmodell	481
7.5	GPRS-Netzwerkelemente	484
7.5.1	Die Packet Control Unit (PCU)	484
7.5.2	Der Serving GPRS Support Node (SGSN)	485
7.5.3	Der Gateway GPRS Support Node (GGSN)	487
7.6	GPRS Radio Resource Management	488
7.7	GPRS-Schnittstellen und Protokolle	491
7.8	GPRS Mobility und Session Management (GMM/SM)	496
7.8.1	Mobility Management-Aufgaben	496
7.8.2	GPRS Session Management	499
7.9	Session Management aus Anwendersicht	502
7.9.1	Leitungsvermittelter Verbindungsaufbau	503
7.9.2	GPRS-Verbindungsaufbau	504
7.10	Fragen und Aufgaben	506
Stichwortverzeichnis	509	



Long Term Evolution (LTE) und LTE-Advanced

1

1.1 Einleitung und Überblick

LTE, der Mobilfunkstandard der vierten Generation (4G), ist heute die Grundlage aller Mobilfunknetze weltweit. Entwickelt von Herstellern und Mobilfunkbetreibern in der 3rd Generation Partnership Project (3GPP) Organisation, brach 4G deutlich mit der Architektur bisheriger Mobilfunkstandards. Während zuvor die Sprachtelefonie fest in allen Schichten des Netzwerkes integriert war, ist LTE ein rein für die paketvermittelnde Datenübertragung konzipiertes System und für den schnellen Internetzugang optimiert. Wie auch heute im Festnetz ist die Sprachübertragung somit nur noch eine von vielen Anwendungen, die das IP Protokoll verwendet. Während es bei 2G und 3G Systemen noch unterschiedliche Standards in diversen Teilen der Welt gab, gelang es mit LTE zum ersten Mal, einen einheitlichen Standard zu schaffen. Dies ermöglicht heute weltweites Roaming für Kunden und günstigere Netzwerkkomponenten für Mobilfunkbetreiber. Stark gewachsen ist mit LTE auch das genutzte Spektrum. Während GSM Netzwerke mit 200 kHz (0,2 MHz) breiten Kanälen arbeiten, wurde dies bei UMTS auf 5 MHz ausgedehnt. Mit LTE wurde die maximale Kanalbandbreite dann auf 20 MHz erweitert. Mit Kanalbündelung (Carrier Aggregation) wird heute in Städten typischerweise 60–80 MHz Spektrum für die Datenübertragung zwischen LTE Netzwerk und einem Endgerät verwendet. Auch bei der Übertragungstechnik zwischen den Mobilfunkstandorten und dem dahinterliegenden Weitverkehrsnetzwerk fand mit LTE eine Revolution statt. Statt spezieller Standards und Übertragungsverfahren wird bei LTE auch für den Datentransport durchgehend auf das Internet Protokoll gesetzt, während auf den tieferen Schichten heute Glasfaser oder Mikrowelle, sowie Ethernet auf Layer 2 dominieren.

Während der Fokus auf IP deutliche Vorteile bei der Architektur des Systems brachte, erschwerte dies gleichzeitig die Einführung eines LTE Sprachdienstes. Da entschieden wurde, den leitungsvermittelten Sprachdienst von GSM und UMTS bei LTE

durch ein gänzlich neues System zu ersetzen, starteten alle Netzwerke zunächst mit der CS-Fallback Technik. Wie der Name andeutet, mussten LTE Endgeräte für die Sprachtelefonie in den ersten Jahren in ein GSM oder UMTS Radionetzwerk wechseln. Mittlerweile hat sich jedoch das standardisierte IP Multimedia Subsystem (IMS) und das Voice over LTE (VoLTE) Profil für die Sprachtelefonie durchgesetzt. Nach wie vor wird jedoch auch noch der alte Sprachdienst über GSM verwendet, z. B. mit älteren Endgeräten, beim Wechsel in ein nur mit GSM versorgtes Gebiet, sowie vielfach noch im internationalen Roaming.

Die ursprüngliche Version des LTE-Standards wurde in 3GPP Release 8 spezifiziert und erste Netzwerke wurden 2009 in Betrieb genommen. In nachfolgenden 3GPP Releases wurden dann zahlreiche Erweiterungen wie z. B. das schon erwähnte Carrier Aggregation eingeführt, um mit dem wachsenden Bandbreitenbedarf schrittzuhalten.

Ganz andere Anforderungen an ein Mobilfunknetz haben Anwendungen und Geräte aus dem Bereich des Internet of Things (IoT). IoT Geräte sind oft nur mit kleinen Batterien ausgestattet und können in vielen Fällen auch wegen ihrer Anbringung nicht einfach ersetzt oder neu aufgeladen werden. Außerdem kommunizieren solche Geräte nur sehr sporadisch. Somit unterscheiden sich solche Geräte bezüglich Effizienz und Stromverbrauch stark von Smartphones und anderen Geräten, die eine hohe Übertragungsgeschwindigkeit benötigen. Aus diesem Grund wurden von 3GPP eine Anzahl von Erweiterungen spezifiziert, von denen die Wichtigste die Narrow-Band IoT (NB-IoT) Erweiterung des LTE Radionetzwerkes ist.

Das erste Kapitel dieses Buches ist nun wie folgt aufgebaut: Nach einem kurzen Architekturüberblick und Einführung der Funktionen im Kernnetzwerk, widmet sich das Kapitel dann dem Radionetzwerk, sowie den Verbindungs- und Mobilitätsprozeduren. Im zweiten Teil des Kapitels wird dann nochmals genauer auf die heutige Architektur des Kernnetzwerkes eingegangen, sowie die LTE Erweiterungen für das Internet der Dinge besprochen.

1.2 Netzwerk-Architektur und Schnittstellen

Wie auch frühere Netzgenerationen ist das LTE Netzwerk in ein Radionetz und ein Kernnetz getrennt. Die Anzahl der logischen Komponenten wurde jedoch reduziert, um dadurch die Effizienz zu steigern, um Kosten zu senken und Latenzzeiten zu minimieren. Abb. 1.1 gibt einen Überblick über die Komponenten eines LTE-Netzwerkes und wie diese miteinander verbunden sind. Die nachfolgenden Unterkapitel geben dann einen Überblick über die einzelnen Komponenten und deren Aufgaben.

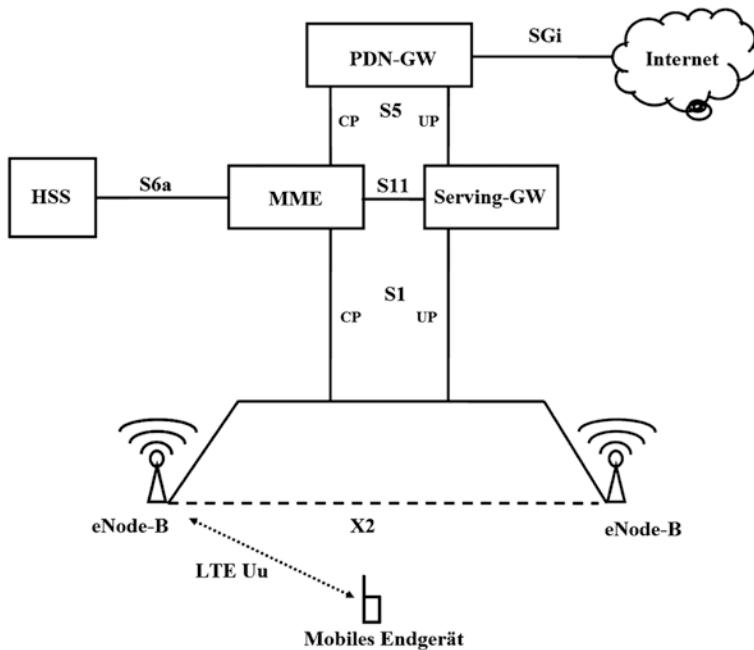


Abb. 1.1 LTE-Netzwerk im Überblick

1.2.1 LTE-Endgeräte und die LTE Uu-Schnittstelle

In der LTE-Spezifikation werden Endgeräte wie Smartphones, Tablets und andere Geräte als User Equipment (UE) bezeichnet. In 3GPP Release 8 wurden die Endgerätekategorien (UE Categories) 1–5 in 3GPP TS 36.306¹ spezifiziert und definieren, mit welcher maximalen Datenübertragungsrate ein Endgerät senden und empfangen kann. In späteren 3GPP Releases wurden dann zusätzliche UE Klassen für Geräte spezifiziert, die Carrier Aggregation unterstützen. Im Unterschied zur ursprünglichen Definition gibt es hier nun für Downlink und Uplink getrennte Kategorien. Abb. 1.2 zeigt eine Auswahl an Endgerätekategorien, die heute in der Praxis verwendet werden. Wichtig für das Verständnis ist an dieser Stelle, dass eine UE-Kategorie nur eine Aussage über die maximal unterstützte Datenrate eines Endgerätes in Downlink und Uplinkrichtung macht. Mit welcher Kombination von Funktionalitäten diese Geschwindigkeit erreicht werden kann (Anzahl der aggregierten Carrier, Anzahl der MIMO Kanäle, Verwendung von 256-QAM), ist nicht Teil der Definition. Die Endgerätekategorie macht auch keine Aussage, ob diese Funktionen im Netzwerk vorhanden sind. Während Geräte in hohen Kategorien theoretisch bis zu 8 MIMO Streams unterstützen können, aber nicht müssen, werden heute von den meisten Netzwerken maximal 4 Streams verwendet.

Kategorie (DL/UL)	4	6	9	16 / 5	20 / 18
Max. Downlink-Geschwindigkeit (Mbit/s)	150	300	450	1000	2000
Typische Anzahl von Carriern im Downlink	1	2	3	4	5
Anzahl der typischen / max. MIMO Downlink Streams	2	2 / 4	4	4	4 / 8
Max. Modulation im Downlink	64-QAM	64-QAM	64-QAM	256-QAM	256-QAM
Max. Uplink-Geschwindigkeit (Mbit/s)	50	50	50	75	150
Typische Anzahl von Carriern im Uplink	1	1	1	1	2
Max. Modulation im Uplink	16-QAM	16-QAM	16-QAM	64-QAM	256-QAM

Abb. 1.2 LTE-Endgerätekategorien (Auswahl)

In Uplink-Richtung unterstützen alle Endgeräte die etwas langsamere aber dafür robustere 16-QAM Modulation. Manche Endgerätekategorien unterstützen zusätzlich 64- oder sogar 256-QAM im Uplink.

Für alle aufgelisteten Endgeräteklassen ist die Unterstützung von mindestens 2×2 MIMO Übertragung im Downlink Pflicht. Mit dieser Übertragungstechnik können zwei Datenströme auf dem gleichen Kanal mit zwei Antennen in der Basisstation zu zwei Antennen im Endgerät übertragen werden. Erreichen die Signale den Empfänger über unterschiedliche Pfade, die z. B. durch Reflexionen an Objekten und der räumlichen Trennung zwischen den Sendeantennen, und auf der Empfängerseite durch räumliche Trennung der Empfangsantennen, kann der Empfänger zwischen den unterschiedlichen Datenströmen unterscheiden. Die Anzahl der Sende- und Empfangsantennen entscheidet, wie viele Datenströme parallel gesendet werden können. Die meisten LTE-Endgeräte

haben heute zwei Empfangsantennen, teurere Geräte unterstützen mittlerweile auch vier Empfangsantennen. Bezeichnungen wie 2×2 MIMO oder 4×4 MIMO geben die Anzahl der Antennen auf der Sender- und Empfängerseite an. Endgeräte haben heute jedoch üblicherweise noch weitere Antennen, da LTE in zahlreichen Frequenzbändern ausgestrahlt wird und weitere Funktechnologien wie WLAN und Bluetooth parallel verwendet werden. Dies lässt die Komplexität des Antennendesigns weiter steigen.

In der Praxis sind heute die meisten Endgeräte in den Kategorien 4–20 angesiedelt und erreichen eine maximale Datenrate unter sehr guten Übertragungsbedingungen von 100 bis 150 Mbit/s pro 20 MHz Kanal, von denen die Endgeräte je nach UE-Kategorie mehrere bündeln können. In der Praxis werden solche Geschwindigkeiten jedoch nur selten erreicht, da Störungen (Interferenz) von Nachbarzellen, viele gleichzeitige Nutzer in einer Zelle und eine vom Zellmittelpunkt abnehmende Empfangsstärke den Durchsatz einschränken.

Während frühere Netze generationen nur wenige Frequenzbänder nutzten, startete LTE in einer großen Anzahl an unterschiedlichen Bändern, um Spektrum weltweit und auch lokal bestmöglich zu nutzen. Abb. 1.3 gibt einen Überblick, in welchen Frequenzbändern LTE in Europa heute vorwiegend verwendet wird. Diese und weitere Bänder für andere Weltregionen sind im Detail in 3GPP TS 36.101² beschrieben. Aktuell verwenden die meisten Netzbetreiber in Europa eine Kombination der Bänder 1, 3, 8, 7 und 20. Während in Band 20 im 800 MHz Bereich in den meisten Ländern nur Kanalbreiten von 10 MHz und somit Datenraten von etwa 75 Mbit/s unter optimalen Bedingungen möglich sind, sind in Band 1 im 2100 MHz Bereich, in Band 3 im 1800 MHz Bereich und in Band 7 im 2600 MHz Bereich Kanalbreiten bis zu 20 MHz möglich und werden in der Praxis auch eingesetzt. Der Vorteil von Band 20 im 800 MHz Bereich ist eine größere Reichweite in ländlichen Gebieten und eine bessere Versorgung innerhalb von Gebäuden aufgrund der geringeren Dämpfungseigenschaften. Vorteil von Band 1, 3 und 7 von 1800 bis 2600 MHz ist dagegen eine deutlich höhere maximale Datenrate aufgrund des breiteren Kanals.

Viele LTE Endgeräte, die in Europa verkauft werden, unterstützen heute typischerweise alle genannten Bänder, sowie zusätzlich viele weitere, die in anderen Regionen der

Abb. 1.3 Die gebräuchlichsten LTE-Bänder in Europa

Band	Downlink (MHz)	Uplink (MHz)	Auch verwendet für
1	2110 – 2170	1920 – 1980	UMTS
3	1805 – 1880	1710 – 1785	GSM
7	2620 – 2690	2500 – 2570	
8	925 – 960	880 – 915	GSM, UMTS
20	791 – 821	832 – 862	
38	2570 – 2620	2570 – 2620	TD-LTE (Time Division Duplex)

Welt verwendet werden. Viele hochpreisige Smartphones beherrschen heute mehr als 20 LTE Bänder. Neben allen Frequenzbändern für Europa können diese Endgeräte auch in vielen Bändern, die nur in Nordamerika und Asien anzutreffen sind, verwendet werden. Günstigere Endgeräte unterstützen zwar weniger Bänder, beherrschen jedoch typischerweise zumindest die wichtigsten Bänder für den nordamerikanischen Kontinent.

Durch die Kombination von allgemein unterstützten Bändern in Endgeräten, akzeptablen Preisen für das Datenroaming außerhalb Europas, und der generellen Verfügbarkeit von LTE Roamingabkommen ist somit ein globaler und schneller Internetzugang über LTE ohne Wechsel der SIM Karte möglich geworden.

1.2.2 Der eNodeB und die S1 und X2-Schnittstellen

Die komplexeste Baugruppe in einem LTE-Netzwerk ist die Basisstation, die in den Standards als eNodeB bezeichnet wird. Der Name leitet sich von NodeB ab, die Bezeichnung einer Basisstation in UMTS. Das vorangestellte ‚e‘ steht für ‚evolved‘. Auch anderen Abkürzungen in den LTE-Standards, die schon aus anderen Systemen bekannt sind, wurde ein ‚e‘ vorangestellt. Ein Beispiel ist das Radionetzwerk, das in UMTS als UTRAN (UMTS Terrestrial Radio Access Network) bezeichnet wird. In LTE wird das Radio Netzwerk somit als E-UTRAN bezeichnet.

Ein eNodeB setzt sich aus folgenden Komponenten zusammen:

- Einem Digitalmodul, auch Baseband Unit (BBU) genannt, das die eigentliche Signalverarbeitung übernimmt und auch als digitales Interface zum Kernnetzwerk dient.
- Einem Radiomodul, auch Remote Radio Unit (RRU) genannt. Es empfängt das digitale Signal der BBU und wandelt es in ein Radiosignal um. In der Empfangsrichtung wandelt die RRU ein analoges Signal in einen digitalen Datenstrom um, der dann im Basebandmodul verarbeitet wird.
- Den Antennen, der sichtbarste Teil eines Mobilfunknetzwerkes. Diese sind direkt an die RRU angeschlossen.

Viele Netzbetreiber verwenden heute eine optische Verbindung zwischen Radiomodul und Digitalmodul. Auf diese Weise kann das Radiomodul sehr nahe an der Antenne angebracht werden und somit Kosten für teure Koaxial-Kupferkabel gespart werden.

An dieser Stelle ist es wichtig, auch das Konzept des ‚Bearers‘ einzuführen, da dieser Begriff an zahlreichen Stellen dieses Kapitels verwendet wird. Ein Bearer ist eine logische Verbindung zwischen Netzwerkelementen und beschreibt die Quality of Service (QoS) Attribute einer Verbindung, wie z. B. die maximal zulässige Verzögerung und die maximale Übertragungsgeschwindigkeit einer Verbindung für einen Teilnehmer. Übertragungen zwischen einem Endgerät und der Basisstation werden über sogenannte Radio Access Bearer (RAB) geleitet. Der RAB wird einem Endgerät während des Verbindungsaufbaus zugeteilt und beinhaltet zum einen den Signaling Radio Bearer (SRB),

um Nachrichten für das Session Management, das Mobility Management und die Radio Resource Kontrolle (RRC) zu übertragen. Zum anderen wird über einen RAB mindestens ein Data Radio Bearer (DRB) übertragen, über den IP Nutzdatenpakete ausgetauscht werden.

Während in früheren Netzgenerationen die Basisstation wenig mehr als ein intelligentes Modem war, bilden LTE-Basisstationen autonome Einheiten. Beim Design der Netzwerkarchitektur entschloss man sich, die meisten der Funktionalitäten, die früher in einem zentralen Radioknoten untergebracht waren (RNC in UMTS, BSC in GSM), in die Basisstationen zu übertragen. Somit ist ein eNodeB nicht nur für die Luftschnittstelle (Air Interface) verantwortlich, sondern auch für:

- das User Management und die Aufteilung der Ressourcen auf dem Air Interface an mehrere gleichzeitige Teilnehmer.
- die Sicherstellung von Quality of Service (QoS)-Attributen für einzelne Verbindungen wie eine maximale Verzögerungszeit und der Bereitstellung einer minimalen Bandbreite in Abhängigkeit des Nutzerprofils.
- das Mobilitätsmanagement.
- das Interferenzmanagement, also die Reduktion des Einflusses der eigenen Sendetätigkeit auf die Übertragungen der Nachbarstationen.

Beispielsweise entscheidet der eNodeB selber, eine Verbindung während eines Datentransfers an einen benachbarten eNodeB zu übergeben. Auch Handover werden unabhängig von anderen Komponenten im Netzwerk durchgeführt und das Kernnetzwerk erst nach Durchführung der Prozedur unterrichtet.

Die Luftschnittstelle wird bei LTE als Uu-Interface bezeichnet und ist die einzige Schnittstelle in einem Mobilfunknetzwerk, die immer drahtlos ist. Die theoretische Spitzengeschwindigkeit hängt von der Bandbreite des Kanals ab. LTE ist hier sehr flexibel und definiert eine Anzahl unterschiedlicher Kanalbandbreiten von 1,25 bis 20 MHz. In einer 20 MHz Konfiguration mit 2×2 MIMO und 256-QAM Modulation können Datenraten bis zu 200 Mbit/s erreicht werden. Die in der Praxis erreichbaren Geschwindigkeiten hängen jedoch von vielen Faktoren ab, wie zum Beispiel dem Abstand eines Endgerätes zum eNodeB, der verwendeten Leistung, der Interferenz von Nachbarstationen, etc. Aus diesem Grund sind die erzielbaren Geschwindigkeiten in der Praxis meist wesentlich geringer.

Die Schnittstelle zwischen Basisstation und dem Kernnetz wird als S1 Interface bezeichnet. Üblicherweise wird diese Verbindung über eine optische Verbindung oder alternativ auch über eine Mikrowellenverbindung geführt. Mit auf Ethernet basierender Technologie lassen sich bei allen Übertragungsformen Geschwindigkeiten im Gbit/s-Bereich erreichen. Dies reicht aus, um eine Basisstation mit drei oder mehr Sektoren zu versorgen, über die heute jeweils mehr als ein 20 MHz Kanal (Carrier) übertragen wird. Zusätzlich werden über eine solche Backhaul-Verbindung auch Datenströme von GSM, UMTS und 5G NR Basisstationen geleitet, die am gleichen Ort installiert sind. Aus

diesen Gründen übersteigt der Bandbreitenbedarf der Backhaul-Verbindung die Bandbreite eines einzelnen LTE-Sektors bei weitem.

Das S1 Interface hat zwei logische Hälften, die jedoch zusammen über die gleiche physische Verbindung übertragen werden. Die Nutzdaten, also die IP Pakete der Endgeräte, werden über den S1-UP (S1 User Plane)-Protokollstack übertragen. Jedes Endgerät hat dafür mindestens einen GTP (GPRS Tunneling Protocol) Tunnel, in dem dessen IP Pakete eingepackt und dann übertragen werden. In einem solchen Paket wird das IP Protokoll somit zweimal verwendet. Wie in Abb. 1.4 und auf der rechten Seite in Abb. 1.5 gezeigt, kommt auf Layer 3 des Protokoll Stacks das IP Protokoll zum ersten Mal zum Einsatz. Hier sind die IP Adressen des eNodeB und des Serving-Gateways zu finden, um das Datenpaket innerhalb des Mobilfunknetzwerks weiterzuleiten. Die IP Adresse des Endgerätes und des Servers im Internet befinden sich, eingepackt im GTP Tunnel, weiter oben im Protokollstack und werden erst am Gateway zum Internet wieder ausgepackt. Auf diese Weise ist es möglich, die Mobilität des Nutzers im Mobilfunknetzwerk zu gewährleisten. Wechselt das Endgerät zu einem anderen eNodeB, wird nur die IP Adresse auf Layer 3 eines GTP Paketes modifiziert, die IP Adresse des Endgeräts weiter oben im Protokoll Stack ändert sich nicht. Somit ist die Mobilität des Endgerätes für die Gegenstelle im Internet transparent. Außerdem ist es damit möglich, dass ein

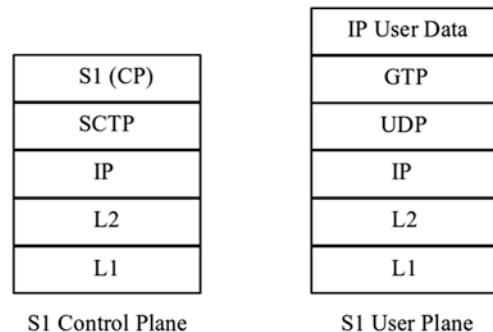
```

> Frame 1: 118 bytes on wire (944 bits), 118 bytes captured (944 bits)
> Ethernet II, Src: 
> Internet Protocol Version 4, Src: 10.108.8.8 (10.108.8.8), Dst: 10.122.18.20 (10.122.18.20)
  0100 .... = Version: 4
  .... 0101 = Header Length: 20 bytes (5)
> Differentiated Services Field: 0x38 (DSCP: AF13, ECN: Not-ECT)
  Total Length: 104
  Identification: 0x6a92 (27282)
> Flags: 0x0000
  Fragment offset: 0
  Time to live: 59
  Protocol: UDP (17)
  Header checksum: 0x8865 [validation disabled]
  [Header checksum status: Unverified]
  Source: 10.108.8.8 (10.108.8.8)
  Destination: 10.122.18.20 (10.122.18.20)
> User Datagram Protocol, Src Port: 2152, Dst Port: 2152
> GPRS Tunneling Protocol
> Internet Protocol Version 4, Src: 172.4.3.34 (172.4.3.34), Dst: 8.8.8.8 (8.8.8.8)
  0100 .... = Version: 4
  .... 0101 = Header Length: 20 bytes (5)
> Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
  Total Length: 60
  Identification: 0x4462 (17506)
> Flags: 0x4000, Don't fragment
  Fragment offset: 0
  Time to live: 64
  Protocol: UDP (17)
  Header checksum: 0x3726 [validation disabled]
  [Header checksum status: Unverified]
  Source: 172.4.3.34 (172.4.3.34)
  Destination: 8.8.8.8 (8.8.8.8)
> User Datagram Protocol, Src Port: 50467, Dst Port: 53
> Domain Name System (query)

```

Abb. 1.4 Protokollsichten eines GTP Pakets

Abb. 1.5 S1 Control Plane und User Plane-Protokoll-Stack



Endgerät auch zwischen unterschiedlichen Radiotechnologien (z. B. GSM, LTE, 5G-NR) wechseln kann und dabei seine IP Adresse behält.

Die Protokolle auf den Schichten 1 und 2 der S1-Schnittstelle sind in den Standards nicht weiter definiert und werden somit in Abb. 1.5 nur als Layer 1 (L1) und Layer 2 (L2) bezeichnet. Somit können hier die unterschiedlichsten Protokolle zum Einsatz kommen, die für den Transport von IP-Paketen verwendet werden können.

Der S1-Control Plane (S1-CP)-Protokollstack, definiert in 3GPP TS 36.413³, wird für zwei Aufgaben benötigt: Zum einen kommuniziert der eNodeB über S1-CP mit dem Kernnetz für seine eigenen Zwecke, zum Beispiel um sich beim Netzwerk anzumelden, um Status- und Keep-Alive-Nachrichten zu senden und um Konfigurationsinformationen vom Netzwerk zu erhalten. Zum anderen wird S1-CP verwendet, um benutzerspezifische Signalisierung auszutauschen. Möchte sich beispielsweise ein Endgerät am Netzwerk anmelden, wird für diesen Nutzer eine logische Verbindung mit dem Kernnetz aufgebaut und darüber dann die Authentisierung und der Aufbau des Tunnels für die Nutzerdaten abgewickelt. Nachdem die Verbindung für einen Nutzer aufgebaut wurde, wird S1-CP dann verwendet, um die Verbindung aufrecht zu erhalten und gegebenenfalls an eine andere LTE, UMTS oder GSM-Basisstation weiterzugeben.

Abb. 1.5 zeigt den S1-CP-Protokoll-Stack auf der linken Seite. Das IP-Protokoll ist wiederum die Grundlage. Statt jedoch TCP oder UDP auf Layer 4 zu verwenden, wird das Stream Control Transmission Protocol (SCTP) verwendet, das in RFC 4960⁴ definiert wurde. Seine Aufgaben sind die Verwaltung vieler gleichzeitiger Signalisierungsverbindungen, das Sicherstellen der korrekten Reihenfolge von Datenpaketen mit Signalisierungsinformationen, sowie Flusskontrolle und Überlastmanagement.

In den vorangegangenen 3GPP-Radionetzwerken werden die Basisstationen von einer zentralen Instanz kontrolliert. In GSM ist dies der Base Station Controller (BSC) und in UMTS der Radio Network Controller (RNC). In diesen Systemen ist die BSC bzw. der RNC für den Aufbau einer Verbindung über die Luftschnittstelle zuständig, für die Kontrolle und Aufrechterhaltung der Verbindung, für die Sicherstellung des QoS und für den Handover einer Verbindung von einer Basisstation zur nächsten. Bei LTE wurde

dieses Konzept nicht weitergeführt, um zum einen die Latenzzeit bei der Übertragung von Nutzdatenpaketen zu reduzieren und zum anderen die für die Verwaltungsaufgaben beträchtliche Prozessorleistung von einem zentralen Element auf viele Basisstationen zu verteilen. Dies ist besonders bei der paketorientierten Datenübertragung von Vorteil, da hier die Verbindung über die Luftschnittstelle oft an die aktuelle Aktivitätssituation angepasst werden muss, um den Stromverbrauch des Endgeräts zu reduzieren.

Um ohne eine zentrale Koordination im Radionetzwerk eine Verbindung weiterreichen zu können, kommunizieren LTE Basisstationen direkt über die X2 Schnittstelle miteinander. Auf diese Weise können eNodeBs das sogenannte Handover einer Verbindung selber kontrollieren. Falls die Nachbarzelle bei einem Handover bekannt ist, können die eNodeBs über das X2 Interface direkt kommunizieren. Falls dies nicht möglich ist, kann der Handover auch mithilfe des Kernnetzes über das S1 Interface abgewickelt werden. Nachbarschaftsbeziehungen werden entweder vom Netzwerkbetreiber vorkonfiguriert oder können während des Betriebs mithilfe der Endgeräte auch dynamisch ermittelt werden. Diese Funktionalität wird auch als Automatic Neighbor Relation (ANR) bezeichnet.

1.2.3 Die Mobility Management Entity (MME)

Während eNodeBs alle Entscheidungen zur Aufrechterhaltung einer Verbindung zu einem Endgerät über die Luftschnittstelle treffen, ist die Benutzerverwaltung eine zentralisierte Aufgabe im Kernnetzwerk. Dies ist notwendig, da es im Mobilfunknetzwerk zumindest einen festen Router geben muss, über den IP-Pakete für eine Verbindung zwischen Endgerät und Internet übertragen werden. Außerdem wird eine Nutzerdatenbank benötigt, um Authentisierungsdaten und Nutzerprofile an einem zentralen Ort zu speichern. Somit kann sich ein Endgerät an jeder Stelle des Netzwerkes anmelden und sich auch in ausländische Netzwerke einbuchen.

Der Netzwerkknoten, der für die Signalisierung zwischen den eNodeBs und dem Kernnetzwerk verantwortlich ist, wird als Mobility Management Entity (MME) bezeichnet. Abb. 1.1 zeigt die logische Anordnung der MME im Überblick. In großen Netzwerken werden üblicherweise viele MMEs verwendet, da zum einen das Signalisierungsaufkommen die Kapazität einer MME übersteigt und auch, um eine Redundanz bei auftretenden Defekten zu haben. Da sich die Aufgaben der MME nicht auf die Luftschnittstelle beziehen, sondern sich auf die Benutzerverwaltung beschränkt, wird die Kommunikation zwischen eNodeB und MME auch als Non-Access Stratum (NAS)-Signalisierung bezeichnet. Konkret ist die MME für folgende Aufgaben zuständig:

- **Authentifizierung:** Bei der ersten Anmeldung eines Teilnehmers nach Einschalten des Geräts baut dieses automatisch eine Verbindung zum LTE-Netzwerk auf, bevor Daten übertragen werden können. Der eNodeB kommuniziert dann mit der MME über das

S1 Interface, um das Endgerät zu identifizieren. Die MME fordert dazu Authentifizierungsinformationen vom Home Subscriber Server (HSS) an, der weiter unten noch genauer beschrieben wird. Nach positiver Authentifizierung wird dann die Verschlüsselung auf der Luftschnittstelle aktiviert.

- Aufbau von Bearern: Da die MME nur für die Signalisierung, nicht jedoch für den eigentlichen Nutzdatentransfer zuständig ist, ist eine ihrer weiteren Aufgaben die Kommunikation mit anderen Netzwerknoten für die Erstellung eines IP-Tunnels zwischen einem eNodeB und einem Gateway zu einem externen Netzwerk, üblicherweise dem Internet.
- NAS Mobility Management: Ist ein Endgerät für eine längere Zeit inaktiv (üblich sind in der Praxis Zeiten zwischen 10 und 30 s), wird die Verbindung über die Luftschnittstelle und den logischen Tunnel des Benutzers im Radionetzwerk deaktiviert. Das Endgerät entscheidet dann selber über Zellwechsel und es findet keinerlei Signalisierung mit dem Netzwerk statt, solange sich die neuen Zellen in der gleichen Tracking Area befinden. Dies spart Energie im Endgerät und mindert die Signalisierungslast im Netzwerk. Sollten in dieser Inaktivitätsphase IP-Pakete aus dem Internet für das Endgerät ankommen, sendet die MME eine Paging-Nachricht an alle eNodeBs, die Teil der letzten bekannten Tracking Area des Endgeräts sind. Nachdem sich das Endgerät auf das Paging meldet, ist der genaue Aufenthaltsort wieder bekannt und die Bearer können wieder aufgebaut werden.
- Handover-Unterstützung: Falls zwischen zwei eNodeBs kein X2 Interface vorhanden ist, hilft die MME bei der Koordinierung des Handovers zwischen den Zellen.
- Interworking mit anderen Radionetzwerken: Befindet sich ein Endgerät an der Grenze eines LTE-Abdeckungsbereiches, kann sich eine eNodeB entschließen, ein Endgerät an ein GSM oder UMTS-Netzwerk zu übergeben. Im Idle-Zustand wird diese Entscheidung vom Endgerät selber ohne das LTE-Netzwerk getroffen. In beiden Fällen, die später noch genauer beschrieben werden, ist die MME für die Koordination des Wechsels verantwortlich.

Für diese und weitere Aufgaben wurden zwischen den Netzwerkkomponenten eine Anzahl von Schnittstellen wie S5, S6a, S11 und SGs definiert. Diese sind in Abb. 1.1 in einer Übersicht gezeigt und werden in den nachfolgenden Unterkapiteln nun genauer beschrieben.

Für die Weiterleitung der Nutzdaten zwischen Kern- und Radionetzwerk ist in LTE das Serving-Gateway (S-GW) zuständig, das nachfolgend beschrieben wird.

1.2.4 Das Serving-Gateway

Das Serving-Gateway (S-GW) ist für die Weiterleitung von Nutzerdaten in IP-Tunneln zwischen den eNodeBs und dem PDN-Gateway verantwortlich, das die Verbindung im LTE Kernnetzwerk zum Internet herstellt. Auf der Radionetzwerkseite terminiert das

S-GW die S1-UP (User Plane) GTP-Tunnel und auf der Kernnetzseite die S5-UP GTP-Tunnel zum Internet Gateway. Die S1 und S5-Tunnel für einen Nutzer sind unabhängig voneinander und können bei einem Handover bei Bedarf dynamisch geändert werden. Wenn beispielsweise während eines Handovers der Tunnel zwischen zwei Basisstationen umgeschaltet werden muss, ändert sich somit nur der Endpunkt des S1-Tunnels für die Weiterleitung der Nutzerdaten im Radionetzwerk zum neuen eNodeB. Wird der neue eNodeB jedoch von einer anderen MME und einem anderen Serving-Gateway bedient, muss auch der S5-Tunnel modifiziert werden. Das Erstellen von neuen Tunneln und deren Modifikation wird vom MME kontrolliert, und Kommandos zum Serving-Gateway werden über die S11-Schnittstelle, wie in Abb. 1.1 gezeigt, übertragen. Für die S11-Schnittstelle wird das GTP-C (Control)-Protokoll verwendet.

1.2.5 Das PDN-Gateway

Die dritte Kernnetzkomponente ist das Packet Data Network Gateway (PDN-Gateway). In der Praxis bildet diese Komponente den Übergang zum Internet. In manchen Netzwerken wird das PDN-Gateway auch für Anbindung an Firmennetzwerke über eine verschlüsselte Verbindung verwendet, um Mitarbeitern einen direkten mobilen Zugang in ihr Firmennetzwerk zu ermöglichen. Wie zuvor schon erwähnt, terminiert das PDN-Gateway die S5 Schnittstelle.

Auf der User Plane werden die Nutzdatenpakete für die Übertragung zwischen Serving-Gateway und dem PDN-Gateway in S5 GTP-Tunneln übertragen. Das S-GW sendet dann Nutzdaten, die vom PDN-Gateway eintreffen, über die S1-Schnittstelle an den aktuellen eNodeB der Verbindung weiter, welcher dann die Daten über die Luftschnittstelle an das Endgerät weiterleitet.

Eine weitere Aufgabe des PDN-Gateway ist die Vergabe von IP-Adressen an Endgeräte. Während der Verbindungsaufnahme nach dem Einschalten kontaktiert der eNodeB zunächst die MME für die Authentifizierung des Endgeräts. Nach erfolgreicher Authentifizierung fordert die MME dann vom PDN-Gateway eine IP-Adresse für das Endgerät an. Für diese Prozedur wird das S5-CP (Control Plane)-Protokoll verwendet. Wenn das PDN-Gateway den Zugang zum Netzwerk genehmigt, gibt es eine IP-Adresse zurück an die MME, die diese dann an das Endgerät weiterleitet. Teil dieser Prozedur ist auch der Aufbau der S1 und S5 Nutzdatentunnel.

In der Praxis werden einem Endgerät typischerweise mehrere IP Adressen gleichzeitig zugewiesen. Eine Adresse wird üblicherweise für den Internetzugang verwendet, eine Weitere für Voice over LTE (VoLTE). Das Endgerät ist somit mit zwei Netzwerken gleichzeitig verbunden, dem Internet und dem internen IP Multimedia Subsystem (IMS) Netzwerk für den Sprachdienst. Auf Betriebssystemebene sind diese zwei Verbindungen als zwei unabhängige Netzwerkschnittstellen sichtbar.

Durch den Mangel an verfügbaren öffentlichen IPv4 Adressen teilen die meisten Netzbetreiber nur private IP Adressen an Endgeräte zu. Mit Network Address Trans-

lation (NAT) werden dann viele private IPv4 Adressen auf wenige öffentliche IPv4 Adressen abgebildet. Das gleiche Verfahren wird auch von DSL-, Kabel- oder Glasfaser routern im Festnetz verwendet. Vergibt der Netzbetreiber öffentliche IPv4 Adressen an diese Geräte, vergeben diese dann private IPv4 Adressen an alle Geräte, die im Haushalt angeschlossen sind. Der Nachteil dieses Ansatzes ist, dass Endgeräte nicht direkt aus dem Internet erreichbar sind, da der Verbindungsauflauf immer vom Endgerät selber ausgelöst werden muss. Nur beim Aufbau der Verbindung von innen kann der NAT Router eine Verknüpfung (Mapping) zwischen der internen IP Adresse und dem TCP oder UDP Port mit der externen IP Adresse und dem dort verwendeten TCP oder UDP Port erzeugen.

Ein Vorteil von NAT im Mobilfunk ist jedoch, dass unberechtigte Verbindungsanfragen, z. B. von Schadprogrammen, die das Netzwerk auf anfällige Geräte absuchen, automatisch abgewehrt werden. Auch IP Pakete, die aus dem Internet an einen Port bei der Neuvergabe eines Mappings an einen anderes Endgerät eingehen, können somit erkannt und am PDN-GW verworfen werden. Dies schützt nicht nur das Endgerät bis zu einem gewissen Grad, sondern senkt auch den Energieverbrauch des Endgeräts, da solche Pakete nicht zugestellt werden und somit der Stromsparmodus nicht verlassen werden muss⁵.

Das PDN-GW spielt auch beim internationalen Roaming eine wichtige Rolle. Damit auch im Ausland auf LTE Netzwerke zugegriffen werden kann, verbinden Netzbetreiber weltweit ihre Netze, damit auch vom Ausland aus auf die Teilnehmerdatenbank im Heimatnetzwerkes eines ausländischen Teilnehmers zugegriffen werden kann. Wird nach der Authentifizierung des Teilnehmers dann der Internetzugang für den Teilnehmer gewährt, wird ein GPRS Tunneling Protocol (GTP) Tunnel zwischen dem S-GW im besuchten (visited) Netzwerk und dem PDN-GW im Heimatnetzwerk aufgebaut. Dieser Prozess ist großteils mit dem Aufbau eines Tunnels über das S5 Interface identisch, wenn sich sowohl das S-GW und das PDN-GW im Heimatnetzwerk befinden. Um diese zwei Szenarien auseinanderzuhalten, wird im internationalen Roaming nicht vom S5, sondern vom S8 Interface gesprochen. In der Praxis werden Mobilfunknetzwerke über das IP Roaming Exchange (IPX) Netzwerk miteinander verbunden, einem privaten und vom Internet abgetrennten IP Netzwerk. Der Nachteil dieses sogenannten Home Routings ist, dass die Nutzdatenpakete zunächst zurück zum Heimnetzwerk des Teilnehmers gesendet werden und erst von dort dann zum Internet. Die Alternative dazu wird ‚Local Breakout‘ genannt und wurde ebenfalls spezifiziert. In dieser Variante wird zusätzlich das PDN-GW des ausländischen Netzbetreibers verwendet und Pakete können somit direkt ins Internet weitergeleitet werden. In der Praxis wird dies jedoch heute nicht verwendet, da hierzu unter anderem andere Abrechnungsverfahren notwendig wären.

In vielen Fällen wird heute in Netzwerken das S-GW und das PDN-GW im gleichen physischen Netzknoten betrieben. In diesem Fall ist das S5 Interface nur virtuell vorhanden. Während des Roamings sind S-GW und PDN-GW jedoch getrennt. Auch die MME könnte im gleichen Netzknoten betrieben werden. Da die MME jedoch nur für die

Signalisierung verantwortlich ist und somit mehr CPU Leistung und weniger Netzwerkverkehr erzeugt, werden für die MME oft andere Server verwendet.

1.2.6 Das HSS

Für die Verwaltung von Teilnehmern verwendet LTE eine Datenbank, die als Home Subscriber Server (HSS) bezeichnet wird. Für die Kommunikation mit der Datenbank wird in LTE das IP-basierte DIAMETER-Protokoll für den Informationsaustausch verwendet. Standardisiert ist dieses Protokoll in RFC 3588⁶, und die Schnittstellenbezeichnung in LTE lautet S6a.

Wichtigster Parameter jedes Teilnehmereintrags ist die International Mobile Subscriber Identity, kurz IMSI genannt. Die IMSI ist eine weltweit eindeutige Nummer, die einen Teilnehmer identifiziert und bei fast allen teilnehmerbezogenen Signalisierungsvorgängen in GSM, UMTS, LTE und 5G NR verwendet wird. Neben dem HSS ist die IMSI eines Teilnehmers auch auf der SIM Karte gespeichert und besteht aus folgenden Teilen:

- Dem Mobile Country Code (MCC): Dieser gibt an, aus welchem Land der Teilnehmer stammt. Nachfolgend eine Tabelle mit einigen MCCs:

MCC	Land
262	Deutschland
232	Österreich
228	Schweiz
208	Frankreich
310	USA
604	Marokko
505	Australien

- Dem Mobile Network Code (MNC): Dieser bestimmt, aus welchem Netzwerk der Teilnehmer stammt. Dies ist notwendig, da es in einem Land üblicherweise mehrere unabhängige Mobilfunknetzwerke gibt. In Deutschland gibt es z. B. folgende Mobile Network Codes: 01 für T-Mobile, 02 für Vodafone und 03 für Telefonica O₂.
- Die Mobile Subscriber Identification Number (MSIN): Diese Nummer ist im nationalen Netzwerk eindeutig (Abb. 1.6).

Da die IMSI international eindeutig ist, kann sich ein Endgerät damit auch in ausländischen Netzwerken einbuchen. Beim Einschalten übermittelt das Endgerät, wie auch im Heimatnetzwerk, die auf der SIM-Karte des Teilnehmers gespeicherte IMSI an das

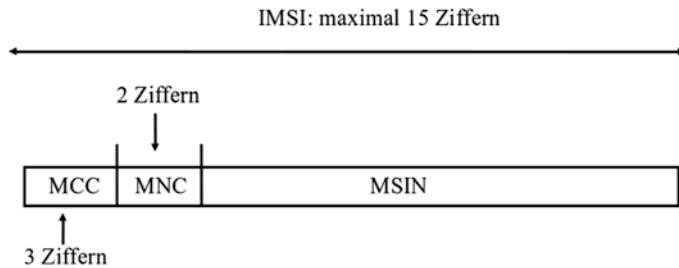


Abb. 1.6 Die IMSI

ausländische Netzwerk. Anhand der ersten Ziffern erkennt die MME dann, aus welchem Land (MCC) und aus welchem Netzwerk (MNC) der Teilnehmer stammt und kann somit das HSS im Heimnetzwerk des Teilnehmers nach dessen Daten fragen.

Mit der IMSI als Schlüssel verwaltet das HSS für jeden Teilnehmer zahlreiche Parameter, die beschreiben, welche Dienste ein Teilnehmer nutzen darf. Dies sind zum Beispiel:

- Die maximale Datenrate im Up- und Downlink.
- Informationen über die Tarif des Teilnehmers (nicht standardisiert).
- Welches Quality of Service Profil für den Internet Zugang verwendet wird.
- Ob der Teilnehmer VoLTE Nutzen darf.
- Welche Netztechnologien (GSM, UMTS, LTE, 5G NR) der Teilnehmer verwenden darf.
- Welche Dienste und Netztechnologien im internationalen Roaming verwendet werden dürfen.

Auch für den GSM und LTE Sprachdienst enthält das HSS zahlreiche Parameter. Diese werden später in den jeweiligen Kapiteln genauer beschrieben.

Eine weitere wichtige Funktion des HSS und des darin integrierten Authentication Centers (AuC) ist die Aufbewahrung des geheimen Schlüssels jedes Teilnehmers. Auf der Teilnehmerseite wird dieser auf der SIM Karte gespeichert und kann nicht ausgelesen werden. Wie später noch detaillierter beschrieben, wird bei Verbindungsaufnahme eines Endgerätes mit dem Netzwerk eine Authentifizierungsprozedur ausgeführt, an der auch das HSS mit dem AuC beteiligt ist. Nachdem der Teilnehmer authentifiziert ist, generiert das HSS mit dem AuC temporäre Integritätscheck- und Ciperhing Keys und sendet diese über die MME an das Endgerät und den eNodeB. Die geheimen Schlüssel der Teilnehmer hingegen werden nie über das Netzwerk gesendet.

In GSM und UMTS wird das HSS als Home Location Register (HLR) bezeichnet. Im 5G Kernnetzwerk wurden das HSS sowie die Authentifizierungskomponente in drei Teile getrennt, in den Universal Data Management (UDM) Service, das Universal Data Repository (UDR), sowie die Authentication Server Function (AUSF). In der Praxis ist

es üblich, dass es für alle Netzgenerationen eines Betreibers eine gemeinsame Datenbank und Authentifizierungsfunktion gibt. Diese stellt dann für die jeweilige Kernnetzgeneration die entsprechenden Schnittstellen bereit. Nur so ist es möglich, dass ein Endgerät nahtlos zwischen unterschiedlichen Netzgenerationen wechseln kann.

1.2.7 Abrechnungssysteme

Die Netzwerknoten und Interfaces, die in den letzten Abschnitten vorgestellt wurden, sind zwingend in einem LTE Netzwerk notwendig. Zusätzlich gibt es noch eine Reihe weiterer Netzwerknoten und Funktionen für die Abrechnung und Qualitätssteuerung.

Für die Abrechnung der verbrauchten Datenmenge eines Nutzers werden sogenannte „Billing Records“ im Netzwerk, z. B. in der MME, erzeugt. Diese Records werden gesammelt und an ein Abrechnungssystem (Charging Server) weitergeleitet. Hier wird dann einmal im Monat eine Rechnung für den Nutzer ausgestellt. Diese Art der Abrechnung wird auch als „offline billing“ oder „postpaid billing“ bezeichnet.

Eine andere Art der Abrechnung ist das sogenannte „Online Charging“, das auch als „Prepaid Billing“ bekannt ist. Diese Abrechnungsmethode ist nicht standardisiert und die Interaktion mit dem Nutzer ist implementationsabhängig. Typischerweise wird für die Interaktion mit dem Nutzer eine Webseite oder eine App verwendet. Dort kann der Nutzer dann Tarife buchen oder bei aufgebrauchtem Datenvolumen weiteres Datenvolumen freischalten.

1.2.8 Qualitäts- und Prioritätssteuerung

Für Echtzeitanwendungen wie z. B. Sprachtelefonie über LTE (VoLTE) ist es wichtig, eine minimale Bandbreite für die Verbindung sowie eine gleichbleibende Paketverzögerungszeit auf allen Schnittstellen im Netzwerk zu garantieren. Dies ist besonders dann notwendig, wenn eine Zelle schon stark ausgelastet ist. Diese Quality of Service (QoS) Voraussetzungen können mit der Policy Control Resource Function (PCRF) gewährleistet werden. Dienste eines Netzbetreibers wie z. B. VoLTE können über die standardisierte Rx Schnittstelle eine Anforderung für ein QoS Profil an die PCRF senden. Die PCRF übersetzt dann diese Anfrage in Kommandos für das PDN-GW und das S-GW. Diese wiederum sorgen dann für die Einhaltung der QoS Anforderungen für eine Verbindung im Kern- und Zugangsnetzwerk. Die PCRF ist Teil der 3GPP IMS Spezifikation und war ursprünglich für die Verwendung als Teil des IMS gedacht. In der Praxis wird diese jedoch heute auch von anderen nicht-IMS Diensten eines Netzbetreibers, wie z. B. für Bandbreiten- und Volumenlimitierung verwendet. Da die PCRF Teil des Mobilfunknetzwerkes ist und somit nicht über das Internet erreichbar ist, können nur Dienste eines Netzbetreibers QoS Anforderungen stellen. Internet basierte

Dienste für Sprachübertragung wie z. B. Signal, WhatsApp, etc. haben hingegen keinen Zugriff auf die PCRF.

1.3 Die LTE-Luftschnittstelle und das Radionetzwerk

Die wichtigste Neuerung von LTE im Vergleich zu früheren 3GPP-Mobilfunknetzwerken ist die komplett neu definierte Luftschnittstelle. Um den Grund für diesen Richtungswechsel besser zu verstehen, ist ein kurzer Blick auf die mit der Zeit entstandenen Limitationen von GSM und UMTS nötig:

Die GSM-Luftschnittstelle basiert auf 200 kHz-Kanälen, die in jeweils 8 sich wiederholende Zeitschlitzte (Timeslots) für Sprachtelefonie aufgeteilt sind. Ein Timeslot war ursprünglich für die Übertragung eines Telefongesprächs gedacht, und die Anzahl der Timeslots entspricht somit der maximalen Anzahl gleichzeitiger Gespräche auf einem Kanal. Um die Anzahl gleichzeitiger Gespräche zu erhöhen, werden mehrere Kanäle pro Basisstation verwendet. Später wurde die Luftschnittstelle mit GPRS für die paketorientierte Datenübertragung erweitert, indem definiert wurde, wie die Timeslots auch für IP-basierte Datenübertragung verwendet werden können. Das Limit der 200 kHz-Kanäle und die geringe Anzahl an Timeslots pro Kanal blieb jedoch bestehen.

Mit UMTS wurden diese Limitationen dann durch Einführung eines im Vergleich zu GSM sehr breitbandigen Kanals von 5 MHz überwunden. Statt Datenströme durch Zeitschlitzte zu trennen, verwendet UMTS das Code Division Multiple Access (CDMA) Verfahren, das unterschiedliche Datenströme gleichzeitig, aber mit unterschiedlichen Codes überträgt. Auf der Empfängerseite sind die verwendeten Codes bekannt und die gleichzeitigen Datenströme können somit wieder voneinander getrennt werden. Mit der HSPA (High Speed Packet Access) Erweiterung von UMTS wurde dieses Verfahren weiterverwendet, es kam jedoch wieder eine Timeslot-Struktur hinzu, um die paketorientierte Datenübertragung zu optimieren. Diese Timeslots sind jedoch nicht für die Sprachübertragung optimiert, sondern für die Übertragung von IP-Datenpaketen.

Durch die anhaltende Weiterentwicklung der zur Verfügung stehenden Technik und Prozessorleistung war es einige Jahre später dann abermals möglich, die Kanalbandbreite zu erhöhen. UMTS war jedoch in dieser Hinsicht sehr inflexibel, da sich das CDMA-Übertragungsverfahren nicht für noch breitere Kanäle eignet. Wird die Kanalbandbreite erhöht, verkürzen sich mit CDMA als Konsequenz die Zeiten für jeden Übertragungsschritt. Dies ist zwar bezüglich der Rechenleistung kein Problem, bereitet aber im Zusammenhang mit Effekten, die das Signal zwischen Sender und Empfänger verändern, große Probleme. Während der Übertragung wird das Signal an Objekten reflektiert, und der Empfänger sieht nicht nur ein Signal, sondern mehrere leicht zeitversetzte Echos von abgelenkten Signalteilen, die einen längeren Übertragungsweg hatten. Wird nun die Zeit für jeden Übertragungsschritt zu weit gesenkt, überlappen sich die Echos von zwei aufeinander folgenden Übertragungsschritten.

1.3.1 OFDMA für die Datenübertragung im Downlink

Um die Limitierungen bezüglich der verwendeten Bandbreite von GSM und UMTS zu überwinden, verwendet LTE ein Übertragungsverfahren namens Orthogonal Frequency Division Multiple Access (OFDMA). Anstatt Daten mit einer hohen Geschwindigkeit über einen einzigen Kanal zu senden, teilt OFDMA den Datenstrom in viele langsamere Datenströme auf, die dann gleichzeitig über viele Kanäle übertragen werden. Der große Vorteil dieses Verfahrens ist, dass die einzelnen Übertragungsschritte nun sehr viel länger dauern und der oben beschriebene Überlappungseffekt, auch Multi-Path Fading genannt, nicht dominiert. Abb. 1.7 zeigt eine Übersicht über die Anzahl der Subkanäle in Abhängigkeit der verwendeten Bandbreite für LTE. Je mehr Bandbreite zur Verfügung steht, desto mehr LTE Subcarrier werden verwendet. Für einen 10 MHz-Kanal werden beispielsweise 600 Subcarrier verwendet. Jeder einzelne parallele Datenstrom kann somit 600 mal langsamer sein als der gesamte Datenstrom.

Um Bandbreite zu sparen sind die Subcarrier so angeordnet, dass die Seitenausläufer des Signals auf Frequenzebene beim Scheitelpunkt des benachbarten Trägers genau null sind. Diese Eigenschaft wird auch als Orthogonalität bezeichnet. Um die vielen gleichzeitigen Datenströme zu dekodieren, wird eine mathematische Funktion verwendet, die als Inverse Fast Fourier Transformation (I-FFT) bezeichnet wird. Der Input für die I-FFT sind viele zeitbasierte parallele Signale, die dann als Amplitude auf unterschiedliche Frequenzen (Kanäle) abgebildet werden. Abb. 1.8 zeigt das Prinzip dieses Ansatzes. Oben links wird ein sehr schneller eingehender Datenstrom zunächst in viele langsamere, dafür aber parallele Datenströme aufgeteilt, die dann den einzelnen Subcarriern auf der Frequenzachse zugewiesen werden. Mit der I-FFT-Funktion wird dann ein einzelnes zeitbasiertes Signal erzeugt, das danach moduliert und über die Luftschnittstelle gesendet wird.

Bandbreite	Anzahl der Subcarrier	FFT Größe
1.25 MHz	76	128
2.5 MHz	150	256
5 MHz	300	512
10 MHz	600	1024
15 MHz	900	1536
20 MHz	1200	2048

Abb. 1.7 Definierte Bandbreiten für LTE

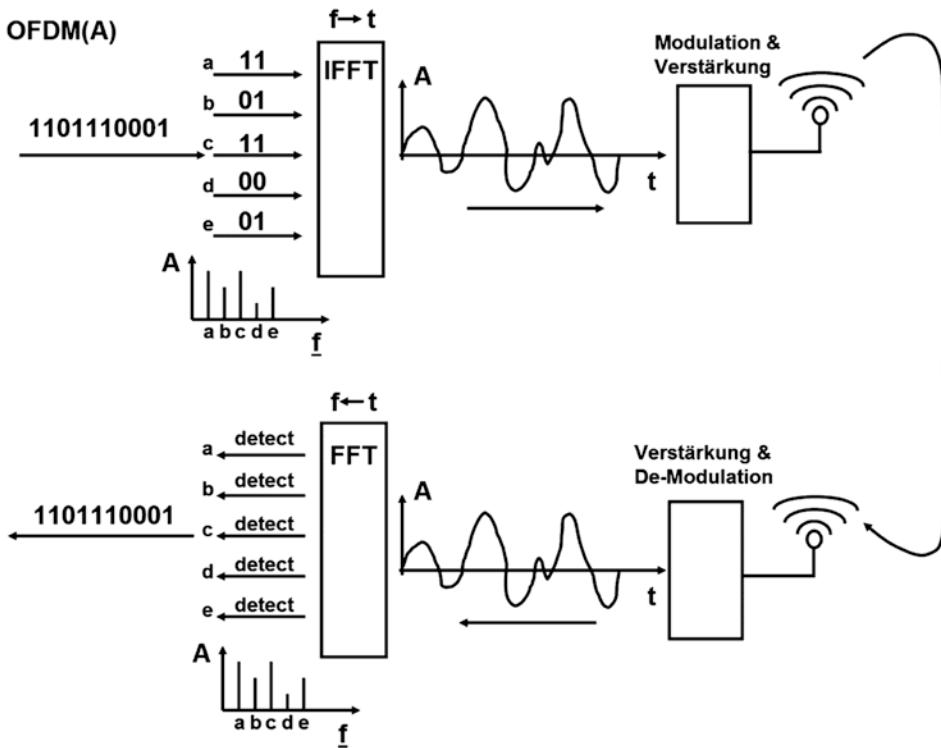


Abb. 1.8 Prinzip der OFDM-Datenübertragung in Downlink-Richtung

Die Empfangsseite ist im unteren Teil von Abb. 1.8 zu sehen. Nach der Demodulation wird das Eingangssignal in einen Fast Fourier Transformation (FFT)-Block geleitet, der das Zeitsignal in ein Frequenzsignal umwandelt, in dem dann die einzelnen Subcarrier wieder detektiert werden können. Im letzten Schritt werden dann die langsamten Datenströme wieder zu einem schnellen Datenstrom zusammengefügt und an die höheren Schichten des Protokollstacks übergeben.

LTE verwendet die folgenden physikalischen Parameter für die Subcarrier:

- Subcarrier-Abstand: 15 kHz
- Länge eines Übertragungsschritts (OFDM-Symbollänge): $66,667 \mu\text{s}$
- Länge des normalen Cyclic Prefix: $4,7 \mu\text{s}$. Das Cyclic Prefix wird vor jedem OFDM-Symbol übertragen, um die Inter-Symbol Interferenz zu minimieren. In sehr schwierigen Umgebungen mit vielen verschiedenen und sehr langen Übertragungswegen kann auch ein längeres Cyclic Prefix von $16,67 \mu\text{s}$ verwendet werden. Ein Nachteil ist jedoch ein geringerer Durchsatz, da die eigentliche Symbollänge konstant bleibt und somit weniger Symbole pro Zeiteinheit übertragen werden können.

An dieser Stelle ist es interessant, die mit 15 kHz sehr schmalbandigen LTE Subcarrier mit den 200 kHz-Kanälen von GSM zu vergleichen. Dies verdeutlicht, wie langsam die Datenübertragung ist und die dadurch gewonnene Robustheit gegen Inter-Symbol Interferenz. Weiterhin ist wichtig, dass die Bandbreite der Subcarrier unabhängig von der Bandbreite des gesamten LTE-Kanals ist. Für breitere Kanäle werden einfach mehr Subcarrier verwendet. Mehrere Nutzer können in diesem System durch Verteilung auf unterschiedliche Subcarrier gleichzeitig bedient werden. Zusätzlich ist auch die Verteilung der Nutzer auf unterschiedliche Übertragungszeiten möglich.

1.3.2 SC-FDMA für Uplink Übertragungen

Für die Datenübertragung in Uplink-Richtung ist OFDMA weniger ideal, da die Leistungsdifferenz zwischen der durchschnittlichen Übertragungsleistung und maximaler Übertragungsleistung (Peak to Average Power Ratio, PAPR) sehr groß ist. In der Praxis muss ein Verstärker in der Lage sein, die maximale Leistung zu unterstützen. Dies legt den durchschnittlichen Leistungsverbrauch fest, der unabhängig von der mittleren benötigten Übertragungsleistung ist. Bei OFDM wird jedoch die maximale Leistung nur selten benötigt und somit ist der PAPR sehr groß. Somit ist es besser, ein Übertragungsverfahren zu wählen, das einen geringeren PAPR aufweist, um somit eine bessere Balance zwischen Leistungsaufnahme und maximaler Datenrate zu finden.

Für eine Basisstation ist ein hoher PAPR-Wert hingegen kein Problem, da der Stromverbrauch für den Verstärker hier nur eine untergeordnete Rolle spielt. Für batteriebetriebene Endgeräte sollte der Sender jedoch so effizient wie möglich arbeiten. Aus diesem Grund entschloss sich die 3GPP, eine andere Übertragungsart für den Uplink zu wählen, Single Carrier – Frequency Division Multiple Access (SC-FDMA). Dieses Übertragungsverfahren ist OFDMA sehr ähnlich, die Übertragungskette hat jedoch, wie in Abb. 1.9 gezeigt, zwei zusätzliche Funktionsblöcke.

In einem ersten Schritt, der links oben in der Abbildung gezeigt wird, wird der Eingangsdatenstrom geliefert. Statt diesen aufzuteilen und dann direkt auf die Substreams abzubilden, wird der Datenstrom zunächst in ein frequenzbasiertes Signal mit einer Fast Fourier Transformation (FFT) umgewandelt. Dadurch wird erreicht, dass jedes einzelne Bit über alle Subcarrier verteilt wird. Die Anzahl der Subcarrier, die im Uplink verwendet werden, richtet sich nach den aktuellen Übertragungsbedingungen, der maximalen Sendeleistung des Endgeräts und der Anzahl der gleichzeitigen Nutzer. Subchannels, die nicht von einem Endgerät verwendet werden, werden bei der weiteren Verarbeitung im Sendemodul auf 0 gesetzt. Der aus der FFT entstehende Frequenzvektor wird dann wie bei OFDMA einer I-FFT-Funktion übergeben, die aus den Daten wieder ein zeitbasiertes Signal erstellt. Mathematisch kann gezeigt werden, dass der PAPR dieses Signals wesentlich niedriger ist als ein Signal, das ohne die zusätzliche FFT-Stufe erzeugt wurde.

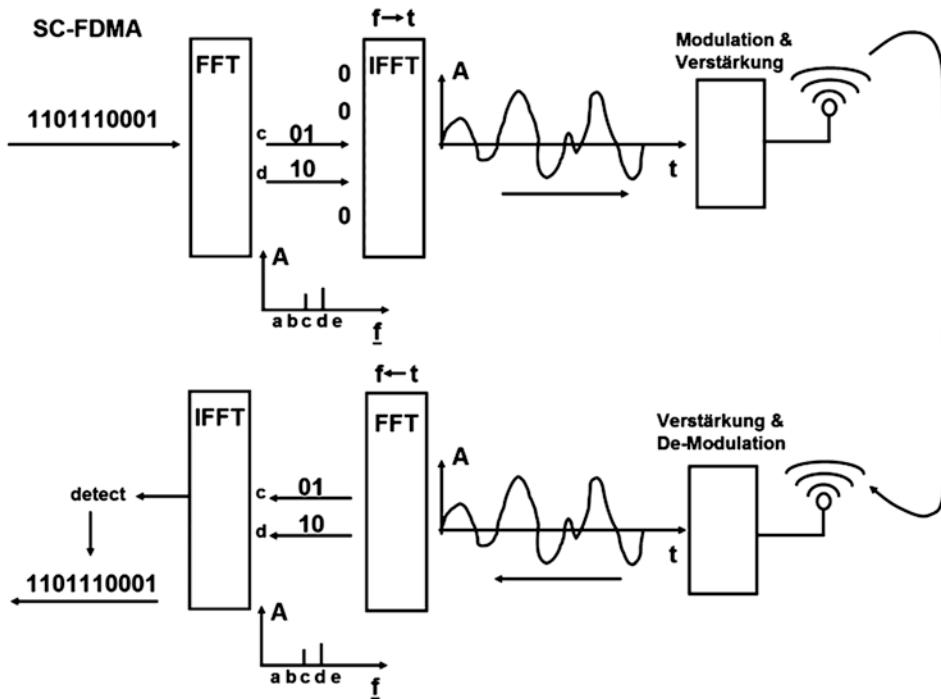


Abb. 1.9 SC-FDMA-Datenübertragung in Uplink-Richtung

Auf der Empfängerseite, dargestellt im unteren Teil von Abb. 1.9, wird das Signal zunächst demoduliert und dann in gleicher Weise wie bei OFDMA der FFT-Funktion zugeführt. Um das ursprüngliche Signal zu erhalten, wird das Ergebnis dann einer I-FFT-Funktion übergeben, die den ersten Umwandlungsschritt der Senderseite wieder rückgängig macht.

Von den zusätzlichen Bearbeitungsschritten für SC-FDMA abgesehen, werden für die Datenübertragung die gleichen physikalischen Parameter verwendet, wie für OFDMA in der Downlink-Richtung.

1.3.3 Quadrature Amplitude Modulation für Subcarrier

Wie zuvor beschrieben, wird auf der LTE Luftschnittstelle OFDM verwendet, um viele langsame Datenströme (Subcarrier) parallel zu übertragen und somit eine sehr hohe Übertragungsgeschwindigkeit zu erreichen. Auf jedem Subcarrier können Daten je nach Signalqualität mit unterschiedlichen Modulationsarten wie 16-, 64- oder 256-QAM übertragen werden. QAM steht für Quadrature Amplitude Modulation und ist eine Modulationstechnik, mit der mehrere Bits pro Übertragungsschritt in eine Amplitude

eines Sinussignals und zusätzlich in eine Phasenverschiebung bezogen auf ein Referenzsignal kodiert werden. Die Bits werden also in zwei Dimensionen übertragen (Amplitude und Phasenverschiebung).

Aus mathematischer Sicht können diese zwei Dimensionen als komplexe Zahl mit einer I- und einer Q-Komponente beschrieben werden. Abb. 1.10 zeigt, wie Bits in zwei Dimensionen in einem kartesischen Koordinatensystem dargestellt werden. Jeder Punkt, der auf dem Gitter eine I- und eine Q-Amplitude besitzt, repräsentiert 4 Bits. Insgesamt gibt es 16 Kombinationen, weshalb man von einer 16-QAM Modulation spricht.

In einem zeitbasierten Signal werden diese Amplituden wie folgt dargestellt: Der Punkt, der die 4-Bit Kombination 1101 repräsentiert, hat eine I-Amplitude von +1 und auch eine Q-Amplitude von +1. In einem Zeitsignal wird dies durch eine Sinusschwingung dargestellt, deren Amplitude der Länge des Pfeils entspricht und eine Phasenverschiebung von 45 Grad bezogen auf ein nicht modifiziertes Signal aufweist.

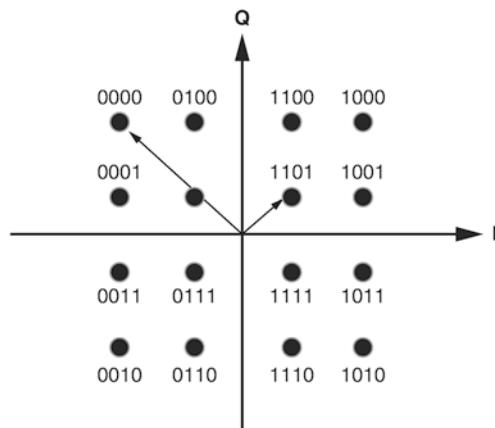
Der Punkt, der die 4-Bit Kombination 0000 repräsentiert, hat eine I-Amplitude von -3 und eine Q-Amplitude von +3. In einem Zeitsignal wird dies durch eine Sinusschwingung dargestellt, deren Amplitude der Länge des Pfeils zu diesem Punkt entspricht und einer Phasenverschiebung von 135 Grad bezogen auf das Referenzsignal.

Mathematisch kann gezeigt werden, dass die Amplitude und Phase einer Sinusschwingung durch die Kombination zweier Sinusschwingungen erzeugt werden kann, die mit der gleichen Frequenz oszillieren. Der Unterschied zwischen den zwei Signalen ist, dass diese genau 90 Grad zueinander phasenverschoben sind. Das bedeutet, dass eine Sinuswelle durch den Nullpunkt auf der Zeitachse ein Viertel einer vollen Sinuswelle früher läuft, als die andere Welle. Mit einer solchen Anordnung kann nun die Phase und die Amplitude des gemischten Signals gesteuert werden, indem die Amplituden der zwei Eingangssignale verändert werden.

Die Amplitude eines der Eingangssignale repräsentiert dabei die I-Komponente des Signals und die andere die Q-Komponente. Mit anderen Worten können somit die I- und

Abb. 1.10 16-QAM

Modulation



Q-Komponenten aus Abb. 1.10 für jeden der dargestellten Punkte durch Veränderung von den zwei Amplituden der Eingangssignale erreicht werden. Auf der Empfängerseite wird die umgekehrte Prozedur durchgeführt, also die Amplitude und Phase des Eingangssignals in zwei Kanälen bearbeitet, von denen jeder jeweils die Amplitude eines der ursprünglichen Eingangssignale wiederherstellt.

Die I- und Q-Signale werden auch als Baseband Signale bezeichnet, da die I/Q Werte sich nicht mit der Frequenz des Carrier Signals ändern, also z. B. nicht mit 2600 MHz, sondern nur einmal pro Übertragungsschritt. Des Weiteren sind die I/Q Werte unabhängig von der Übertragungsfrequenz. Dies bedeutet, dass die gleichen Werte verwendet werden, um das resultierende Carrier Signal zu erzeugen, unabhängig davon, ob die Daten bei 800 MHz oder 2600 MHz übertragen werden. Wie nachfolgend gezeigt wird, dauert jeder Übertragungsschritt 66.667 µs. Somit wird ein neues I/Q Paar für jeden Subcarrier etwa 15.000 mal pro Sekunde erzeugt. Da die I/Q Werte vom digitalen Teil des Modemchips berechnet werden, wird das GSM/UMTS/LTE Modem auch als Baseband Prozessor bezeichnet, der unabhängig vom Applikationsprozessor ist, auf dem das Betriebssystem des Endgerätes, wie z. B. Android, ausgeführt wird.

1.3.4 Symbole, Slots, Radio Blocks und Frames

Für die Datenübertragung ist die LTE-Luftschnittstelle wie folgt organisiert: Die kleinste Übertragungseinheit auf jedem Subcarrier ist ein Übertragungsschritt mit einer Länge von 66,667 µs. Ein Übertragungsschritt wird auch als Symbol bezeichnet und mehrere Bits können in Abhängigkeit des Modulationsverfahrens in ein Symbol kodiert werden. Bei sehr guten Übertragungsbedingungen kann z. B. das 64 Quadrature Amplitude Modulation (64-QAM)-Verfahren verwendet werden, mit dem 6 Bits ($2^6 = 64$) in einem Symbol abgebildet werden. In 3GPP Release 12 wurde die 256-QAM Modulation hinzugefügt, um 8 Bits pro Symbol ($2^8 = 256$) übertragen zu können. Diese Modulation bietet sich vor allem für sehr kleine Zellen an, in denen Nutzer nahe am Zellmittelpunkt sind und somit ein starkes Signal mit wenig Rauschen empfangen können. Unter weniger guten Übertragungsbedingungen werden 16-QAM oder QPSK (Quadrature Phase Shift Keying) verwendet, um 4 oder 2 Bits pro Symbol zu übertragen. Ein Symbol wird auch als ein Resource Element (RE) bezeichnet.

Da der Overhead für die Zuweisung von einzelnen Symbolen an einen Nutzer zu groß wäre, werden diese nun in mehreren Schritten zu Gruppen zusammengefasst. Abb. 1.11 zeigt zunächst, wie 7 zusammenhängende Symbole auf 12 Subcarriern zu einem Resource Block (RB) zusammengefasst werden. Ein Resource Block belegt einen Slot mit einer Dauer von 0,5 ms.

Zwei Slots bilden dann einen Subframe mit einer Länge von einer Millisekunde. Dieser bildet in LTE die kleinste Zuweisungseinheit. Das bedeutet, dass der eNodeB jede Millisekunde einmal entscheidet, welche Endgeräte in einem Slot Daten senden und empfangen. Die Anzahl der parallelen Resource Blocks in jedem Subframe hängt von

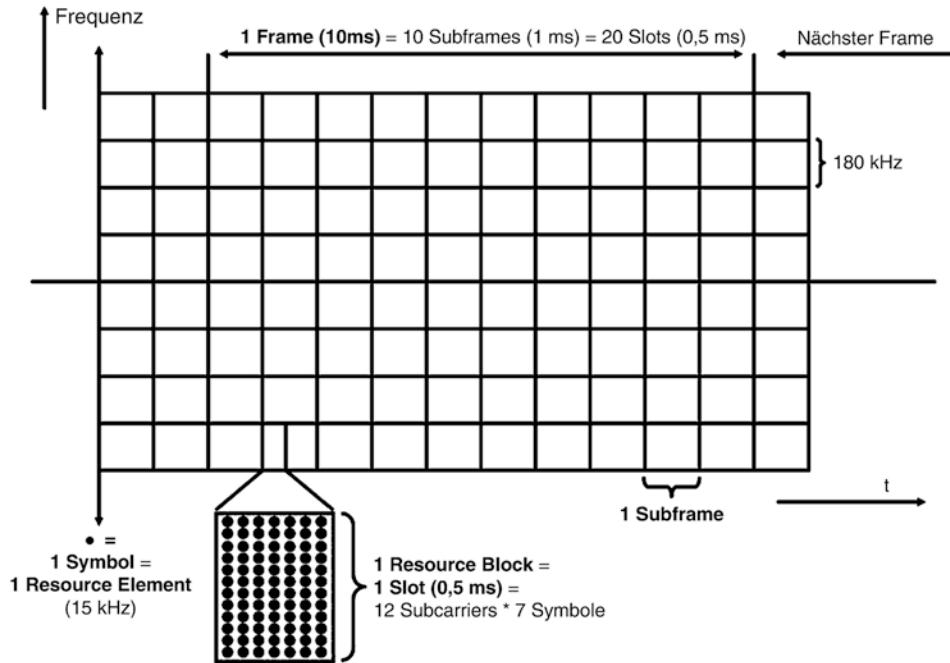


Abb. 1.11 Das LTE Resource Grid

der Bandbreite des Kanals ab. Ein 10 MHz breiter Kanal hat beispielsweise 600 Subcarrier. Da ein Resource Block 12 Subcarrier bündelt, gibt es somit 50 Resource Blocks, die einem oder mehreren Endgeräten pro Subframe zugewiesen werden können.

Für das Netzwerk gibt es zwei Möglichkeiten, einen Subframe zu übertragen. Die erste Variante sind Localized Virtual Resource Blocks (LVRBs), die zusammenhängend wie in Abb. 1.11 übertragen werden. Bei dieser Übertragungsmethode benötigt der eNodeB ein Narrowband Kanalfeedback vom Endgerät, um die RBs an die optimale Stelle des Kanals zu legen. Die zweite Übertragungsart sind Distributed Virtual Resource Blocks (DVRBs), mit der die Symbole eines Blocks über die gesamte Bandbreite des Kanals verteilt werden. Bei dieser Übertragungsart sendet das Endgerät entweder kein Feedback über die Empfangsqualität des Kanals zurück, oder nur ein Wideband Feedback über die gesamte Bandbreite.

Und schließlich werden 10 Subframes zu einem LTE Radio Frame mit einer Länge von 10 ms zusammengefasst. Frame-Abstände werden zum Beispiel für periodische System Information-Nachrichten verwendet, die im Anschluss besprochen werden. An dieser Stelle sollte noch erwähnt werden, dass Abb. 1.11 nur eine vereinfachte Darstellung ist, da ein LTE-Träger mehr als 8 Resource Blocks auf der Y-Achse verwendet. Auf einem 10 MHz-Kanal gibt es beispielsweise 50 parallele Resource Blocks.

1.3.5 Referenz- und Synchronisations-Signale

Wie bereits zuvor beschrieben, teilt das Netzwerk eine Anzahl von Resource Blocks pro Subframe einem bestimmten User zu, also einmal pro Millisekunde. Dabei gilt es jedoch zu berücksichtigen, dass nicht all Symbole für Nutzdaten verwendet werden können. Welche Symbole für andere Aufgaben verwendet werden, hängt vom Ort eines Resource Blocks im Resource Grid ab.

Damit Endgeräte ein LTE-Signal während einer Netzwerksuche finden und später die Signalqualität einschätzen können, werden Referenz-Symbole, die auch als Referenz-Signale bezeichnet werden, in einem vordefinierten Muster über die gesamte Kanalbandbreite verteilt. Referenzsignale werden in jedem siebten Symbol auf der Zeitachse und auf jedem sechsten Subcarrier auf der Frequenzachse eingefügt. Dies ist in Abb. 1.12 dargestellt und Details finden sich in 3GPP TS 36.211⁷. Insgesamt gibt es 504 unterschiedliche Referenz-Signal-Sequenzen, die dem Endgerät dabei helfen, die Signale von unterschiedlichen Basisstationen zu unterscheiden. Die Sequenzen werden auch als Physical Cell ID (PCI) bezeichnet. Damit benachbarte Basisstationen von einem Endgerät voneinander unterscheiden werden können, müssen diese unterschiedliche Referenzsignale verwenden. Um die Organisation der PCIs zu vereinfachen, wurden sechs PCI-Gruppen definiert, die jeweils um einen Subcarrier verschoben sind.

Für die erste Synchronisation werden zwei zusätzliche Signaltypen verwendet. Diese werden als Primary- und Secondary Synchronization Signals bezeichnet und werden in jedem ersten und sechsten Subframe auf den inneren 72 Subcarriern eines Kanals übertragen. In jedem dieser Subcarrier wird ein Symbol für jedes der beiden Synchronisations-Signale verwendet. Somit werden diese alle 5 ms übertragen.

1.3.6 Das LTE-Kanalmodel in Downlink-Richtung

Alle Signalisierungsnachrichten aus höheren Protokollsichten und der Nutzdatenverkehr werden auf der LTE-Luftschnittstelle in sogenannten Kanälen übertragen. Ursprünglich mit UMTS eingeführt, gibt es logische-, transport- und physikalische Kanäle, die in Abb. 1.13 für den Downlink dargestellt werden. Die unterschiedlichen Kanäle werden für unterschiedliche Arten von Daten auf der logischen Schicht ver-

Abb. 1.12 Referenz-Signale in Resource Blocks

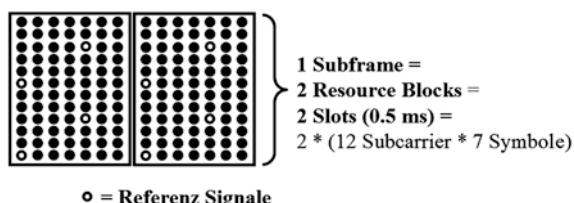
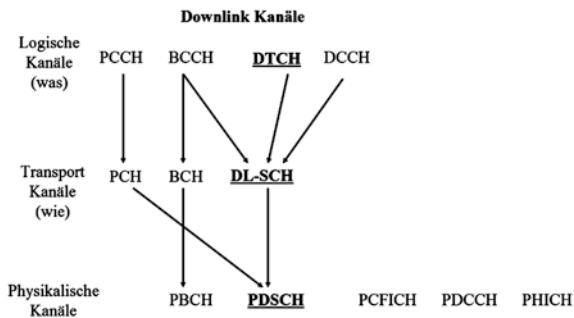


Abb. 1.13 LTE-Kanalstruktur
in Downlink-Richtung



wendet und trennen logische Datenströme von den Eigenschaften der physikalischen Datenübertragung.

Auf logischen Kanälen werden die Daten jedes Nutzers über einen eigenen logischen Dedicated Traffic Channel (DTCH) übertragen. Auf der Luftschnittstelle jedoch werden alle Dedicated Channels auf einen einzigen Shared Channel abgebildet, der alle Resource Blocks belegt. Wie zuvor beschrieben, werden einige Symbole in den Resource Blocks jedoch für andere Aufgaben verwendet. Welche Resource Blocks welchem Nutzer zugeteilt werden, entscheidet der Scheduler im eNodeB für jeden Subframe, d. h. einmal pro Millisekunde.

Die Abbildung von Dedicated Channels auf einen einzigen Shared Channel erfolgt in zwei Schritten. Zunächst werden die logischen DTCHs von allen Nutzern auf den Transport Layer Downlink Shared Channel (DL-SCH) abgebildet. In einem zweiten Schritt wird dieser Datenstrom dann auf den Physical Downlink Shared Channel (PDSCH) übertragen.

Transportkanäle können nicht nur Datenströme unterschiedlicher Nutzer multiplexen, sondern auch mehrere logische Kanäle eines einzelnen Nutzers, bevor der Datenstrom auf einen physikalischen Kanal abgebildet wird. Ein Beispiel: Ein Nutzer, dem ein dedizierter Übertragungskanal zugeordnet wurde, benötigt auch einen Kontrollkanal für das Management der Verbindung. Auf diesem Kanal werden z. B. Nachrichten für die Handover-Kontrolle, Messergebnisse von Nachbarzellenmessungen und Anweisungen für die Kanalmodifikation übertragen. Der DTCH und der DCCH werden somit auf den DL-SCH abgebildet und von dort auch zusammen weiter auf den PDSCH, also auf individuelle Resource Blocks. Außerdem werden auch spezifische Informationen der Zelle, die im logischen Broadcast Kanal (BCCH) übertragen werden, auf den Downlink Shared Transport-Kanal abgebildet, wie in Abb. 1.13 dargestellt ist.

Auch der Paging Control Channel (PCCH), der für die Benachrichtigung von Endgeräten über anstehende Datenpakete dient, während das Endgerät keine aktive Verbindung über die Luftschnittstelle aufgebaut hat, wird ebenfalls über den Shared Channel übertragen. Dazu wird der PCCH zunächst auf den Transport Layer Paging Channel (PCH) übertragen und dann ebenfalls auf den Physical Downlink Shared Channel (PDSCH).

Einzig eine geringe Anzahl von Systemparametern ist von der Übertragung auf dem Shared Channel ausgenommen. Diese Parameter werden von Endgeräten für die Synchronisation benötigt und auf dem Physical Broadcast Channel (PBCH) übertragen, der die ersten 3 Symbole auf 72 Subcarrieren (= 6 RBs) in der Mitte eines Kanals in jedem vierten Frame belegt. Das Übertragungsintervall für diese Informationen ist somit 40 ms und der PBCH folgt auf die Primary- und Secondary-Synchronization-Signale.

1.3.7 Downlink Management-Kanäle

Wie im vorigen Unterkapitel beschrieben, werden die meisten Kanäle auf einem einzigen Downlink Physical Shared Channel (PDSCH) übertragen, der alle Resource Blocks belegt. Somit ist ein Mechanismus notwendig, der einem Endgerät vermittelt, wann, wo und welche Art von Daten auf dem Shared Channel übertragen werden und welche Resource Blocks sie in Uplink-Richtung verwenden dürfen. Dies Aufgabe übernehmen Physical Downlink Control Channel (PDCCH)-Nachrichten.

Die Downlink Control-Informationen werden auf den ersten 1–4 Symbolen jedes Subframes auf der ganzen Kanalbandbreite übertragen. Die Anzahl der Symbole, die für diesen Zweck genutzt werden, wird über den Physical Control Format Indicator Channel (PCFICH) bekannt gegeben, der 16 Symbole belegt. Diese Flexibilität wurde eingeführt, damit sich das System dynamisch auf eine im Laufe der Zeit sich ändernde Anzahl von Nutzern einstellen kann.

Die Downlink Control-Nachrichten sind in sogenannten Control Channel Elements (CCEs) organisiert. Ein oder mehrere CCEs enthalten eine Signalisierungs-Nachricht, die entweder an ein Endgerät adressiert ist, oder, im Falle von Broadcast-Nachrichten, an alle Endgeräte in der Zelle. Um die für die Dekodierung der Nachrichten nötige Prozessorleistung und damit die Stromaufnahme zu senken, ist die Control Region in mehrere Bereiche, auch Search Spaces genannt, aufgeteilt. Ein Endgerät muss somit nicht alle CCEs dekodieren, um eine an sich adressierte Nachricht zu finden.

Schließlich gibt es auch noch Symbole, die für Rückmeldungen reserviert sind, ob ein Datenblock korrekt empfangen wurde oder nicht. Diese Funktion wird als Hybrid Automatic Retransmission reQuest (HARQ) bezeichnet, und der dafür notwendige Kontrollkanal wird Physical HARQ Indicator Channel (PHICH) genannt.

Zusammenfassend kann man sagen, dass der Physical Downlink Shared Channel (PDSCH) in allen Resource Blocks über die gesamte Kanalbandbreite übertragen wird. In den einzelnen Resource Blocks sind jedoch eine Anzahl von Symbolen für andere Aufgaben reserviert, wie z. B. Referenz-Signale, Synchronisierungs-Signale, den Broadcast-Kanal, den Kontroll-Kanal, den Physical Control Format Indicator Channel und den HARQ Indicator Channel. Die Anzahl der Symbole in einem Resource Block, die so belegt sind, hängen von der Position eines Blocks im Resource Grid ab. Für jeden Kanal oder Signal gibt es eine mathematische Formel, über die ein Endgerät berechnen kann, wo diese Informationen zu finden sind.

1.3.8 System Information Messages (SIBs)

Informationen, die für alle Endgeräte in einer Zelle wichtig sind, werden in System Information Broadcast (SIB) Nachrichten übertragen. Über den Broadcast Kanal wird jedoch nur der Master Information Block (MIB) übertragen. Alle anderen System Information Broadcast Nachrichten werden über den Physical Downlink Shared Channel (PDSCH) übertragen und deren Position im Resource Grid wird über den Physical Downlink Control Channel (PDCCH) bekannt gegeben.

Abb. 1.14 gibt einen Überblick über die in LTE verwendeten System Information Blocks (SIBs). Details finden sich in 3GPP TS 36.331⁸. Die wichtigsten Systeminformationen werden im Master Information Block übertragen, der alle 40 ms ausgestrahlt wird. Informationen über die aktuelle Zelle werden in SIB 1 übertragen, der alle 80 ms gesendet wird. Alle anderen SIBs werden in System Information Messages übertragen, deren Periode variabel ist. Welche SIBs mit welcher Periode übertragen werden, wird in SIB 1 übermittelt.

Nicht alle der in Abb. 1.14 gezeigten Broadcast Nachrichten werden in der Praxis in allen Netzwerken verwendet. Immer anzutreffen sind der MIB, sowie SIB 1 bis SIB 3. SIB 5 wird ebenfalls in der Praxis ausgestrahlt, da hier die Informationen über Zellen in anderen Frequenzbereichen übermittelt werden.

Message	Inhalt
MIB	Die wichtigsten Parameter für den initialen Zugriff.
SIB 1	Cell ID und für den Zugriff auf diese Zelle relevanten Parameter, so wie die Periodizität von anderen SIB Nachrichten.
SIB 2	Common und Shared Channel Konfigurationsparameter.
SIB 3	Parameter für Intra-Frequency Cell Reselection.
SIB 4	Intra-Frequency Neighbor Cell Reselection Parameter.
SIB 5	Inter-Frequency Neighbor Cell Reselection Parameter.
SIB 6	Informationen für Cell Reselection nach UMTS.
SIB 7	Informationen für Cell Reselection nach GSM.
SIB 10	Broadcast Informationen des japanischen Erdbeben und Tsunami Warnsystems (ETWS).
SIB 11	ETWS, Secondary Information.
SIB 12	Broadcast Information des Commercial Mobile Alert System (CMAS).
SIB 24	Informationen über 5G New Radio Standalone (SA) Nachbarzellen.

Abb. 1.14 System Information Blocks (SIBs) und deren Inhalt

1.3.9 Das LTE-Kanalmodell in der Uplink-Richtung

In der Uplink-Richtung gibt es ein ähnliches Kanalmodell wie in der Downlink-Richtung. Wiederum gibt es logische-, transport- und physikalische Kanäle, um logische Datenströme von der physischen Übertragung über die Luftschnittstelle zu trennen und mehrere Datenströme über einen Kanal zu übertragen. Wie in Abb. 1.15 gezeigt, ist der wichtigste Kanal der Physical Uplink Shared Channel (PUSCH). Seine Hauptaufgabe ist die Übertragung von Nutzdaten, Signalisierungsinformationen und eines Signal Quality Feedback.

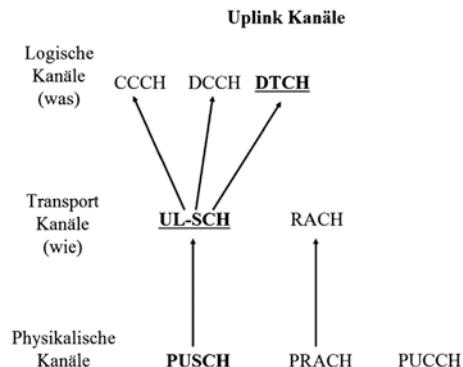
Daten, die über den PUSCH an der Basisstation eintreffen, werden in drei logische Kanäle aufgeteilt. Nutzdaten fließen in Dedicated Traffic Channels (DTCH). Zusätzlich gibt es noch Dedicated Control Channels (DCCH) für Higher Layer Signalisierungsinformationen. Während der Verbindungsauftnahme werden Signalisierungsinformationen über den Common Control Channel (CCCH) übertragen.

Bevor ein Endgerät Daten in Uplink-Richtung schicken darf, muss es sich zunächst mit dem Netzwerk synchronisieren und Ressourcen auf dem PUSCH beantragen. Dies geschieht mit einer Random Access Prozedur auf dem Physical Random Access Channel (PRACH). In den meisten Fällen weiß das Netzwerk nicht im Voraus, dass ein Endgerät eine Verbindung aufbauen will, und es ist somit möglich, dass dies mehrere Endgeräte gleichzeitig mit den gleichen RACH-Parametern versuchen. In diesen Fällen wird eine sogenannte Contention-based RACH-Prozedur durchgeführt und somit sichergestellt, dass entweder nur ein Endgerät Ressourcen auf dem Uplink Shared Channel bekommt, oder beide Requests fehlschlagen und wiederholt werden müssen.

1.3.10 Multiple Input Multiple Output Übertragungen

Zusätzlich zu effizienten Modulationsarten wie 64-Quadrature Amplitude Modulation (64-QAM) und 256-QAM, mit denen 6 Bits bzw. 8 Bits pro Übertragungsschritt kodiert

Abb. 1.15 Die LTE Uplink-Kanalstruktur



werden, spezifiziert 3GPP Release 8 auch die Verwendung von mehreren Antennen für die sogenannte Multiple Input Multiple Output (MIMO) Datenübertragung in Downlink-Richtung. Die grundsätzliche Idee bei MIMO ist, Daten gleichzeitig in mehreren unabhängigen Datenströmen über den gleichen Kanal über die Luftschnittstelle zu übertragen. In der Praxis werden heute zwei, sowie in manchen Fällen und mit geeigneten Endgeräten auch vier gleichzeitige Datenströme für den Shared Channel bei guten Übertragungsbedingungen verwendet. Für die anderen Kanäle wird nur ein einziger Datenstrom und eine sehr robuste Modulation und Kodierung verwendet, da der eNodeB sicherstellen muss, dass die übertragenen Informationen alle Endgeräte erreichen, unabhängig von deren aktueller Position und Übertragungsbedingungen.

Die gleichzeitige Übertragung von mehreren Datenströmen ist nur möglich, wenn die einzelnen Ströme weitergehend unabhängig auf dem Weg vom Sender zum Empfänger bleiben. Dies ist möglich, wenn folgende zwei Bedingungen erfüllt sind:

Auf der Senderseite sind zwei oder vier unabhängige Sendermodule für die Erzeugung der Datenströme notwendig. Zusätzlich wird für jeden Datenstrom eine eigene Antenne benötigt. Für zwei gleichzeitige Streams werden also zwei Antennen benötigt. In der Praxis werden dazu üblicherweise in einem einzigen Antennengehäuse zwei Antennen angebracht, die jeweils vertikal und horizontal polarisierte Signale aussenden.

Auch der MIMO-Empfänger benötigt zwei oder vier Antennen für die zwei oder vier unabhängigen Empfangswege. Dies ist vor allem für kleine mobile Endgeräte eine Herausforderung. Für Endgeräte wie Notebooks oder LTE Home Router ist es hingegen etwas einfacher, mehrere Antennen unterzubringen.

Die zweite Anforderung, die für MIMO erfüllt sein muss, sind möglichst unabhängige Übertragungswege für die unterschiedlichen Datenströme. Dies kann, wie in Abb. 1.16 gezeigt, erreicht werden, wenn die Signale unterschiedliche Wege zum Empfänger nehmen. Dies funktioniert am besten, wenn es keine direkte Sichtverbindung zwischen Sendern und Empfängern gibt. Abb. 1.16 ist jedoch eine Vereinfachung, da in den meisten Situationen die Signale sich gegenseitig bis zu einem gewissen Grad beeinflussen und somit die erreichbaren Datenraten geringer werden.

Theoretisch kann mit zwei unabhängigen Datenströmen der Durchsatz verdoppelt werden und mit vier Datenströmen sogar vervierfacht werden. In der Praxis fällt dies jedoch geringer aus, da die Signale interferieren. Wenn die Interferenz zu stark wird, kann man entweder das Modulationsverfahren ändern, also statt z. B. 64-QAM und MIMO gleichzeitig zu verwenden, wird das Modulationsverfahren auf 16-QAM reduziert. Es kann jedoch auch vorteilhaft sein, statt zwei Datenströme nur einen Datenstrom zu übertragen, diesen jedoch mit 64-QAM anstatt mit 16-QAM und MIMO. Da Modulation, Coding und MIMO jede Millisekunde angepasst werden können, ist es dem eNodeB somit möglich, sehr schnell auf sich ändernde Übertragungsbedingungen zu reagieren. In manchen Dokumenten finden sich auch die Begriffe 2×2 und 4×4 MIMO, um zu verdeutlichen, dass an beiden Enden der Verbindung zwei bzw. vier Antennen benötigt werden.

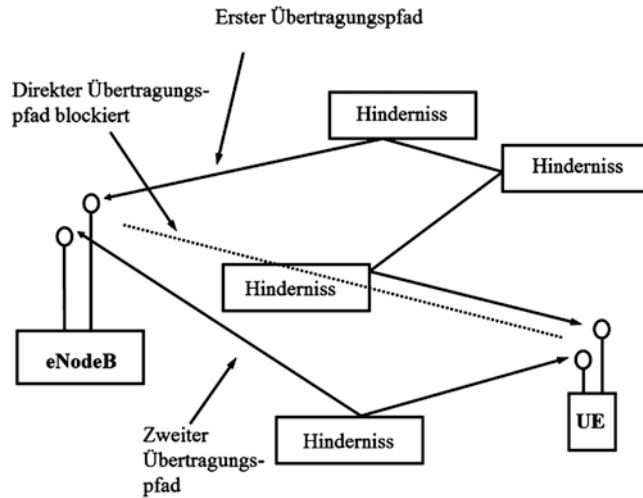


Abb. 1.16 Vereinfachtes Prinzip der MIMO-Übertragung

Im Uplink wurde in 3GPP Release 8 aus Sicht des Endgeräts nur eine Single Stream-Übertragung spezifiziert. In neueren Versionen des Standards wurde dann auch die Möglichkeit spezifiziert, dass ein eNodeB mehrere Endgeräte anweisen kann, im gleichen Resource Block zu senden. Da dies für die Endgeräte nicht sichtbar ist, wird hier auch von „Multi-User MIMO“ gesprochen. Um auf der eNodeB Seite die Signale unterschiedlicher Geräte auseinanderzuhalten, werden diese angewiesen, verschiedene Cyclic Shifts für ihre Referenzsignale zu verwenden. In der Praxis wird dies jedoch heute noch nicht verwendet. Außerdem wurde auch „Single-User MIMO“ spezifiziert, um es einem einzelnen Endgerät zu ermöglichen, mehrere Datenströme an den eNodeB zu senden. Auch von dieser Option wird in der Praxis noch kein Gebrauch gemacht.

In der Praxis gibt es zwei unterschiedliche MIMO Feedback Mechanismen für die Downlink Datenübertragung. Im ‚Closed Loop‘ Modus wird eine Precoding Matrix auf den Datenstrom angewendet. Die zwei oder vier ausgehenden Datenströme werden vor der Übertragung durch diese Matrix modifiziert, um die Ausbreitung der unterschiedlichen Datenströme in eine geeignete Richtung zu fokussieren. Das Endgerät kann dazu dem eNodeB mit einem Precoding Matrix Indicator (PMI) mitteilen, welche Precoding Matrix für die aktuelle Kanalsituation die Beste wäre. Mit dem Rank Indicator (RI) kann das Endgerät dem eNodeB zusätzlich mitteilen, wie viele Datenströme gleichzeitig übertragen werden sollen. Und schließlich sendet das Endgerät auch noch einen Channel Quality Indicator (CQI), aus dem der eNodeB entnehmen kann, welche Modulationsart (QPSK, 16-QAM, 64-QAM, 256-QAM) und welche Coding Rate verwendet werden soll. Die Coding Rate gibt das Verhältnis zwischen Nutzdatenbits und Fehlererkennungsbits an.

Für sich schnell bewegende Nutzer ist es schwierig, die Precoding Matrix schnell genug anzupassen. Für solche Situationen ist deswegen das Open-Loop MIMO Verfahren besser geeignet. Hier werden nur der RI und der CQI vom Endgerät an das Netzwerk übermittelt.

Und schließlich können mehrere Antennen auch für Sende- und Empfangs-Diversity verwendet werden. Hier wird der gleiche Datenstrom über mehrere Antennen mit einem jeweils eigenen Coding Scheme übertragen. Dies erhöht zwar die Übertragungsgeschwindigkeit nicht, hilft aber dem Empfänger das Signal zu dekodieren. Dies resultiert dann in Datenraten, die höher sind, als bei einer Übertragung mit nur einer Antenne. Wie in Abb. 1.17 gezeigt, gibt es eine große Anzahl Übertragungsmodi, die in 3GPP TS 36.213 spezifiziert sind⁹. Der eNodeB hat somit eine große Auswahl an Verfahren, um sich an wechselnde Übertragungsbedingungen anzupassen.

Im Uplink wird heute nur eine Single-Stream Datenübertragung verwendet, die schon in 3GPP Release 8 spezifiziert wurde. In manchen Netzwerken ist jedoch zu beobachten, dass zusätzlich zu 16-QAM und 64-QAM auch 256-QAM im Uplink unter sehr guten Übertragungsbedingungen verwendet wird.

1.3.11 HARQ und ARQ

Trotz adaptiver Modulation und Kodierung kann nicht verhindert werden, dass manche Datenpakete nicht korrekt empfangen werden können. Dies ist auch nicht wünschenswert, da dies bedeuten würde, dass die Kodierung zu robust gewählt wurde und somit die verfügbare Kapazität auf der Luftschnittstelle nicht ausgereizt wird. In der Praxis wird die Luftschnittstelle am effizientesten genutzt, wenn etwa 10 % der Pakete aufgrund von Übertragungsfehlern erneut übermittelt werden müssen. Die Herausforderung hierbei ist, Übertragungsfehler möglichst schnell zu erkennen und die fehlerhaften Daten dann möglichst schnell erneut zu übertragen. Außerdem muss der Scheduler in der Lage sein, Modulation und Kodierung schnell zu ändern, um die Fehlerrate konstant zu halten. In LTE kommt dafür das Hybrid Automatic Retransmission reQuest (HARQ) Verfahren auf dem MAC Layer zu Einsatz, sowie ARQ auf dem RLC Layer.

Transmission Mode	Übertragungsverfahren auf dem PDSCH
Mode 1	Single Antenna Port
Mode 2	Transmit Diversity
Mode 3	Transmit Diversity oder Large Delay CDD
Mode 4	Transmit Diversity oder Closed-Loop MIMO
Mode 5	Multiuser MIMO
Mode 6	Transmit Diversity oder Closed-Loop MIMO
Mode 7	Single Antenna Port 0 oder Port 5

Abb. 1.17 Spezifizierte LTE Übertragungsmodi (Transmission Modes)

HARQ auf dem MAC Layer

In der Downlink Richtung wird HARQ in LTE asynchron verwendet. Das bedeutet, dass fehlerhafte Daten nicht sofort neu übertragen werden müssen. Nach der Übertragung eines Datenblocks in einem 1 ms Subframe wartet der eNodeB auf ein Acknowledgement (ACK) des Endgeräts. Ein Negative Acknowledgement (NACK) wird gesendet, wenn das Endgerät den Block nicht korrekt empfangen hat. Das HARQ ACK/NACK wird entweder auf dem PUSCH oder auf dem PUCCH gesendet, wenn dem Endgerät gerade keine Ressourcen in Uplink Richtung zugeteilt sind. Das kann zum Beispiel der Fall sein, wenn vom Netzwerk viele Daten ankommen, während das Endgerät selber gerade keine Daten zu übertragen hat.

Empfängt der eNodeB ein ACK, werden die Daten, die in diesem Subframe übertragen werden, aus dem Sendepuffer gelöscht und der nächste Datenblock im Puffer kann gesendet werden. Wird ein NACK empfangen, wird der eNodeB versuchen, den entsprechenden Datenblock erneut zu übertragen. Diese Wiederholung kann sofort geschehen oder auch etwas aufgeschoben werden, wenn z. B. die Übertragungsbedingungen gerade schlecht sind.

Bevor ein Datenblock gesendet wird, werden Redundanzbits zum Datenstrom hinzugefügt. Mit diesen können Fehler erkannt werden und zu einem Teil sogar auch korrigiert werden. Wie viele Redundanzbits eingefügt werden, hängt von den Radiobedingungen und dem Scheduler ab. Bei guten Radiobedingungen werden die meisten Redundanzbits wieder aus dem Datenstrom entfernt, bevor dieser übertragen wird. Dieses Verfahren wird auch als „Puncturing“ bezeichnet. Tritt ein Übertragungsfehler auf und die Daten können mit den Redundanzbits nicht korrigiert werden, hat der eNodeB mehrere Möglichkeiten, die Daten erneut zu übertragen:

- Der Datenblock kann einfach erneut übertragen werden.
- Eine andere Redundancy Version (RV) des Datenstroms mit einem anderen Redundancy Bitmuster wird gesendet. Auf der Empfängerseite werden die Datenströme kombiniert und somit die Anzahl der Fehlererkennungs- und Korrekturbits erhöht.
- Das Netzwerk kann sich auch entscheiden, die Modulation und das Kodierungsverfahren zu ändern und die Daten damit erneut zu übertragen.

Einen Datenblock zu wiederholen benötigt Zeit, da zunächst der fehlerhafte Empfang erkannt und dann die Daten erneut übertragen werden müssen. In LTE wird das ACK/NACK für Downlink Übertragungen nach vier Subframes gesendet. Das gibt dem Empfänger Zeit, die Daten zu dekodieren und zu prüfen. Somit kann der Datenblock frühestens 5 Subframes nach der ursprünglichen Übertragung wiederholt werden. Der eNodeB kann die Übertragung falls nötig auch erst in einem späteren Subframe senden. In Abhängigkeit der Radiobedingungen, sowie der Modulation und der Kodierung kann es sein, dass ein Datenblock sogar mehrere Male wiederholt werden muss. Dies ist jedoch nicht gewünscht, da dann die Datenrate sehr stark reduziert ist. Im eNodeB

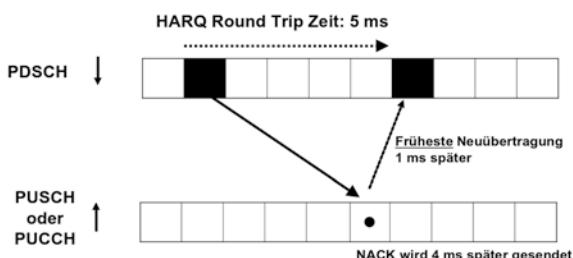
kann die maximale Anzahl der Wiederholungen eines Datenblocks konfiguriert werden. Wird diese Grenze erreicht, wird der Datenblock verworfen. In diesem Fall liegt es dann bei den höheren Schichten des Protokollstacks, den Datenverlust zu erkennen und gegebenenfalls zu korrigieren. Dies ist jedoch nicht für alle Datenströme wünschenswert. Für den VoLTE Sprachdienst ist es besser, Sprachdaten nicht erneut zu übertragen, da deutlich zu spät eintreffende Sprachpakete nicht mehr nutzbar sind. Aus diesem Grund können Sprachcodecs auch fehlende Daten bis zu einem gewissen Grad verschleieren.

Da eine erneute Übertragung erst nach 5 ms stattfinden kann, werden bis zu 8 HARQ Prozesse parallel ausgeführt, damit keine Lücken in der Übertragung entstehen können. Abb. 1.18 zeigt eine Datenübertragung mit 5 HARQ Prozessen. Werden Daten im Downlink übertragen, enthalten die Scheduling Grant Nachrichten im PDCCH unter anderem die Modulations- und Kodierungsart, die HARQ Prozessnummer und ob es sich um eine neue Übertragung oder eine Wiederholung handelt. Außerdem wird die Redundancy Version (RV) angegeben.

Im Uplink wird HARQ im synchronen Modus verwendet. Eine Wiederholung eines vom Netz nicht korrekt empfangenen Datenblocks findet also nach einer fixen Zeit statt. Wird ein Datenblock in einem 1 ms Subframe korrekt vom eNodeB empfangen, wird dies dem Endgerät 4 Subframes später bestätigt. Das ACK wird auf dem PHICH übertragen, der über eine Anzahl Symbole in den ersten zwei Symbolreihen jedes Subframes gesendet wird. Da mehrere Endgeräte ein Scheduling Grant für unterschiedlichen RBs in einem Subframe haben können, wird eine mathematische Funktion verwendet, die beschreibt, welche Symbole des PHICHs das Feedback für welches Endgerät enthält. Nachdem ein positives ACK vom Endgerät empfangen wurde, kann dann der nächsten Datenblock im Uplink übertragen werden.

Wenn das Netzwerk den Datenblock nicht korrekt empfangen hat, muss es eine Wiederholung anfordern. Dies kann auf zwei Arten geschehen. Die erste Möglichkeit ist, ein NACK zu senden und eine erneute Übertragung in einem neuen Format und an einer neuen Stelle im Resource Grid anzufordern. Die geschieht über ein Scheduling Grant auf dem PDCCH. Die zweite Möglichkeit ist, nur ein NACK ohne weitere Details über den PDCCH zu senden. In diesem Fall wiederholt das Endgerät die Übertragung an der gleichen Stelle im Resource Grid wie die ursprüngliche Übertragung.

Abb. 1.18 Synchrones HARQ in der Downlink Richtung



ARQ auf dem RLC Layer

Pakete, die trotz des HARQ Mechanismus verloren gehen, werden über die Automatic Retransmission Request (ARQ) Funktion auf dem nächst höheren Layer im Protokollstapel, dem Radio Link Control (RLC) Layer erkannt und erneut übertragen. Details hierzu finden sich in 3GPP TS 36.322¹⁰. Während ARQ für die meisten Bearer verwendet wird, gibt es eine Ausnahme: Für den VoLTE Dedicated Bearer wird ARQ deaktiviert, da zu spät eintreffende Sprachpakete vom Sprachcodec nicht mehr verwendet werden können.

Für alle anderen Bearer wird ARQ wie folgt verwendet:

- Sobald der Empfänger ein fehlendes Paket bemerkt, sendet er eine Nachricht zum Sender. Dieser wiederholt dann den fehlenden RLC Frame. Ein ‚sliding window‘ Ansatz sorgt dafür, dass ein fehlender Frame den Datenstrom nicht unterbricht. Fehlt ein RLC Frame, werden alle nachfolgenden RLC Frames beim Empfänger zwischen gespeichert und erst dann weitergesendet, wenn auch der fehlende RLC Frame empfangen wurde.
- Während der normalen Datenübertragung fordert der Sender vom Empfänger periodisch einen ARQ Status Report mit einem Polling Indicator Bit im RLC Header eines Datenframes an. Dies reduziert den Umfang des Feedbacks und stellt gleichzeitig sicher, dass keine RLC Fehlermeldung übersehen wird.

Auch auf höheren Schichten haben Protokolle weitere Mechanismen, um fehlende oder fehlerhafte Pakete zu erkennen. Im IP Protokoll gibt es beispielsweise ein CRC Feld, das von Routern im Sendepfad überprüft werden kann. Und schließlich verwendet auch das TCP Protokoll ähnlich wie der RLC Layer ein ‚sliding window‘ Verfahren, um verlorene TCP Pakete zu erkennen und vom Empfänger erneut anzufragen.

1.3.12 PDCP – Komprimierung und Verschlüsselung

Über den MAC und RLC-Layern die zuvor beschrieben wurden, liegt der Packet Data Convergence Protocol (PDCP) Layer. Seine Hauptaufgabe ist die Verschlüsselung von Nutzdaten und der Signalisierungsnachrichten vor der Übertragung über die Luftschnittstelle. Außerdem werden Signalisierungsnachrichten mit einem Integritätscode versehen, um gezielte Veränderungen von außen erkennen zu können.

Eine weitere jedoch optionale Aufgabe des PDCP-Layers ist die Komprimierung der IP Header. In Abhängigkeit der Größe der Nutzdaten in einem Paket wird ein mehr oder weniger großer Anteil der Kapazität auf der Luftschnittstelle für die verschiedenen Header des Protokollstacks benötigt. Dies ist vor allem bei VoIP-Paketen der Fall, die, um die Verzögerung zu minimieren, alle 20 ms gesendet werden, und deshalb entsprechend wenige Nutzdaten enthalten.

Effiziente Sprachcodecs wie z. B. WB-AMR erzeugen IP Pakete mit einer Länge von nur etwa 90 Bytes. Zwei Drittel davon werden für diverse Header (IP, UDP und RTP) benötigt und eine Komprimierung der Header ist somit sehr vorteilhaft. Im Unterschied dazu werden für andere Applikationen wie z. B. Web Browsing sehr viel größere Datenpakete verwendet, typisch sind Paketgrößen von über 1400 Bytes. Bei solchen Paketen ist eine Komprimierung der Header weniger wichtig. Aus diesem Grund wird Header Compression für solchen Datenverkehr in der Praxis bisher nicht eingesetzt.

Auf der LTE Luftschnittstelle wird der Robust Header Compression (RoHC) Algorithmus hauptsächlich für VoLTE Sprachdatenpakete verwendet. Spezifiziert ist das Protokoll in RFC 4995 und 5795¹¹. Die Idee von RoHC ist es, die Header Komprimierung nicht für alle IP Pakete in gleicher Weise durchzuführen, sondern diese in zusammenhängende Datenströme (Streams) zu gruppieren. IP Pakete in einem Stream haben gemeinsame Eigenschaften wie z. B. die gleichen IP Quell- und Zieladressen und die gleichen TCP oder UDP Portnummern. Für einen solchen IP Paketstream wird dann eines von mehreren RoHC Profilen für die Komprimierung verwendet. Für VoIP Pakete kommt eines der Profile aus RFC 5225¹² zum Einsatz, das die Header von IP, UDP und RTP (Real-time Transport Protocol für Sprache und andere Multimediadaten) komprimiert. Für andere Streams können andere Profile verwendet werden, die z. B. nur IP und TCP Header komprimieren.

Nachdem ein Profil für einen Stream ausgewählt wurde, wird eine RoHC Session gestartet. Mehrere unabhängige Sessions können gleichzeitig über einen Radio Bearer übertragen werden. Nach der RoHC Session Setup Phase werden statische Parameter wie IP Adressen, UDP Portnummern, etc., komplett ausgelassen und variable Parameter wie Zähler werden komprimiert. Eine Checksumme wird verwendet, um eine RoHC Session gegen Übertragungsfehler abzusichern.

Zusätzlich wird der PDCP Layer verwendet, um während eines Handovers die Nutzdaten sowie die Control-Plane korrekt zum eNodeB zu transferieren. Für Signalisierungs- und RLC Acknowledged Datenströme werden Datenpakete zwischengespeichert und wiederholt, sollten diese während des Handovers verloren gegangen sein. Dies wird anhand der PDCP Sequenznummer erkannt. Für RLC Unacknowledged Übertragungen (z. B. VoLTE Sprachpakete) findet dies nicht statt, da die Datenpakete aufgrund der Real Time Anforderungen durch die entstandene Verzögerung wahrscheinlich nicht mehr verwendet werden können.

1.3.13 Der LTE Protokoll Stack

In den vorherigen Unterkapiteln wurden die wichtigsten Funktionen der LTE Luftschnittstelle wie z. B. HARQ, ARQ und Verschlüsselung besprochen. Diese Funktionen sind über mehrere Layer des Protokollstacks verteilt. Abb. 1.19 zeigt, wie die Layer aufeinander aufbauen und wo die unterschiedlichen Funktionalitäten implementiert sind. Auf der vertikalen Achse ist der Protokollstack zweigeteilt. Auf der linken Seite von

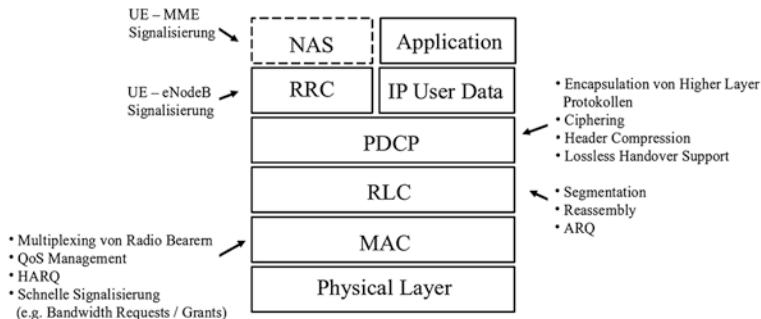


Abb. 1.19 Der Protokollstack der Luftschnittstelle und seine Funktionen

Abb. 1.19 ist die Control Plane dargestellt. Die oberste Schicht bildet das NAS Protokoll, das hauptsächlich für Mobility Management und andere Funktionen zwischen Endgerät und MME verwendet wird. NAS Messages werden über die Luftschnittstelle getunnelt, der eNodeB leitet diese Nachrichten transparent weiter. NAS Messages sind immer in Radio Resource Control (RRC) Nachrichten eingebettet, die über die Luftschnittstelle gesendet werden. Des Weiteren werden RRC Nachrichten für das Management der Luftschnittstelle verwendet, z. B. um einen Handover anzukündigen oder einen Bearer zu modifizieren. Eine RRC Nachricht muss deshalb nicht unbedingt eine NAS Message enthalten. Auf der User Plane, die in Abb. 1.19 auf der rechten Seite gezeigt wird, ist dies nicht der Fall. Hier enthalten IP Pakete immer Nutzdaten, die nur gesendet werden, wenn eine Applikation Daten übertragen will.

Die erste gemeinsame Protokollsicht, die IP, RRC und NAS Signalisierungsnachrichten überträgt, ist die PDCP Schicht. Wie schon im letzten Unterkapitel beschrieben, dient dieses Protokoll dazu, IP Pakete und Signalisierungsnachrichten einzupacken, dann zu verschlüsseln, Header Compression durchzuführen und für Handover ohne Paketverluste zu sorgen. Auf PDCP folgt eine Schicht darunter die RLC Schicht. Diese ist für die Aufteilung und das Zusammenfügen von Paketen aus höheren Schichten verantwortlich, damit diese über die Luftschnittstelle in kleineren Teilen übertragen werden können. Zusätzlich können auf diesem Layer fehlende Pakete erkannt und erneut übertragen werden. Danach folgt der MAC Layer. Diese Schicht dient dem Multiplexing von Daten aus verschiedenen Radio Bearern und steuert die Einhaltung von Quality of Service (QoS) Zielen, indem es der RLC Schicht mitteilt, wie viele Datenpakete bereitgestellt werden sollen. Zusätzlich ist der MAC Layer für HARQ Retransmissions zuständig. Und schließlich sind im MAC Header Felder für die Adressierung von Endgeräten und für Funktionen wie Bandwidth Requests, Grants, Power Management und Timing Advance Kontrolle vorhanden.

1.4 Scheduling

In LTE wird die Datenübertragung sowohl in Downlink- als auch in Uplink-Richtung vom Netzwerk kontrolliert. Dies geschieht in ähnlicher Weise wie auch bei anderen Technologien wie GSM und UMTS. Durch das vom eNodeB kontrollierte Scheduling ergeben sich eine Reihe von Vorteilen gegenüber Systemen wie Wireless LAN (Wi-Fi), die ein dezentrales Scheduling verwenden:

- Das Netzwerk kann schnell auf sich ändernde Bedingungen auf der Luftschnittstelle reagieren und somit den allgemeinen Durchsatz optimieren.
- Das Netzwerk kann die Quality of Service (QoS) Anforderungen eines Nutzers überwachen und steuern.
- Überlastsituationen können kontrolliert werden.

Andere Technologien wie Wireless LAN verwenden kein zentralisiertes Scheduling und überlassen es den Endgeräten und einem kooperativen Zugriffsmechanismus, die Luftschnittstelle zu belegen. Eine zentrale Kontrolle ist hier nicht unbedingt notwendig, da die Anzahl der Endgeräte pro Access Point sehr viel kleiner und das Abdeckungsgebiet des Netzwerkes deutlich geringer ist.

1.4.1 Downlink Scheduling

In der Downlink-Richtung ist der Scheduler in einem eNodeB dafür verantwortlich, alle Daten, die vom Netzwerk für die Benutzer einer Zelle angeliefert werden, über die Luftschnittstelle weiterzugeben. In der Praxis wird pro Endgerät meist ein einziger Default Bearer verwendet, über den die Daten von und zum Internet übertragen werden. Die Scheduling-Prozedur ist einfach, wenn sich nur ein Benutzer in einer Zelle befindet und weniger Daten vom Netzwerk angeliefert werden, als über die Luftschnittstelle übertragen werden können. Wenn der eNodeB aber mehrere Nutzer bzw. Bearer gleichzeitig bedient und die im Sendepuffer wartenden Daten die Kapazität der Luftschnittstelle übersteigen, muss der Scheduler entscheiden, welche Daten pro Subframe gesendet werden sollen und welche Nutzer wieviele Daten bekommen. Diese Entscheidung wird von einer Vielzahl unterschiedlicher Faktoren bestimmt.

Zunächst muss der Scheduler die Übertragung von Daten zur richtigen Zeit sicherstellen, falls für einen Bearer eine bestimmte Bandbreite und Verzögerung gewährleistet werden muss. Diese Daten werden dann anderen Daten auf der Luftschnittstelle bevorzugt. In der Praxis ist dies jedoch bislang eher unüblich, da die meisten Datenströme die gleiche Priorität haben.

Für Bearer mit der gleichen Priorität können andere Faktoren die Entscheidungen des Schedulers beeinflussen, welcher Nutzer wie viele Resource Blocks pro Zeiteinheit bekommt. Würde jeder Bearer mit der gleichen Priorität auch gleich behandelt

werden, würde unter Umständen viel Bandbreite verschenkt werden, da bei diesem Ansatz eine überproportional hohe Anzahl an RBs verwendet werden müsste, um auch Endgeräte unter schlechten Radiobedingungen die gleiche Bandbreite zu gewähren, wie für Nutzer, die unter guten Bedingungen Daten viel schneller empfangen könnten. Das andere Extrem wäre, immer diejenigen Nutzer zu bevorzugen, die gerade gute Empfangsbedingungen haben, da dann die Datenraten für Nutzer mit schlechten Empfangsbedingungen extrem langsam wären. In der Praxis werden deshalb sogenannte Proportional Fair Scheduler verwendet, die auch die Empfangsbedingungen pro Endgerät über die Zeit in die Scheduling-Entscheidung einbeziehen und so eine Balance zwischen bester Auslastung der Zelle und Durchsatz pro Nutzer erreichen.

Scheduling in Downlink-Richtung für einen Nutzer wird wie folgt durchgeführt. Für jeden Subframe entscheidet der eNodeB, wieviele Nutzer Daten empfangen sollen und wieviele RBs jedem Nutzer zugeteilt werden. Das Endgerät berechnet im ersten Schritt, wo innerhalb der Control Region eigene Scheduling-Nachrichten zu finden sind. Um herauszufinden, ob eine empfangene Nachricht für sich bestimmt ist, berechnet das Endgerät eine Checksumme und vergleicht diese dann mit der am Ende der Nachricht angegebenen Checksumme. In die Berechnung der Checksumme geht auch die ID des Endgeräts ein. Somit kann zum einen ermittelt werden, ob die Nachricht korrekt empfangen wurde, und zum anderen, ob die Nachricht auch für das Endgerät bestimmt ist. Ist dies nicht der Fall, wird die Scheduling-Nachricht verworfen.

Die Länge einer PDCCH Scheduling-Nachricht ist variabel und hängt vom Inhalt ab. Um die Dekodierung zu vereinfachen, wurden eine Anzahl von PDCCH-Nachrichtenlängen definiert. Diese sind wie folgt strukturiert: Auf der untersten Schicht bilden 4 Symbole auf der Frequenzachse eine Resource Element Group (REG). 9 REGs bilden ein Control Channel Element (CCE). Diese werden dann weiter in PDCCH-Nachrichten zusammengefasst, welche dann 1, 2, 4 oder 8 CCEs umfassen können. Eine PDCCH-Nachricht, die beispielsweise aus 2 CCEs besteht, enthält 18 REGs und $18 \times 4 = 72$ Symbole. Somit belegt die Nachricht 72 Subcarrier. Mit QPSK Modulation hat die Nachricht dann eine Länge von 144 Bits. Die längste PDCCH-Nachricht hat mit 8 CCEs eine Länge von 576 Bits. Welche Länge verwendet wird, hängt auch von Modulation und Kodierung der PDCCH-Nachricht ab. Hat ein Endgerät gute Empfangsbedingungen, kann eine bessere Modulation und eine geringere Kodierung verwendet werden und es werden somit weniger CCEs für die Nachricht benötigt.

Eine PDCCH-Nachricht kann für mehrere Zwecke verwendet werden, und da sich deren Längen unterscheiden, wurden eine Anzahl unterschiedlicher Nachrichten-Typen definiert. In den Standards werden Nachrichten-Typen auch als Downlink Control Information (DCI)-Format bezeichnet. Abb. 1.20 gibt einen Überblick über die 10 definierten DCI-Formate.

Downlink Assignment-Nachrichten für ein Endgerät enthalten folgende Informationen:

Nachrichten-Typ (DCI-Format)	Inhalt
0	Uplink Scheduling Grants für den PUSCH
1	PDSCH Assignments mit einem Single Codeword.
1a	PDSCH Compact Downlink Assignment
1b	PDSCH Compact Downlink Assignment mit Precoding Vektor
1c	PDSCH Assignments im 'Very Compact' Format
1d	PDSCH Assignments für Multi-User MIMO
2	PDSCH Assignments für Closed-Loop MIMO
2a	PDSCH Assignments für Open-Loop MIMO
3	Uplink Sendeleistungskontrolle mit 2-bit Power Adjustments
3a	Uplink Sendeleistungskontrolle mit 1-bit Power Adjustments

Abb. 1.20 Downlink Control Channel-Nachrichten-Typen (DCI-Formate)

- Der Typ der Resource Allocation (siehe unten)
- Power Control Informationen
- HARQ Informationen (new data bit, redundancy version)
- Modulations- und Kodierungsschema
- Anzahl der MIMO Streams
- Precoding-Informationen (d. h. wie die Daten für die Übertragung vorbereitet werden sollen).

Die eigentliche Zuteilung von Resource Blocks kann wie folgt beschrieben sein:

- Type 0 Resource-Zuteilungen enthalten eine Bitmap über zugeteilte Resource Block-Gruppen. Für einen 10 MHz-Kanal besteht eine Gruppe aus 3 RBs. Für die 50

Resource Blocks eines 10 MHz-Kanals werden somit 17 Bits für die Bitmap benötigt. Für einen 20 MHz-Kanal werden 4 RBs zu einer Gruppe zusammengefasst und für 100 RBs ist somit eine Bitmap von 25 Bits notwendig.

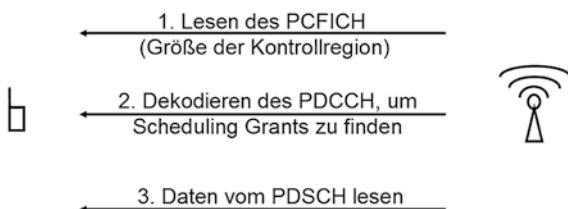
- Auch bei Type 1 Resource-Zuteilungen wird eine Bitmap verwendet. Anstatt jedoch ganze Gruppen mit einer „1“ in der Bitmap zuzuteilen, wird nur ein Resource Block pro Gruppe zugeteilt. Auf diesem Wege kann ein Resource Assignment über die gesamte Bandbreite verteilt werden.
- Type 2 Resource Allocations beinhalten einen Startpunkt auf der Frequenzachse und die Anzahl der zugeteilten RBs. Diese können dann entweder direkt aufeinander folgen oder über den gesamten Kanal verteilt sein.

Aus der Sicht des Endgerätes kann die Zuteilung, wie in Abb. 1.21 gezeigt, wie folgt zusammengefasst werden: Zu Beginn jedes Subframes dekodiert das Endgerät zuerst den Physical Control Format Indicator Channel (PCFICH) um herauszufinden, wie viele Symbole auf der Zeitachse für die Kontroll-Region, also für PDCCH-Nachrichten verwendet werden und sucht dann nach Nachrichten, die für sich bestimmt sind. Falls eine Downlink Assignment-Nachricht gefunden wurde, ermittelt es den Resource Allocation Type (Bitmap oder Startpunkt mit Länge). Mit den damit gewonnenen Informationen können nachfolgend alle Bereiche des Downlink Shared Channels dekodiert werden, um an die eigentlichen Nutzdaten zu gelangen.

In Abhängigkeit der aktuell benötigten Datenrate für ein Endgerät ist es meist nicht nötig, Daten für einen Nutzer in jedem Subframe zu übertragen. In diesem Fall ist es auch nicht nötig, dass das Endgerät in jedem Subframe nach Scheduling Grants sucht. Um die Leistungsaufnahme zu verringern, kann das Netzwerk sogenannte Discontinuous Reception (DRX)-Zeiten definieren, in denen das Endgerät den Kontrollkanal nicht abzuhören braucht.

Während das dynamische Scheduling ideal für kurze, jedoch unter Umständen sehr bandbreitenintensive Übertragungen wie Websurfen, Video Streaming und e-mails ideal ist, ist es für Real-Time Streaming Anwendungen wie Sprachtelefonie weniger geeignet. Bei Sprachtelefonie (VoLTE) werden kurze Datenbursts in regulären Intervallen übertragen. Falls die Datenrate, wie bei der Sprachtelefonie üblich, sehr gering ist, ist der Overhead für Scheduling Nachrichten sehr hoch, da pro Scheduling Nachricht nur sehr wenige Daten übertragen werden. Eine Lösung bietet das semi-persistent Scheduling (SPS). Statt eine Scheduling Nachricht über die Bereitstellung für Ressourcen auf der

Abb. 1.21 Empfang von Daten in der Downlink-Richtung



Luftschnittstelle für jeden Datenburst zu senden, wird ein periodisches Übertragungsmuster definiert. Da dies nur einmal gemacht wird, reduziert dies den Overhead beträchtlich. Während ein Teilnehmer nicht spricht, werden keine Sprachdatenpakete übertragen, sondern nur sogenannte Silence Description Information (SID) Datenpakete. In solchen Zeitabschnitten kann das Persistent Scheduling abgeschaltet werden. In der Uplink Richtung wird das Semi-Persistent Scheduling automatisch deaktiviert, wenn für eine vom Netzwerk festgelegte Zeit keine Datenpakete im Uplink gesendet werden. In Downlink Richtung wird das Semi-Persistent Scheduling mit einer RRC Nachricht beendet. Details hierzu finden sich in Abschn. 3.10 der LTE MAC Spezifikation in 3GPP TS 36.321¹³.

In der Praxis bleibt es dem Netzwerkhersteller und Netzbetreiber überlassen welche Scheduling Art für welche Daten verwendet wird. Die Herausforderung für das Netzwerk ist dabei herauszufinden, wann Semi-Persistent Scheduling verwendet werden kann. Für die VoLTE Sprachtelefonie des Netzbetreiber ist dies jedoch recht einfach möglich, da die Sprachdatenpakete über einen eigenen Bearer geleitet werden, der für ein Gespräch extra aufgebaut wird und es dem Radionetzwerk somit erlaubt, Sprachpakete einfach zu erkennen. Alle anderen Datenpakete vom und zum Nutzer werden über den normalen Default Bearer über die Luftschnittstelle geschickt, für den ein eNodeB das normale dynamische Scheduling verwendet. Internet basierte Sprachdienste wie z. B. Signal oder WhatsApp haben keine Möglichkeit, dem LTE Netzwerk mitzuteilen, dass seine Sprachdatenpakete eine spezielle Quality of Service Behandlung erhalten sollen. Somit werden deren Sprachdatenpakete über den normalen Default Bearer übertragen und somit in gleicher Weise wie andere IP Datenpakete behandelt. Dies kann vor allem in Überlastsituationen zu Aussetzern und Verzögerungen führen.

1.4.2 Uplink Scheduling

Um Daten in Uplink-Richtung auf dem Physical Uplink Shared Channel übertragen zu können, muss das Endgerät einen Scheduling Request an den eNodeB senden. Wenn aktuell keine physikalische Verbindung mit dem eNodeB besteht, muss dafür dann zunächst eine Verbindung hergestellt werden. Dies wird, wie bereits zuvor beschrieben, über eine RRC Connection Request-Nachricht auf dem Random Access Channel durchgeführt. Das Netzwerk baut dann eine Verbindung auf und vergibt Ressourcen auf dem Uplink Shared Channel, auf dem das Endgerät dann weitere Signalisierungsnachrichten und Nutzdaten senden kann. Uplink Ressourcen werden über den Physical Downlink Control Channel zugeteilt.

Während eine aktive Verbindung besteht und Ressourcen auf dem Uplink Shared Channel zugeteilt sind, sendet das Endgerät einen Buffer Status Report im Header jedes Paketes. Diese Information wird dann vom eNodeB verwendet, um in nachfolgenden Subframes Kapazität im Uplink bereitzustellen. Sind trotz aktiver Verbindung

keine Ressourcen auf dem Uplink Shared Channel zugewiesen, werden die Scheduling Requests über den Physical Uplink Control Channel übertragen.

Um den Uplink Shared Channel bestmöglich zu nutzen, muss der eNodeB wissen, wieviel Sendeleistung einem Endgerät im Vergleich zur aktuell verwendeten Sendeleistung noch zur Verfügung steht. Damit kann dann berechnet werden, mit welcher Modulation und mit welchem Kodierungsverfahren wie viele Resource Blocks benötigt werden. Zu diesem Zweck werden periodisch sogenannte Power Headroom Reports an den eNodeB gesendet.

1.5 Grundsätzliche Prozeduren

Nachdem nun die unterschiedlichen Referenzsignale und Kanäle der Luftschnittstelle bekannt sind, werden nun eine Anzahl Prozeduren beschrieben, die für die Kommunikation mit dem Netzwerk benötigt werden.

1.5.1 Netzwerksuche

Nach dem Einschalten eines Endgeräts ist die erste Aufgabe aus funktechnischer Sicht, nach einem geeigneten Netzwerk zu suchen und sich dann zu verbinden. Um diesen Prozess abzukürzen, kann das Endgerät die zuletzt verwendeten Parameter speichern und nach dem Einschalten dann versuchen, zur zuletzt verwendeten Zelle zurückzukehren. Dies ist z. B. möglich, wenn das Endgerät im ausgeschalteten Zustand an keinen anderen Ort transportiert wurde.

Falls die letzte verwendete Zelle nicht gefunden wurde, muss das Endgerät eine vollständige Netzwerksuche durchführen. Im ersten Schritt sucht ein Endgerät zunächst nach einem Signalpegel auf allen Kanälen und in allen Bändern und dann nach einem Primary Synchronisation Signal (PSS), das alle 5 ms ausgestrahlt wird. Wird ein PSS gefunden, bleibt das Endgerät auf dem Kanal und sucht nach dem Secondary Synchronisation Signal (SSS). Während der Inhalt des PSS konstant ist, ändert sich der Inhalt des SSS in jedem Frame, und das Endgerät kann auf diese Weise den Startzeitpunkt eines Frames ermitteln. Abb. 1.22 zeigt, wie Synchronisationssignale auf der Zeitachse gefunden werden.

Um die Netzwerksuche zu vereinfachen, werden das PSS und das SSS nur auf den inneren 1,25 MHz jedes Kanals übertragen, unabhängig von der gesamten Kanalbandbreite. Somit kann eine einfachere FFT Analyse verwendet werden, um das Signal zu finden.

Die Primary und Secondary Synchronization-Signale enthalten auch implizit die Physical Cell ID (PCI). Die PCI wird verwendet, um Nachbarzellen auf der gleichen Frequenz zu unterscheiden. In der Praxis ist das für Endgeräte dann wichtig, wenn sie sich zwischen mehreren Zellen befinden und mehrere PSS und SSS auf der gleichen

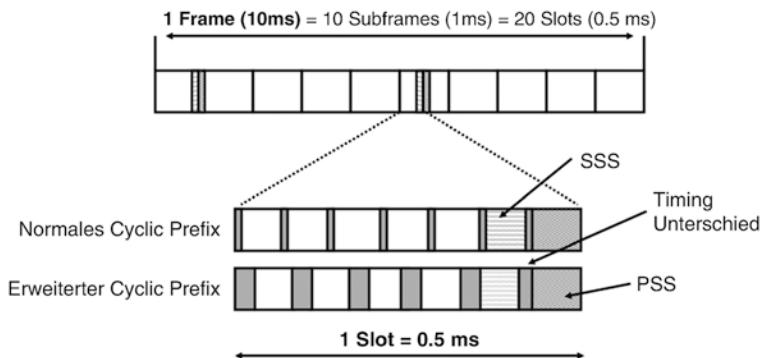


Abb. 1.22 PSS und SSS in einem LTE Frame

Frequenz empfangen können. Nachdem das PSS und SSS gefunden wurde, weiß das Endgerät auch, ob ein Normal- oder ein Extended Cyclic Prefix verwendet wird. Zusätzlich dazu kennt LTE auch eine von der PCI unabhängige Cell ID auf höheren Schichten. Diese dient zur eindeutigen Identifizierung der Zelle im gesamten Netzwerk.

Die Signale, die unterschiedliche Zellen auf dem gleichen Kanal senden, interferieren natürlich miteinander. Da abgesehen von nationalen Grenzgebieten eine Frequenz nur von einem Netzwerkbetreiber verwendet wird, versucht ein Endgerät immer automatisch, sich mit der Zelle zu verbinden, die auf einer Frequenz das stärkste Signal aussendet. Alle anderen Zellen auf dieser Frequenz werden in diesem Schritt noch ignoriert.

Falls das Endgerät direkt die Zelle gefunden hat, deren Parameter es vor dem letzten abschalten gespeichert hat, kann der Suchvorgang an dieser Stelle beendet werden, ohne zunächst nach weiteren, evtl. besser zu empfangenden Zellen zu suchen. Nach einer erfolgreichen Verbindungsprozedur (Attach) kann dann über den normalen Cell Reselection-Mechanismus zu einer besseren Zelle gewechselt werden.

Der nächste Schritt in der Suchprozedur ist das Dekodieren des Master Information Block (MIB), der alle 40 ms über den Physical Broadcast Channel (PCCH) in den inneren 1,25 MHz des Kanals übertragen wird. Der MIB enthält die wichtigsten Informationen über die Konfiguration des Kanals, die das Endgerät für den weiteren Verlauf benötigt. Damit diese Informationen auch unter sehr schlechten Radiobedingungen noch erfolgreich dekodiert werden können, wird der MIB mit einer sehr starken Kodierung und einer einfachen Modulation übertragen. Die erste Information, die das Endgerät dem MIB entnimmt, ist die Bandbreite des Kanals, da alle bisherigen Informationen lediglich in den inneren 1,25 MHz des Kanals vorhanden waren. Zusätzlich enthält der MIB auch die Struktur des HARQ Indicator Channel (PHICH, siehe oben) und die System Frame-Nummer (SFN), die z. B. für die Verschlüsselung und die Berechnung der Paging-Intervalle benötigt wird. Das Paging wird nachfolgend noch genauer beschrieben.

Mit den Informationen des MIB kann dann nach dem System Information Block 1 (SIB-1) gesucht werden. Dieser wird alle 80 ms auf dem Downlink Shared Channel übertragen, und das Endgerät muss deshalb auf dem Downlink Control Channel auf eine Nachricht warten, wann und wo der SIB-1 auf dem Shared Channel übertragen wird. Nachdem der SIB-1 gefunden wurde, erhält das Endgerät folgende Informationen:

- Den Mobile Country Code (MCC) und den Mobile Network Code (MNC) der Zelle. Diese Parameter werden vom Endgerät benötigt um herauszufinden, ob die Zelle zum Heimatnetzwerk gehört.
- Die NAS Cell Identity, ähnlich wie die Cell ID bei GSM und UMTS.
- Der Tracking Area Code (TAC), entspricht der Location Area in GSM.
- Die minimale Empfangsstärke, mit der die Zelle empfangen werden muss, damit eine Verbindung aufgebaut werden darf.
- Cell Barring Status, der das Endgerät informiert, ob die Zelle verwendet werden darf.
- Eine Liste und Periodizität von anderen System Information Blocks.

Mit den Informationen des SIB-1 kann ein Endgerät entscheiden, ob es mit der Zelle kommunizieren möchte. Dies ist z. B. der Fall, wenn die Zelle zum Heimatnetzwerk gehört. Das Endgerät sucht und dekodiert dann im nächsten Schritt SIB-2, der weitere Konfigurationsinformation über die Zelle enthält:

- Die Konfiguration des Random Access Channels.
- Die Konfiguration des Paging Channels.
- Die Konfiguration des Downlink Shared Channels.
- Die Konfiguration des Physical Uplink Control Channel.
- Ob ein Sounding Reference Signal (SRS) im Uplink gesendet werden soll, um dem eNodeB weitere Informationen über die Übertragungsbedingungen zu geben.
- Konfiguration für die Leistungsregelung im Uplink.
- Timer und Konstanten, die z. B. bestimmen, wie lange auf Antworten auf bestimmte Nachrichten gewartet werden soll.
- Die Bandbreite des Uplink Channels.

Weitere System Information Blocks enthalten Informationen, die vor allem für Zellwechsel wichtig sind, nachdem eine Verbindung mit dem Netzwerk aufgebaut wurde. Weitere Details hierzu später in diesem Kapitel.

Gehört die Zelle nicht zum Heimatnetzwerk und auch nicht zum zuletzt verwendeten Netzwerk, z. B. wenn der Benutzer sein Endgerät nach einem Flug ins Ausland wieder einschaltet, setzt das Endgerät die Suche nach geeigneten Zellen im aktuellen Frequenzband und auch in anderen Bändern fort. Wird ein Frequenzband von mehr als nur einer Radiotechnologie verwendet, wie z. B. das 900 MHz-Band für GSM und LTE, muss das Endgerät das Band nach den entsprechenden Technologien absuchen. In der Praxis dauert aus diesen Gründen die Netzwerksuche nach dem Einschalten im Ausland sehr

lange. Während früher ein GSM Endgerät, dass zudem nur zwei Bänder unterstützte, binnen weniger Sekunden eine Netzwerksuche durchführen konnte, brauchen aktuelle LTE Endgeräte, die sehr viele Radiotechnologien und Bänder unterstützen, deutlich länger.

1.5.2 Attach und Aktivierung des Default Bearers

Nachdem ein Endgerät alle Information hat, um erstmals mit dem Netzwerk nach dem Einschalten zu kommunizieren, führt es eine sogenannte Attach-Prozedur aus. Aus Sicht des Anwenders wird während dieser Prozedur dem Endgerät eine IP-Adresse zugewiesen, die für die Kommunikation mit dem Internet benötigt wird.

1.5.2.1 Die erste Kontaktaufnahme

Abb. 1.23 gibt einen Überblick über den ersten Teil der Attach-Prozedur wie in 3GPP TS 23.401¹⁴ beschrieben. Nachdem alle SIBs mit Konfigurationsinformationen über die Zelle ausgewertet wurden, fordert das Endgerät Ressourcen auf dem Uplink Shared Channel für den Attach-Prozess an. Dies geschieht über den Random Access Channel. Nach erfolgreichem Abschluss ist das Endgerät dem eNodeB bekannt, und dem Endgerät

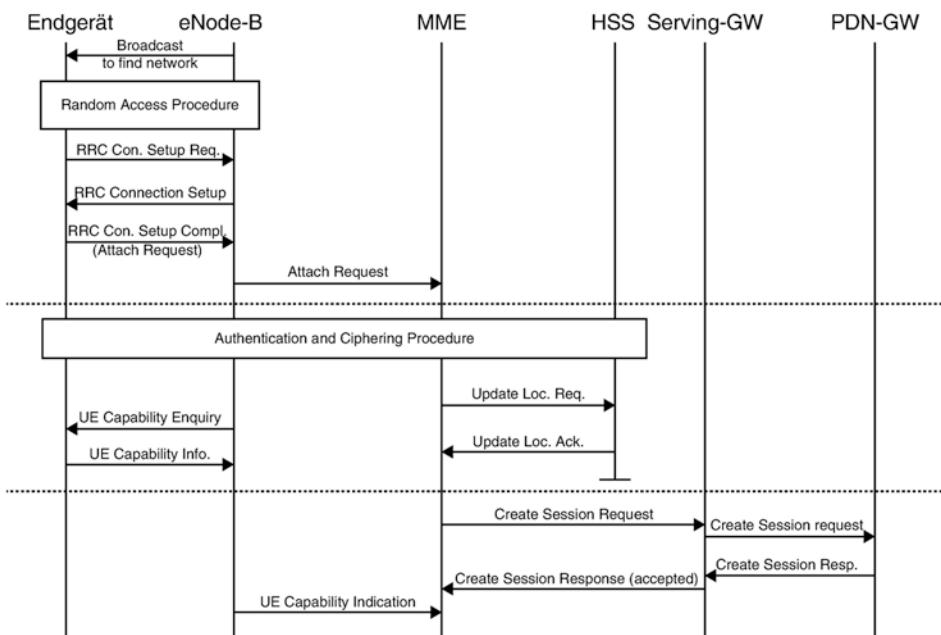


Abb. 1.23 Attach und Default Bearer Activation – Teil 1

wurde eine Cell – Radio Network Temporary ID (C-RNTI) zugeteilt. Diese MAC Layer ID wird z. B. bei der Zuteilung von Ressourcen auf dem PDCCH verwendet.

Im nächsten Schritt wird eine Radio Resource Control (RRC) Verbindung hergestellt, damit Signalisierungsnachrichten mit dem eNodeB und dem Kernnetzwerk ausgetauscht werden können. Dazu sendet das Endgerät eine **RRC Connection Request** Nachricht an das Netzwerk. Die Nachricht enthält unter anderem den Grund für die Verbindungsaufnahme und die temporäre NAS Identity des Endgerätes für das Kernnetz. Diese wurde s-TMSI (Service Architecture Evolution Temporary Mobile Subscriber Identity) genannt.

Wird der Zugriff auf das Netzwerk gewährt, was außer in Überlastsituationen immer der Fall ist, antwortet der eNodeB mit einer **RRC Connection Setup**-Nachricht, die alle Parameter enthält, um einen Dedicated Radio Signaling Bearer (SRB-1) aufzubauen. Über diesen werden dann alle weiteren RRC-Nachrichten übertragen. Diese Nachrichten enthalten z. B. Parameter für die Konfiguration von Physical- und MAC Layer-Eigenschaften wie die Uplink Shared Channel-Konfiguration, Parameter für die Leistungsregelung im Uplink, ob und wie ein Sounding Reference Signal gesendet werden soll, welche Art von Scheduling Requests vom Endgerät gesendet werden sollen, etc. Außerdem wird der SRB-1 auch für Nachrichten an die Mobility Management Entity (MME) im Kernnetz (NAS Signaling) verwendet.

Im nächsten Schritt antwortet das Endgerät mit einer **RRC Connection Setup Complete**-Nachricht an den eNodeB. Im RRC-Teil der Nachricht informiert das Endgerät den eNodeB, mit welcher MME es zuletzt verbunden war. In LTE kann ein eNodeB mit mehr als einer MME kommunizieren, um z. B. die Signalisierungslast zu verteilen und um Redundanz bei Ausfällen zu haben. Falls keine Informationen über die vorherige MME vorhanden sind, wählt der eNodeB selber eine geeignete MME aus.

Die **RRC Connection Setup Complete**-Nachricht enthält weiterhin eine eingebettete NAS-Nachricht, die eigentliche **Attach Request**-Nachricht, die der eNodeB transparent an die MME weitersendet. Teil dieser Nachricht ist die Globally Unique Temporary ID, auch GUTI genannt. Mit dieser ID kann die MME die noch gespeicherten Teilnehmerinformation zur Authentifizierung in ihrem Cache finden, oder gegebenenfalls ermitteln, welche MME das Endgerät zuletzt verwaltet hat und diese dann über den neuen Aufenthaltsort des Endgeräts informieren.

Über den nun vorhandenen Signaling Radio Bearer (SRB) authentisieren sich als nächstes UE und Netzwerk gegenseitig. Somit ist gewährleistet, dass sich nicht nur das Netzwerk über die Identität des Teilnehmers sicher sein kann, sondern auch das Endgerät kann sicher sein, mit einem Netzwerk zu kommunizieren, dem die Authentifizierungsinformationen des Teilnehmers aus dem Home Subscriber Server (HSS) bekannt sind. Dies schließt für beide Seiten einen Man-in-the-Middle Angriff aus. Nach der Authentifizierung schickt die MME dann eine Security Mode Command Nachricht, um das Integrity Checking und optional eine Verschlüsselung aller Nachrichten zwischen der MME und dem Endgerät zu aktivieren. Über den Integritätscheck wird sichergestellt, dass Signalisierungsnachrichten zwischen Endgerät und MME von einem Angreifer nicht verfälscht werden können. Mit einer Security Command Complete Nachricht

bestätigt das Endgerät den Vorgang. Fortan sind alle Nachrichten zwischen UE und MME geschützt.

Nachdem das Endgerät authentifiziert wurde, teilt dies die MME dem HSS mit einer **Update Location Request**-Nachricht mit. Das HSS antwortet darauf mit einer **Update Location Acknowledge**-Nachricht.

Um auch Nutzdatenpakete und Signalisierungsdaten die nur zwischen UE und eNodeB ausgetauscht werden zu schützen, ist im weiteren Verlauf ein zweiter Security Command/Security Complete Prozess notwendig. Dieser findet statt mit der MME jedoch zwischen UE und eNodeB statt.

Weiterhin wird in Abb. 1.23 gezeigt, wie der eNodeB an dieser Stelle der Verbindungsaufnahmeprozedur mit einer **UE Capability Inquiry**-Nachricht die unterstützten Funktionalitäten des Endgerätes für die Luftschnittstelle abfragt. Das Endgerät antwortet mit einer **UE Capability Information**-Nachricht, die unter anderem die unterstützten Radio-Technologien (GSM, UMTS, CDMA, etc.) enthält, welche Frequenzbänder unterstützt werden, ob RoHC Header Komprimierung implementiert ist und Informationen über weitere optionale Funktionalitäten. Diese Informationen sind für den eNodeB wichtig, um später die Luftschnittstelle bestmöglich für das Endgerät zu konfigurieren und zu wissen, in welche anderen Frequenzbänder und anderen Radionetzwerke die Verbindung übergeben werden kann, wenn das Endgerät den LTE-Abdeckungsbereich verlässt. Diese Informationen werden auch an die MME weitergegeben.

1.5.2.2 Das Erzeugen einer Session

Nachdem die MME eine **Update Location Acknowledge**-Nachricht vom HSS bekommen hat, wird die eigentliche Prozedur für den Aufbau einer Datenverbindung gestartet. Dies geschieht durch eine **Create Session Request**-Nachricht, die die MME zu einem Serving-Gateway ihrer Wahl schickt. Für eine optimale Lastverteilung und aus Redundanzgründen kann eine MME mit mehreren Serving-GWs kommunizieren. Das Serving-GW wiederum leitet die Nachricht an ein PDN-GW weiter, das dann eine IP-Adresse für den Teilnehmer auswählt und dem Serving-GW mit einer **Create Session Response**-Nachricht antwortet.

Das Serving-GW leitet dann die Nachricht an die MME weiter, und der Tunnel für die IP-Pakete zwischen Serving-GW und PDN-GW ist somit aufgebaut. Dieser Tunnel wird benötigt, da sich der Aufenthaltsort des Endgeräts ändern kann. Durch das Tunneln der IP-Datenpakete kann das Routing im LTE-Netzwerk für den Benutzer somit jederzeit geändert werden, ohne dass dem Endgerät eine neue IP-Adresse zugeteilt werden muss. Das Erzeugen eines Tunnels wird in den Standards auch als Context Establishment bezeichnet.

1.5.2.3 Context Establishment im Radionetzwerk

Nachdem der Benutzerkontext im Kernnetz aufgebaut wurde, antwortet die MME auf die ursprüngliche Attach Request-Nachricht mit einer **Initial Context Setup Request**-Nach-

richt, die eine Attach Accept-Nachricht beinhaltet. Wie in Abb. 1.24 gezeigt, beginnt mit dieser Nachricht auf dem S1 Interface zwischen MME und eNodeB der Aufbau des Nutzdatentunnels zwischen dem eNodeB und dem Serving-Gateway. Ein wichtiger Parameter in der Nachricht ist die Tunnel Endpoint ID (TEID), die das Serving-Gateway für diese Verbindung benutzt.

Die letzte Verbindung, die nun noch aufgebaut werden muss, ist der Bearer für die IP-Pakete des Benutzers auf der Luftschnittstelle. Dies beginnt mit einer **RRC Connection Reconfiguration**-Nachricht des eNodeBs an das Endgerät. Am Anfang der Attach-Prozedur wurde bereits ein Signaling Radio Bearer (SRB-1) erzeugt. Mit der nun durchgeführten Rekonfiguration wird nun ein weiterer Signaling Radio Bearer für niedrig priorisierte Nachrichten erstellt, sowie ein Data Radio Bearer (DRB), über den die IP-Pakete des Benutzers übertragen werden. Die Nachricht enthält weiterhin zwei NAS-Nachrichten, die **Attach Accept**-Nachricht und eine **Activate Default Bearer Context Request**-Nachricht. Diese konfigurieren die höheren Protokollschichten des Radiostacks des Endgeräts. Zusätzlich wird mit diesen Nachrichten auch die IP-Adresse dem Endgerät zugeteilt und weitere Parameter wie die IP-Adresse des DNS Servers dem Endgerät mitgeteilt.

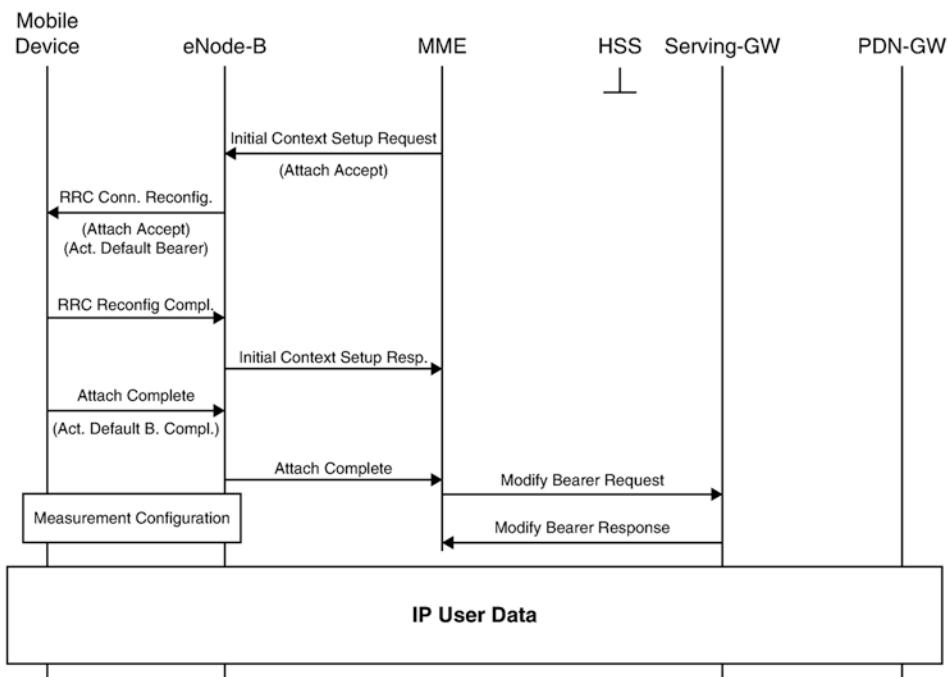


Abb. 1.24 Attach und Default Bearer Activation – Teil 2

Nachdem der RRC-Teil des Protokol-Stacks konfiguriert wurde, antwortet das Endgerät mit einer **RRC Connection Reconfiguration Complete**-Nachricht, und der eNodeB meldet das erfolgreiche Aufsetzen der Verbindung an die MME mit einer **Initial Context Setup Response**-Nachricht.

Nachdem das Endgerät auch den IP Stack des Endgeräts konfiguriert hat, sendet es eine **Attach Complete**-Nachricht an den eNodeB, die eine eingebettete **Activate Default Bearer Complete**-Nachricht enthält. Beide Nachrichten sind für die MME bestimmt, und der eNodeB leitet diese transparent weiter.

Der letzte Schritt der Attach-Prozedur ist der Aufbau des IP-Tunnels auf dem S1 Interface zwischen Serving-Gateway und dem eNodeB. Bis jetzt ist nur dem eNodeB die Tunnel Endpoint ID für diese Verbindung bekannt, da diese Information in der Initial Context Setup Request-Nachricht enthalten war. Zum jetzigen Zeitpunkt kann die MME nun auch dem Serving-Gateway die TEID des eNodeBs für diese Verbindung mit einer **Modify Bearer Request**-Nachricht mitteilen. Das Serving-Gateway speichert die TEID des eNodeBs für diesen Tunnel und kann nun alle IP-Pakete, die für diesen Nutzer vom Netzwerk angeliefert werden, über den richtigen Tunnel an die aktuelle eNodeB dieser Verbindung weiterleiten.

Ab diesem Zeitpunkt ist die Verbindung komplett aufgebaut und das Endgerät kann von nun an IP-Pakete über den eNodeB, das Serving-Gateway und das PDN-Gateway senden und empfangen. Trotz der vielen Nachrichten und Komplexität der Signalisierung wird die gesamte Prozedur in einem Bruchteil einer Sekunde durchgeführt.

An dieser Stelle sendet der eNodeB nun eine Reihe von RRC Reconfiguration-Nachrichten zum Endgerät, um Messungen und Reporting für Nachbarzellen zu konfigurieren, damit die Verbindung später an eine Nachbarzelle übergeben werden kann, wenn der Benutzer seinen Standort wechselt. Dies ist nicht Teil der Attach-Prozedur, sollte aber der Vollständigkeit halber hier trotzdem erwähnt werden.

1.5.3 Handover Szenarien

In Abhängigkeit von der Mess- und Reportkonfiguration, die das Endgerät vom Netzwerk empfangen hat, sucht es ab einer bestimmten Signalstärke nach Nachbarzellen. Bei Erreichen eines weiteren Grenzwertes meldet das Endgerät die gemessenen Werte an den eNodeB. Darauf basierend kann der eNodeB dann entscheiden, ob die Verbindung zu einem anderen eNodeB mit einem für das Endgerät besseren Empfangspegel weitergegeben werden soll.

In LTE gibt es zwei Arten von Handovern. Der effizienteste ist ein Handover, bei dem der aktuelle eNodeB mit dem neuen eNodeB direkt über das X2 Interface kommuniziert. Sollten die zwei eNodeBs keine direkte logische Verbindung haben, wird die Signalisierung über das S1 Interface abgewickelt, indem die MME den Handover Prozess unterstützt.

1.5.3.1 X2 Handover

Basierend auf Measurement Reports des Endgeräts kann sich ein eNodeB entscheiden, eine Verbindung zu einem neuen eNodeB weiterzugeben. Wie in Abb. 1.25 gezeigt, schickt der bisherige eNodeB dazu eine **Handover Request**-Nachricht an den neuen eNodeB. Diese enthält alle notwendigen Informationen über das Endgerät und die aktuelle Verbindung wie im Detail in 3GPP TS 36.423¹⁵ beschrieben ist. Der neue eNodeB überprüft zunächst, ob er noch genügend Bandbreite und Rechenleistung für die zusätzliche Verbindung zur Verfügung hat. Falls der eNodeB die Verbindung akzeptiert, wählt er eine neue Cell Radio Network Temporary ID (C-RNTI) aus und reserviert Ressourcen in Uplink-Richtung, damit das Endgerät während des Handovers eine Non-Contention Based Random Access-Prozedur ausführen kann, um sich mit der neuen Zelle zu synchronisieren. Dies ist notwendig, da das Endgerät noch keinen Timing Advance-Wert für die neue Zelle hat.

Nachdem die Vorbereitungen abgeschlossen sind, antwortet der neue eNodeB mit einem **Handover Request Acknowledge** über die X2-Schnittstelle. Diese Nachricht enthält alle Informationen, die das Endgerät über die neue Zelle benötigt. Da der Handover so schnell wie möglich durchgeführt werden soll, enthält diese Bestätigung unter

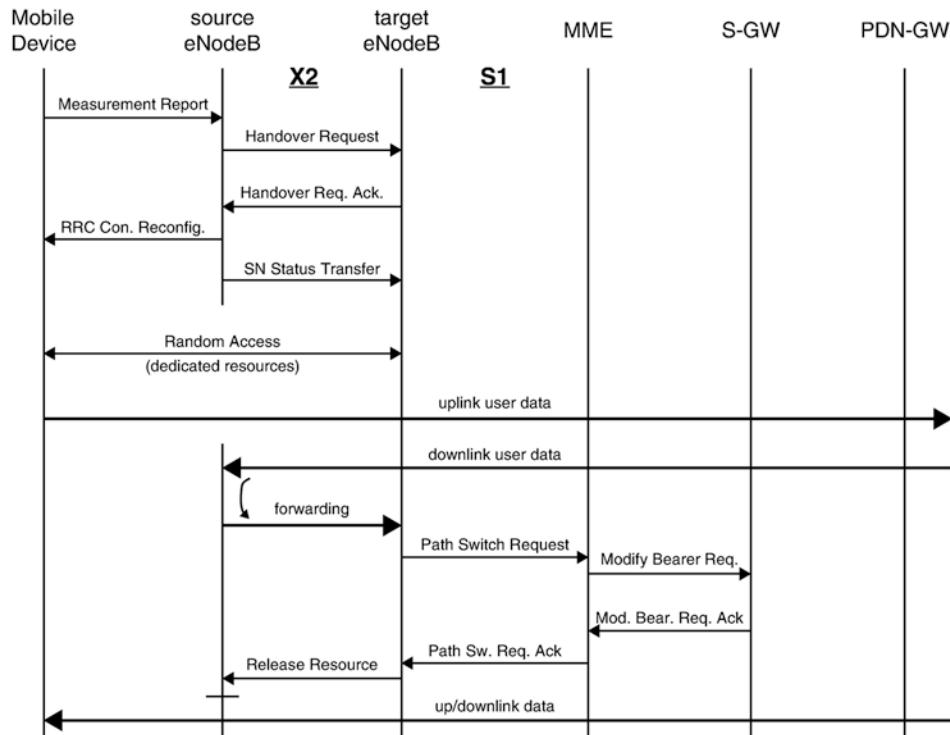


Abb. 1.25 X2 Handover

anderem die Physical Cell-ID (PCI) der neuen Zelle, die Bandbreite des Kanals, und die Konfiguration der Referenzsignale, des PHICH, etc.

Nachdem der bisherige eNodeB diese Bestätigung erhalten hat, sendet dieser sofort ein **Handover**-Kommando an das Endgerät, das danach seine Uplink Übertragungen einstellt. In Downlink-Richtung sendet auch der eNodeB fortan keine weiteren Datenpakete mehr. Sollten nach der Handover Anweisung Daten für das Endgerät aus dem Internet ankommen, werden diese über das X2 Interface an den neuen eNodeB weitergeleitet und von dort aus an das Endgerät geschickt, sobald der Handover Vorgang abgeschlossen ist. Außerdem sendet der bisherige eNodeB eine SN Status Transfer-Nachricht an den neuen eNodeB, die eine Uplink Sequenz Nummer enthält, damit evtl. verloren gegangene Datenpakete im Uplink erkannt und wiederholt werden können.

Bei LTE gibt es kein eigenes Handover-Kommando auf der Luftschnittstelle. Stattdessen wird eine **RRC Connection Reconfiguration**-Nachricht an das Endgerät geschickt, die alle notwendigen Parameter enthält, damit das Endgerät den Wechsel zum neuen eNodeB durchführen kann. Nach Erhalt dieser Nachricht sendet das Endgerät keine weiteren Nutzdaten mehr zum eNodeB.

Da das Endgerät schon zuvor Messungen durchgeführt hat, muss es während des Handovers die neue Zelle nicht mehr suchen. Es kann somit direkt eine **Random Access** Preamble auf dem PRACH senden. Da hierfür dedizierte Ressourcen verwendet werden, ist eine Contention Resolution Prozedur somit nicht notwendig.

Die **Random Access Response**-Nachricht des neuen eNodeB beendet die Handover Prozedur aus Sicht des Endgerätes, und es kann nun sofort Daten in Uplink-Richtung senden.

Da der eNodeB die IP-Adresse des Serving-Gateways sowie die Tunnel Endpoint ID der Verbindung kennt, können die Uplink Daten auch direkt und somit ohne Umweg über den bisherigen eNodeB an das Serving-Gateway weitergegeben werden. Auch Daten in Downlink-Richtung, die noch über den Umweg vom Serving-Gateway über den bisherigen eNodeB zum neuen eNodeB weitergegeben werden, können nun zum Endgerät übertragen werden. Im Radio- und Kernnetz sind jedoch noch weitere Schritte notwendig, um den S1-Tunnel vom bisherigen eNodeB auf den neuen eNodeB zu übertragen. Abb. 1.25 zeigt die einfachste Variante, bei der die MME und das Serving-Gateway unverändert bleiben.

Die Änderung des S1 IP-Tunnels und der Context Update im MME beginnen mit einer **Path Switch Request**-Nachricht, die der neue eNodeB an die MME sendet. Die MME überprüft dann, ob der neue eNodeB weiterhin für diese Verbindung mit dem aktuellen Serving-Gateway kommunizieren soll, oder ob dieses geändert werden soll. Dies könnte dann der Fall sein, falls ein anderes Gateway physisch näher am neuen eNodeB liegt, oder um eine bessere Lastverteilung zu erreichen. Im aktuellen Beispiel bleibt jedoch das Serving-Gateway unverändert, und somit muss nur eine **Modify Bearer Request**-Nachricht an das Serving-Gateway gesendet werden, um ihm den neuen Tunnel Endpoint Identifier des neuen eNodeBs mitzuteilen. Das Serving-Gateway führt daraufhin die notwendigen Änderungen durch und antwortet mit einer **Modify**

Bearer Response-Nachricht an die MME. Die MME wiederum bestätigt die Änderung mit einer **Patch Switch Request Acknowledge**-Nachricht an den neuen eNodeB. Abschließend teilt dann der neue eNodeB dem bisherigen eNodeB mit einer Release Resource-Nachricht mit, dass der Handover komplett ist.

1.5.3.2 MME und Serving-Gateway-Änderungen

Bei dem zuvor beschriebenen Handover-Beispiel waren keine Änderungen im Kernnetz für die Verbindungen erforderlich. Dies ist jedoch unter Umständen in folgenden Situationen notwendig:

- Aus Load-Balancing und Kapazitätsgründen.
- Um die Verbindung zwischen Radionetzwerk und Kernnetzwerk der Verbindung zu optimieren.
- Der neue eNodeB ist in einer Tracking Area, die nicht von der bisherigen MME verwaltet wird.

In diesen Fällen werden während des Handovers weitere Prozeduren ausgeführt, um den Nutzerkontext und die IP-Tunnel an die neuen Netzwerkkomponenten zu übergeben. Dies ist jedoch für das Endgerät völlig transparent und erhöht nur die Zeit, die für den Handover benötigt wird. Weitere Details finden sich in 3GPP TS 23.401.

1.6 Mobility Management und Leistungsoptimierung

Nachdem nun die wichtigsten LTE-Prozeduren beschrieben wurden, gibt der folgende Abschnitt einen Überblick über das Mobilitätsmanagement (Mobility Management) und wie der Stromverbrauch im Endgerät so gering wie möglich gehalten werden kann.

1.6.1 Mobilitätsmanagement im RRC-Connected State

Während sich ein Endgerät im RRC Connected State befindet, ist es mit dem Netzwerk synchronisiert und kann somit Daten jederzeit senden oder empfangen. In diesem Zustand ist somit ein Nutzdatentunnel auf dem S1 Interface zwischen eNodeB und dem Serving-Gateway aufgebaut, sowie ein weiterer Tunnel zwischen dem Serving-Gateway und dem PDN-Gateway. Daten, die für das Endgerät bestimmt sind, können so unmittelbar zugestellt werden. Auch Daten des Endgeräts in Uplink-Richtung können unmittelbar gesendet werden, entweder über kontinuierlich zugeteilte Resource Blocks auf dem Uplink Shared Channel oder, in Zeiten mit geringerer Aktivität, nach Senden eines Scheduling Requests über den Uplink Control Channel. Außerdem misst das Endgerät ständig die Signalqualität der aktuellen Zelle sowie die Signalqualität der Nachbarzellen

und sendet diese periodisch oder bei Erreichen eines Grenzwertes an das Netzwerk. Das Netzwerk kann dann bei Bedarf eine Handoverprozedur einleiten.

Die folgende Liste gibt einen Überblick über die wichtigsten LTE Measurement Events, die während des Verbindungsbaus oder später, wenn der Übertragungskanal schlechter wird, vom Netzwerk im Endgerät konfiguriert werden:

Für Handover von LTE nach LTE und für das Management von Carrier Aggregation:

- Event A1: Die aktuelle Zelle (Serving Cell) wird besser empfangen als ein konfigurierter Grenzwert (Threshold).
- Event A2: Die aktuelle Zelle wird schlechter empfangen als ein konfigurierter Grenzwert.
- Event A3: Eine Nachbarzelle wird besser empfangen als die aktuelle Zelle.
- Event A4: Eine Nachbarzelle wird besser empfangen als ein konfigurierter Grenzwert.
- Event A5: Der Empfang der aktuellen Zelle ist schlechter als der konfigurierte Grenzwert 1, eine Nachbarzelle ist besser zu empfangen als der konfigurierte Grenzwert 2.
- Event A6: Eine Nachbarzelle wird mit einem Offset besser empfangen als eine Secondary Cell (für Handover in Kombination mit Carrier Aggregation)

Für Handover oder ein Release mit Redirect von LTE nach UMTS oder GSM:

- Event B1: Der Empfang einer inter-RAT Nachbarzelle wird besser als ein konfigurierter Grenzwert.
- Event B2: Der Empfang der aktuellen Zelle wird schlechter als Grenzwert 1 und der Empfang einer inter-RAT Nachbarzelle wird besser als Grenzwert 2

Typischerweise werden die Events B1 oder B2 vom Netzwerk konfiguriert, nachdem das Endgerät einen schlechten LTE Pegel mit Event A2 gemeldet hat und keine LTE Nachbarzelle mit einem ausreichend starken Signalpegel gefunden wurde.

Measurement Events werden im Endgerät durch den eNodeB über RRCConnectionReconfiguration Nachrichten konfiguriert. Eine Messkonfiguration besteht dabei aus drei Teilen:

Teil 1 – Measurement Objects: Wenn eine Radioverbindung bei guten Übertragungsbedingungen aufgebaut wird und das Endgerät ist im bevorzugten Frequenzband, wird üblicherweise nur das aktuell genutzte LTE Frequenzband als Measurement Objekt konfiguriert. Falls die Übertragungsbedingungen nicht ideal sind oder das Endgerät auf einem niedriger priorisierten LTE Frequenzband Kontakt zum Netzwerk aufnimmt, werden mehrere LTE Frequenzbänder für Messungen konfiguriert. Um auf anderen Bändern messen zu können, wird zusätzlich ein Muster für Sende- und Empfangspausen definiert, damit das Endgerät Zeit hat, seinen Empfänger auf andere Frequenzen zu schalten und dort nach Zellen zu suchen. Sollten dann auch dort keine LTE Zellen gefunden werden, kann das Netzwerk auch UMTS Frequenzbänder oder GSM Kanäle als Measurement Objects konfigurieren.

Teil 2 – Report Konfigurationen: Ein Report kann entweder als periodisch auftretend konfiguriert werden, oder als Event (A1–A5, B1, B2, etc.), also als einmaliger Report der gesendet wird, wenn eine Bedingung erfüllt ist. Events können auch so konfiguriert werden, dass sie erst periodisch Messergebnisse melden, nachdem eine bestimmte Bedingung erfüllt ist. Dies kann z. B. verwendet werden, wenn das Netzwerk nur Messergebnisse haben möchte, wenn der Signalpegel der aktuellen Zelle einen konfigurierbaren Grenzwert unterschreitet.

Teil 3 – Messungen: In diesem Teil werden Measurement Objekte und Measurement Report Konfigurationen kombiniert. Auf diese Weise wird festgelegt, für welche Zellen oder Frequenzbereiche welche Messungen ausgeführt werden sollen.

1.6.1.1 Discontinuous Reception (DRX) im RRC-Connected State

Das fortwährende Suchen nach Scheduling Grants in jedem Subframe einmal pro Millisekunde bedingt einen recht hohen Stromverbrauch und sollte deshalb vermieden werden, wenn die aktuell benötigte Bandbreite dies zulässt. Aus diesem Grund ist es möglich, die Verbindung so zu konfigurieren, dass in Zeiten verminderter Aktivität nicht in jedem Subframe nach Scheduling Grants gesucht wird. Dies wird auch als Discontinuous Reception (DRX) bezeichnet und funktioniert wie folgt:

Wenn das Netzwerk ein Endgerät für DRX konfiguriert, definiert es einen Countdown-Wert für einen Timer, der gestartet wird, nachdem ein Datenblock gesendet wurde. Nach jedem Datenblock wird der Timer erneut gestartet. Läuft der Timer ab und es wurden noch immer keine Daten gesendet, aktiviert das Endgerät den DRX-Modus mit einem optionalen kurzen DRX-Zyklus. Dies bedeutet, dass das Endgerät seinen Empfänger nur für kurze Zeit abschaltet und dann automatisch wieder aufwacht. Sollten in dieser Zeit Daten für das Endgerät aus dem Internet angekommen sein, können diese noch immer recht zügig zugestellt werden. Auch der kurze DRX-Zyklus hat einen Timer, und nachdem dieser abgelaufen ist und noch immer keine Daten empfangen wurden, startet das Endgerät automatisch einen langen DRX-Zyklus mit einer längeren DRX-Periode. Dies spart mehr Energie, erhöht jedoch auch die Latenzzeit, sollten in dieser Zeit Daten für das Endgerät aus dem Internet beim eNodeB eintreffen. Empfängt das Endgerät einen Scheduling Grant in Zeiten, in denen die Control Region gelesen wird, wird der DRX-Modus deaktiviert, und alle Timer starten von neuem. Abb. 1.26 zeigt diesen Mechanismus grafisch.

Während das Endgerät im DRX Zustand ist, muss es das Netzwerk weiterhin periodisch über die Signalqualität im Downlink informieren (Channel Quality Indication, CQI). Optional kann das Netzwerk auch periodische Kontrollpakete im Uplink anfordern (Sounding Reference Signals, SRS), um auch die Signalqualität in dieser Richtung überprüfen zu können. Außerdem muss das Endgerät auch regelmäßig Power Headroom Reports generieren, damit das Netzwerk weiterhin informiert ist, wie weit die Sendeleistung des Endgeräts, falls notwendig, noch erhört werden kann. All diese Aktionen sind notwendig, damit Endgerät und Netzwerk beim Eintreffen neuer Nutzdatenpakete diese schnellstmöglich übertragen können.

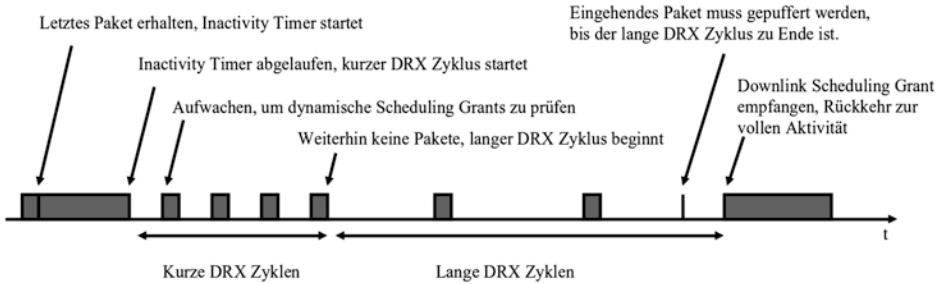


Abb. 1.26 Lange und kurze DRX-Zyklen

Um die Leistungsaufnahme im Endgerät zu reduzieren, kann das Netzwerk optional eine Timing Alignment Zeit definieren. Nachdem dieser Timer abgelaufen ist, muss das Endgerät keine weiteren Reports mehr schicken. Nachteil ist jedoch, dass beim Eintreffen neuer Nutzdaten zunächst wieder eine Timing Alignment Prozedur mit dem Netzwerk durchgeführt werden muss und sich die Übertragung dadurch etwas verzögert.

1.6.2 Mobility Management im Idle State

Werden für längere Zeit keine Daten mehr übertragen, ist es für das Netzwerk und für das Endgerät vorteilhaft, die Verbindung über die Luftschnittstelle in den RRC Idle-Zustand zu versetzen, da dies die Signalisierung und die Leistungsaufnahme reduziert. In diesem Zustand wechselt das Endgerät selber zwischen den Zellen, wenn die Empfangsstärken sich ändern. Das Netzwerk wird dabei nur kontaktiert, wenn sich die Tracking Area (TA) ändert. Als Konsequenz daraus ist die MME im Kernnetzwerk nur über die aktuelle Tracking Area, nicht jedoch über die genaue Zelle informiert, in der sich das Endgerät gerade befindet.

Im RRC Idle-Zustand existiert kein Nutzdatentunnel auf dem S1 Interface zwischen dem eNodeB und dem Serving-Gateway. Der Tunnel zwischen dem Serving-Gateway und dem PDN-Gateway bleibt jedoch auch in diesem Zustand erhalten. Aus logischer Sicht ist die Verbindung weiterhin vorhanden und alle logischen Bearer bleiben aufgebaut. Das bedeutet, dass die IP-Adresse des Endgeräts nicht verloren geht. Bei erneuter Aktivität muss dann die Verbindung über die Luftschnittstelle und der Tunnel im Radionetzwerk wieder aufgebaut werden und das Endgerät wechselt wieder in den RRC Connected State.

Kommt im RRC Idle State ein IP-Paket aus dem Internet für ein Endgerät an, kann es durch das LTE-Kernnetzwerk bis zum Serving-Gateway weitergeleitet werden. Da es jedoch von dort keinen S1-Tunnel zu einem eNodeB gibt, muss die MME zunächst beauftragt werden, das Endgerät zu benachrichtigen und den Tunnel neu aufzubauen. Da die MME nur die Tracking Area, nicht jedoch die genaue Zelle des Endgeräts kennt,

sendet es zunächst eine Paging-Nachricht an alle Zellen der Tracking Area. Die eNodeBs wiederum senden die Paging-Nachricht über die Luftschnittstelle.

Im RRC Idle-Zustand hat das Endgerät seinen Empfänger die meiste Zeit deaktiviert. Nur am Ende des sogenannten Paging Intervals, das üblicherweise 1 bis 2 s dauert, aktiviert es den Empfänger kurz, um Paging-Nachrichten zu empfangen. Dies wird auch Discontinuous Reception (DRX) im RRC Idle-Zustand genannt und ist nicht identisch mit dem DRX-Mode im RRC Connected-Zustand, der im vorherigen Unterkapitel beschrieben wurde.

Findet das Endgerät eine an sich adressierte Paging-Nachricht, baut es von sich aus eine Verbindung mit dem eNodeB mit einer Random Access-Prozedur auf. Der eNodeB, mit dem sich das Endgerät nun erneut verbindet, antwortet auf die Paging-Nachricht des MME, und die Verbindung auf der Luftschnittstelle und der S1-Tunnel zwischen eNodeB und Serving-Gateway werden wieder aufgebaut. Nachdem dies abgeschlossen ist, kontaktiert die MME das Serving-Gateway, das dann die wartenden IP-Pakete an das Endgerät weiterleiten kann.

An diesem Punkt startet der Prozess der Zustandswechsel von neuem. Die folgende Liste gibt noch einmal einen Überblick über die verschiedenen Aktivitätszustände:

- RRC-Connected State mit Suche nach Assignment Grants auf dem Kontrollkanal in jedem Subframe.
- RRC-Connected State mit Suche nach Assignment Grants auf dem Kontrollkanal mit einem kurzen DRX-Zyklus. Der Empfänger wird für kurze Intervalle abgeschaltet (Short DRX-Cycle).
- RRC-Connected State mit Suche nach Assignment Grants auf dem Kontrollkanal mit einem langen DRX-Zyklus. Der Empfänger wird für längere Intervalle abgeschaltet (Long DRX-Cycle).
- RRC-Connected State mit DRX, Time Alignment ausgelaufen, keine Übertragung von Statusinformationen im Uplink mehr.
- RRC-IDLE State, das Endgerät überprüft nur noch alle 1–2 s den Paging Kanal.

Im RRC-IDLE Zustand entscheidet das Endgerät selbstständig, wann es bei Pegeländerungen zu einer anderen Zelle wechselt. Die Parameter dazu werden dem Endgerät vom eNodeB über System Information (SI) Nachrichten mitgeteilt. Somit kann jeder eNodeB unterschiedliche Cell Reselection Kriterien ausstrahlen. Nach Wechsel in eine neue Zelle muss das Endgerät somit nicht nur prüfen, ob die Zelle in einer neuen Tracking Area ist, sondern muss auch die System Information Nachrichten lesen. Erst danach kann das Endgerät dann wieder den Empfänger abschalten und nur noch den Paging Kanal überwachen. Für den Zellwechsel sind folgende Parameter wichtig:

- Cell Barring Status in SIB 1: Gibt Auskunft darüber, ob eine Zelle benutzt werden darf, oder nicht. Eine Zelle kann z. B. im ‚Barred‘ Zustand sein, wenn gerade Wartungsarbeiten durchgeführt werden oder es ein technisches Problem gibt.

- Serving Cell Hysteresis in SIB 3: Gibt an, mit wieviel Dezibel (dB) die aktuelle Zelle gegenüber Nachbarzellen zu bevorzugen ist.
- Geschwindigkeitsabhängige Parameter in SIB 3: In Abhängigkeit ob das Endgerät stationär ist oder sich bewegt (z. B. in einem Auto, Zug, etc.) kann das Netzwerk unterschiedliche Cell Reselection Parameter konfigurieren. So ist es z. B. möglich, die Nachbarzellensuche noch bei einem recht hohen Signalpegel für sich in Bewegung befindliche Endgeräte zu starten und so dem Endgerät genug Zeit zu geben, nach Nachbarzellen zu suchen. Wenn sich das Endgerät nicht oder nur langsam bewegt, könnte eine andere Einstellung die Suche nach Nachbarzellen erst bei einem weit schlechteren Signalpegel starten und die Hysterese für den Zellwechsel höher gesetzt werden. Auf diese Weise kann Energie im Endgerät bei guten Signalpegeln gespart werden. An dieser Stelle sei angemerkt, dass geschwindigkeitsabhängige Messungen zwar sehr interessant erscheinen, in der Praxis bisher aber eher selten verwendet werden.
- Start von Intra-Frequenz Messungen in SIB 3: Definiert die Signalqualität der Serving Cell, bei deren Unterschreiten das Endgerät beginnen soll, nach Nachbarzellen zu suchen.
- Start von Inter-Frequency und Inter-RAT (Radio Access Technology) Messungen in SIB 3: Definiert die Signalqualität der Serving Cell, bei deren Unterschreitung das Endgerät zusätzlich auch noch nach Zellen auf anderen LTE Frequenzen und anderen RATs wie GSM und UMTS suchen soll. Üblicherweise wird hier ein niedrigerer Wert als für Intra-Frequency Messungen verwendet, da der Wechsel auf eine LTE Zelle im gleichen Band bevorzugt wird.
- Informationen über Nachbarzellen in SIB 4 bis 8: Diese System Information Nachrichten enthalten weitere Details über Nachbarzellen auf der gleichen Frequenz, auf anderen Frequenzen und über GSM und UMTS Zellen. SIB 4 mit Informationen über Zellen im gleichen Frequenzband ist optional. Falls SIB 4 nicht ausgestrahlt wird, führt das Endgerät eine ‚blinde‘ Suche durch.

1.6.3 Mobility Management und Zustandsänderungen in der Praxis

In der Praxis haben viele Faktoren Einfluss darauf, wie Netzbetreiber die LTE Luftschnittstelle zu den Endgeräten konfigurieren. Einerseits sind im Connected-Zustand ohne DRX die Verzögerungszeiten sehr gering und es wird keine Signalisierung zwischen den Basisstationen und dem Kernnetzwerk benötigt. Andererseits ist es sehr ineffizient, sich in diesem Zustand zu befinden, wenn keine Daten übertragen werden, da ständig Energie im Endgerät benötigt wird, um die Downlink Kontrollkanäle zu überwachen und ständig Kontrollinformationen und Signalisierung für die Leistungsregelung in Uplink Richtung gesendet wird. Dies wirkt sich sehr negativ auf die Batterielaufzeit aus. Auch für das Netzwerk hat dies Nachteile, da durch viele Endgeräte, die gleichzeitig

Signalisierungsinformationen senden, die Kapazität für Nutzdaten verringert wird. Es muss daher ein Kompromiss gefunden werden, wie lange ein Endgerät im Connected-Zustand ist und keine Daten sendet, bevor DRX aktiviert wird, und wie lange es danach dauert, bis das Netzwerk die Verbindung in den Idle-Zustand setzt. Die folgenden Beispiele zeigen, wie Netzwerke heute typischerweise konfiguriert sind:

Netzwerk 1:

- Zeit, bis DRX aktiviert wird: 100 ms
- DRX Short Cycle: 80 ms
- DRX Long Cycle: 200 ms
- On Duration: 10 ms
- Time Alignment: 10,2 s
- Time bis RRC-IDLE: –

Netzwerk 2:

- Zeit, bis DRX aktiviert wird: 200 ms
- DRX Short Cycle: 40 ms
- DRX Long Cycle 320 ms
- On Duration: 10 ms
- Time Alignment: unendlich
- Time bis RRC-IDLE: –

Die ersten zwei Netzwerke sind sehr ähnlich konfiguriert. Der DRX Modus wird nur nach einer Inaktivität eines Bruchteils einer Sekunde aktiviert und Endgeräte müssen dann nur noch auf Downlink Assignments für 10 ms in jedem Zyklus suchen. Beide Netzwerke haben sehr lange Time Alignment Timer, das erste über 10 s, während das zweite Netzwerk den Timer sogar auf unendlich gesetzt hat. Während diesen Zeiten muss das Endgerät Status und Messinformationen in Uplink Richtung schicken, was wiederum die Stromsparmöglichkeiten reduziert.

Netzwerk 3:

- Zeit, bis DRX aktiviert wird: 200 ms
- DRX Short Cycle: keiner
- DRX Long Cycle: 80 ms
- On Duration: 4 ms
- Time Alignment: 1,92 s
- Time bis RRC-IDLE: 30 s

Die Konfiguration von Netzwerk 3 weist wesentliche Unterschiede zu den zwei anderen Netzwerken auf. Während auch hier der DRX Modus nach einer Inaktivität von nur einem Bruchteil einer Sekunde aktiviert wird, ist der DRX Zyklus sehr viel kürzer als

bei Netzwerk 1 und 2. Auch die On-Duration ist sehr viel kürzer. Und schließlich wird auch das Time Alignment schon nach 1,92 s aufgeben, was eine wesentliche Leistungsreduzierung mit sich bringt. Nach 30 s ohne Datenübertragung wird das Endgerät dann in den RRC-Idle Zustand gesetzt.

Netzwerk 4:

- Kein DRX
- Time bis RRC-Idle: 5 s

Und schließlich gibt es auch Netzwerke in der Praxis, die keinen DRX konfiguriert haben. Stattdessen setzt das Netzwerk die Endgeräte in den RRC-Idle Zustand, nachdem nur 5 s keine Daten mehr übertragen wurden. Das bedeutet, dass typischerweise für jede zu übertragene Webseite ein neuer Kontext zur MME und dem S-GW aufgebaut werden muss. Aus Signalisierungs- und Lastsicht im Kernnetz ist eine solche Konfiguration sehr ineffizient.

1.7 LTE Sicherheitsarchitektur

Das LTE Sicherheitskonzept baut auf einen geheimen Schlüssel auf, der in der SIM Karte, sowie im Home Subscriber Server (HSS) im Netzwerk aufbewahrt wird. Da der gleiche Schlüssel für GSM, UMTS, LTE und 5G NR verwendet wird, ist es möglich, zwischen den unterschiedlichen Sicherheitssystemen der einzelnen Radionetzwerke zu wechseln.

Beim ersten Kontakt mit dem LTE Netzwerk, also während der zuvor beschriebenen Attach Prozedur, wird der Security Prozess zwischen Endgerät, MME und HSS durchgeführt. Während dieses Prozesses authentifiziert sich das Endgerät dem Netzwerk gegenüber und das Netzwerk authentifiziert sich gegenüber dem Endgerät. Dies verhindert Man-in-the-Middle Attacken, da sich ein Angreifer ohne Kenntnis des geheimen Schlüssels nicht für ein bestimmtes Endgerät oder als eine Basisstation eines Netzwerkes ausgeben kann. Auch die Authentifizierungsalgorithmen werden auf der SIM Karte und im HSS in einer geschützten Umgebung ausgeführt. Somit ist es nicht notwendig, den geheimen Schlüssel nach außen zu geben und es besteht somit keine Gefahr, dass Hacker den Schlüssel durch Abhören einer Schnittstelle zwischen SIM Karte und Endgerät oder zwischen HSS und MME abgreifen können.

SIM Karten müssen mindestens den UMTS Authentifizierungsalgorithmus beherrschen. SIM Karten, die nur die GSM Authentifizierung beherrschen und durch die Rückwärtskompatibilität zwischen UMTS und GSM auch mit UMTS Netzwerken funktionieren, können nicht weiterverwendet werden. Dies wurde ganz bewußt in Kauf genommen, da GSM SIM Karten keine Netzwerkauthentifizierung vornehmen können. Wird eine GSM SIM Karte in einem LTE fähigen Endgerät verwendet, und versucht sich dieses dann in einem LTE Netzwerk einzubuchen, kontaktiert die MME zunächst wie

üblich das HSS und fragt nach Authentication- und Ciphering Keys. Da das HSS diese Nachricht von einem LTE Netzwerknoten erhält, wird es diese Nachfrage abweisen, wenn der Eintrag des Nutzers im HSS nur GSM spezifische Sicherheitsparameter enthält. Die MME unterbricht dann den Attach Prozess und weist das Endgerät mit einem Reject Cause 15 (No Suitable Cells In This Tracking Area) an, eine andere Radionetzwerktechnologie zu finden und dort den Vorgang zu wiederholen.

Nachdem die Authentifizierungsprozedur abgeschlossen ist, werden auf Endgeräte- und Netzwerkseite Session Keys generiert. Diese werden verwendet, um alle NAS Nachrichten zwischen Endgerät und MME mit einer Prüfsumme zu versehen (Integrity Checking). Nachrichten können somit auf dem Transportweg nicht verfälscht werden. Außerdem kann optional auch die Verschlüsselung der Signalisierungspakete aktiviert werden.

Nachdem der eNodeB von der MME die Session Keys bekommen hat, aktiviert dann auch dieser das Integrity Checking und das Ciphering für alle RRC Signalisierungsdaten, sowie der Nutzdaten auf der Luftschnittstelle. Da die NAS Nachrichten zwischen Endgerät und eNodeB in RRC Nachrichten auf der Luftschnittstelle eingebettet sind, werden diese doppelt geschützt.

Für die Integritätsprüfung und die Verschlüsselung stehen unter LTE mehrere Algorithmen zur Verfügung. Verschlüsselungsalgorithmen werden EPS Encryption Algorithms genannt und als Abkürzung durchnummertiert (eea0, eea1, eea2, etc.). Eea0 ist ein Nullalgorithmus, die Verbindung wird dann nicht verschlüsselt. LTE Integrity Protection Algorithmen werden als EPS Integrity Algorithms bezeichnet und ebenfalls als Abkürzung durchnumeriert (eia1, eia2, etc.). Da die Integritätsprüfung immer durchgeführt werden muss, gibt es keinen eia0. Eea1 und eia1 implementieren den SNOW3G Algorithmus, der als UEA2/UIA2 mit 3GPP Release 7 auch in UMTS eingeführt wurde. Mit eea2/eia2 wird in LTE der AES (Advanced Encryption Standard) Algorithmus mit einem 128 Bit Schlüssel verwendet. Dieser wird im IT Bereich schon seit vielen Jahren für die Datenverschlüsselung verwendet.

1.8 Zusammenspiel mit UMTS und GSM

Am Rande des LTE-Abdeckungsbereiches soll ein Endgerät, falls noch vorhanden, automatisch in ein GSM oder UMTS Radionetz wechseln, um eine bestehende Verbindung aufrechtzuerhalten. Im ungünstigsten Fall verliert es dabei den Kontakt zum LTE-Netzwerk, weil es keine Zelle mit ausreichender Signalstärke mehr findet, und startet eine Suche nach LTE-Zellen auf anderen Kanälen und anderen Frequenzbändern. Ist dies nicht erfolgreich, wird auch nach UMTS und GSM-Zellen gesucht. Dies dauert jedoch typischerweise zwischen 10 und 30 s, während dessen das Endgerät keine Daten senden oder empfangen kann. Deshalb ist es besser, wenn das Netzwerk diesen Vorgang unterstützt. Die folgenden Prozeduren sind hierfür standardisiert und werden nachfolgend genauer beschrieben:

- Cell Reselection von LTE nach UMTS oder GSM.
- RRC Connection Release With Redirection von LTE zu UMTS oder GSM.

Unabhängig davon, ob das Endgerät ein GSM oder UMTS-Netzwerk selber findet oder vom Netzwerk Hilfe bekommt, ist es notwendig, dass das LTE-Kernnetzwerk mit dem GSM und UMTS-Kernnetzwerk verbunden ist, damit der Kontext der Verbindung, also z. B. die IP-Adresse, beim Übergang nicht verloren geht. Netzwerke verwenden dazu die Schnittstellen, die in 3GPP TS 23.401¹⁶ definiert sind und in Abb. 1.27 gezeigt werden. In der Praxis werden heute kombinierte 2G/3G/4G Kernnetzknoten verwendet und die gezeigten Interfaces sind somit nur logisch innerhalb des kombinierten Netzwerknotens vorhanden.

1.8.1 Cell Reselection zwischen LTE und GSM/UMTS

Die einfachste Art des Wechsels aus Signalisierungssicht zwischen LTE und GSM/UMTS ist eine Cell Reselection Prozedur im RRC Idle State. Um dies zu unterstützen, senden die eNodeBs die Konfigurationsparameter der GSM und UMTS-Nachbarzellen im Broadcastkanal aus. Sinkt dann die LTE-Signalstärke unter einen vom Netzwerk

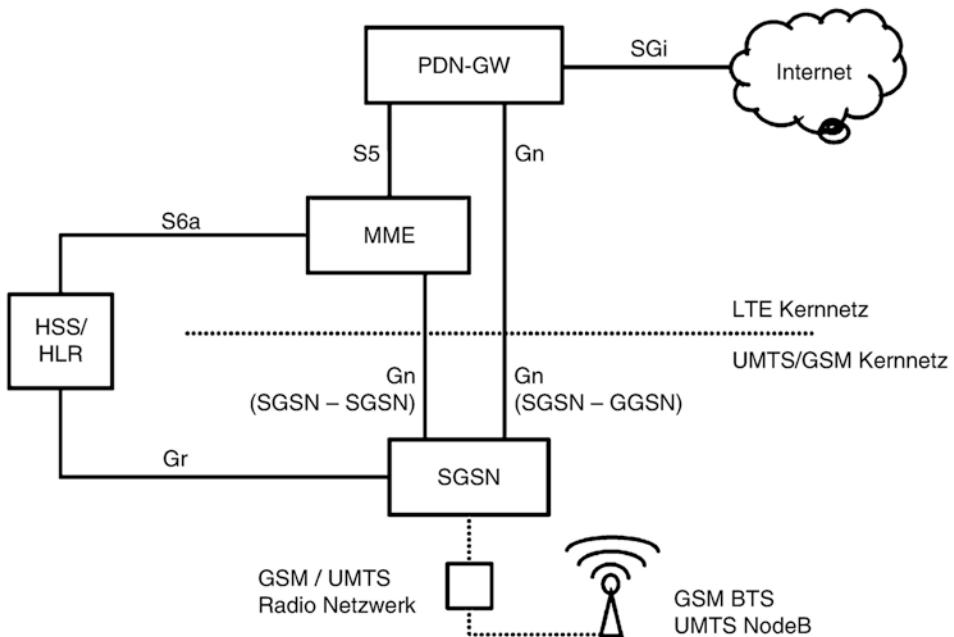


Abb. 1.27 Zusammenschaltung der LTE und GSM/UMTS-Kernnetzwerke

definierten Mindestwert, beginnt das Endgerät mit der Suche nach nicht-LTE-Zellen und wechselt selbständig dorthin, sobald die Empfangsstärke einen kritischen Wert unterschreitet.

Nach dem Wechsel zu einer GSM oder UMTS-Zelle startet das Endgerät dann ein Location Area Update mit dem leitungsvermittelnden Netzwerk und ein Routing Area Update mit dem paketvermittelnden Netzwerk. Mehr zu dieser Zweiteilung ist in den Kapiteln über GSM und GPRS am Ende des Buches beschrieben. Während dieser Prozeduren teilt das Endgerät den GSM bzw. UMTS Netzkomponenten mit, welches MME/S-GW zuletzt verwendet wurde. Da GSM und UMTS keine LTE Netzwerk-knoten kennen, gibt es in der LTE Netzarchitektur Schnittstellen, über die sich die MME und das P-GW wie die entsprechenden GSM oder UMTS Netzwerknoten ansprechen lassen. In GSM und UMTS sind dies die Serving- und Gateway GPRS Support Nodes (SGSNs, GGSNs).

Mit den Informationen des MMEs authentisiert das GSM oder UMTS Netzwerk dann das Endgerät und fordert die Modifikation des IP-Tunnels vom PDN-Gateway an. Schließlich wird der Routing Area Update mit einem Update des aktuellen Aufenthaltsorts des Endgeräts in der Teilnehmerdatenbank (HLR, HSS) abgeschlossen und das Endgerät kehrt in den Idle-Zustand zurück. Details zu dieser Prozedur werden in 3GPP TS 23.060¹⁷ beschrieben.

Wenn ein LTE-Endgerät sich im GSM oder UMTS-Netzwerk befindet, sollte es zu LTE zurückkehren, sobald dieses wieder verfügbar ist. Auch dies wird wieder über Informationen zu LTE-Nachbarzellen auf dem GSM oder UMTS Broadcast-Kanal ermöglicht. In GSM werden dazu die System Information 2-quarter Nachrichten verwendet und in UMTS der System Information Block (SIB) 19. Über Prioritäten für Bänder und Radiotechnologien kann zusätzlich zu den gemessenen Signalstärken der Netzwechsel beeinflusst werden. Somit können LTE Zellen mit einer niedrigeren Signalstärke gegenüber GSM und UMTS Zellen bevorzugt werden. Nach Rückkehr in das LTE-Netzwerk führt das Endgerät eine Tracking Area Update-Prozedur mit der MME durch. In der Tracking Area Update-Nachricht befinden sich dann auch Informationen über die bisherige Location- und Routing Area, über die die MME das zuletzt genutzte GSM oder UMTS Gateway ausfindig machen kann. Dieses wird dann von der MME über das Gn Interface kontaktiert, um den Kontext des Endgeräts zurück in das LTE-Kernnetzwerk zu holen. Nach Authentifizierung des Endgeräts kontaktiert die MME dann das GSM oder UMTS Netzwerk, sowie das PDN-Gateway, um den Kernnetztunnel entsprechend zu modifizieren. Schließlich wird noch das HSS über den neuen Aufenthaltsort des Endgeräts informiert.

1.8.2 RRC Connection Release mit Redirect zwischen LTE und GSM/UMTS

Während sich ein Endgerät im LTE RRC Connected State befindet, ist das Netzwerk für das Mobility Management zuständig. Dies ist notwendig, damit die Zeiten, in denen das Endgerät nach anderen Zellen und anderen Netzwerken sucht und somit keine Daten austauschen kann, koordiniert werden können.

In Abhängigkeit der unterstützten Funktionalitäten des Endgeräts und des Netzwerks kann der eNodeB das Endgerät anweisen, nach LTE, UMTS und GSM-Zellen auf anderen Kanälen und anderen Frequenzbändern zu suchen, nachdem ein konfigurierter Empfangspegel unterschritten wird. Die Suche erfolgt nach einem zwischen Endgerät und eNodeB definierten Muster, damit in diesen Zeiten keine Daten in Downlink-Richtung übertragen werden.

Wenn der Empfangspegel einen Wert erreicht, in dem Daten nicht mehr korrekt übertragen werden können, ist ein RRC Connection Release mit Redirect die einfachste Methode, nach UMTS oder GSM zu wechseln. Wenn zuvor UMTS und GSM inter-RAT-Messungen durchgeführt wurden, wählt das Netzwerk die beste Zelle aus und das Endgerät wechselt dann dorthin. Wurden zuvor keine Messungen ausgeführt, wählt das Netzwerk vorher definierte UMTS oder GSM Frequenzen aus und überlässt es dem Endgerät, die vorhandenen Zellen zu identifizieren.

Nachdem die RRC Connection Release mit Redirect-Nachricht empfangen wurde, beendet das Endgerät die Kommunikation mit der aktuellen Zelle, wechselt das Frequenzband und die Radiotechnologie und sucht bzw. synchronisiert sich mit der neuen Zelle. Kann die neue Zelle nicht gefunden werden, startet das Endgerät eine normale Netzwerksuche und wählt dann eine geeignete Zelle basierend auf Signalstärke und Priorität aus. Dieser zusätzliche Vorgang dauert jedoch einige Sekunden.

Nachdem sich das Endgerät mit der neuen Zelle synchronisiert hat, stellt es eine Signalisierungsverbindung her, um ein Location Area Update und ein Routing Area Update durchzuführen. Dies funktioniert in gleicher Weise wie zuvor für den Idle State beschrieben. Wird die Zelle, die in der RRC Release mit Redirect Nachricht beschrieben wird, sofort gefunden, dauert der Prozess typischerweise nur wenige Sekunden.

1.9 Carrier Aggregation

Bei der Einführung von LTE war eine Kanalbandbreite von 20 MHz geradezu revolutionär, eine Vervierfachung gegenüber den 5 MHz Kanälen von UMTS, die damals noch reichlich Kapazität boten. Über die Jahre wuchs jedoch der Kapazitätsbedarf pro Basisstation und 3GPP erweiterte deshalb den LTE Standard um ein Verfahren, mit dem mehrere Kanäle (Carrier) zu einem gemeinsamen Übertragungskanal zusammengefasst werden können. Dieses Verfahren wird Carrier Aggregation (CA) genannt. Um rückwärtskompatibel zu 3GPP Release 8 zu bleiben, wurde dabei die Carrier Bandbreite

von 20 MHz nicht geändert. Stattdessen kombiniert Carrier Aggregation die Kapazität von mehreren individuellen Carriern. Eine typische Konfiguration, die in der Praxis verwendet wird, ist die Kombination von einem oder mehreren Carriern in LTE Band 7 (2600 MHz Band) und einem oder mehreren Carriern in Band 3 (1800 MHz Band), um eine Kanalbandbreite von 40 MHz, 60 MHz oder sogar mehr im Downlink zu erreichen. Weitere Frequenzbänder, die ebenfalls für Carrier Aggregation verwendet werden, ist das 800 MHz Band (Band 20), das ursprünglich nur für GSM genutzte Band bei 900 MHz und das ursprüngliche UMTS Band bei 2100 MHz. Auch neu zugeteilte Frequenzbänder wie das 700 MHz Band und das 1500 MHz Downlink-Only Band werden in Zukunft für Carrier Aggregation verwendet werden.

Carrier Aggregation wird heute typischerweise asymmetrisch im Downlink und Uplink angewendet, da es heute wichtiger ist, die Kapazität im Downlink zu erhöhen. In der Downlink Richtung werden z. B. zwei bis fünf Carrier zu einer Kanalbreite von 40–80 MHz kombiniert. High-End Geräte können im Uplink heute bis zu zwei Carrier bündeln. Ob und welche Carrier aggregiert werden hängt davon ab, wie viele Carrier die Basisstation unterstützt und den Hardwareeigenschaften des Endgeräts, welche es dem Netzwerk als Teil der UE Capability Nachrichten beim Verbindungsaufbau mitteilt. Tab. 1.1 zeigt typische Endgeräte (UE) Kategorien die heute verwendet werden, sowie die dazugehörige maximale Anzahl von unterstützten gleichzeitigen Carriern.

Für die Zukunft wurden weitere Carrier Aggregation Konfigurationen mit einer noch höheren Kanalanzahl spezifiziert. Nach der ursprünglichen Limitierung auf 5 Carrier ist aktuell in den 3GPP Spezifikationen Carrier Aggregation mit bis zu 32 Carriern enthalten.

Aktuell ist zu beobachten, dass Carrier Aggregation nicht mehr nur eine Funktionalität von sehr teuren Endgeräten ist, sondern 2- oder 3-Carrier Aggregation auch im Mittelklassensegment angekommen ist. Im High-End Segment ist die Bündelung von 4–5 Kanälen aktuell Stand der Technik. In der Praxis hat Carrier Aggregation zwei Ziele: Zum einen erhöht es die theoretische Datenrate eines Nutzers. Zum anderen verringert die weiter zunehmende Anzahl von Nutzern mit steigenden Datenvolumen diesen Vorteil. Aus Netzkapazitätssicht ist Carrier Aggregation somit viel nützlicher, um die Downlink Daten vieler gleichzeitiger Nutzer pro Zelle dynamisch zu verteilen. Dies ist möglich, da CA-fähige Endgeräte im Downlink Daten in unterschiedlichen Teilen des Spektrums ohne Konfigurationsänderung empfangen können. In Abhängigkeit von wechselnden

Tab. 1.1 UE Kategorien und die Anzahl der unterstützten Carrier für Carrier Aggregation

UE Kategorie	Anzahl der Unterstützten CA Carrier
3,4	1
6	2
9,10	3
11,12,16,18	4
20	5

Signalbedingungen aller aktiven Geräte kann der eNodeB Scheduler sehr schnell den Teil des Spektrums ändern, in dem ein Gerät seine Daten empfängt, ohne dabei das Endgerät anzuweisen, das Frequenzband zu wechseln.

1.9.1 CA Varianten, Bandbreitenklassen und Bandkombinationen

In den Standards wurden 4 verschiedene Carrier Aggregation Varianten definiert:

- In der Praxis kombinieren Netzbetreiber oft Carrier in mehreren unterschiedlichen Frequenzbändern. Dies wird als Inter-Band Carrier Aggregation bezeichnet und ist heute die häufigste Form von Carrier Aggregation.
- Wenn ein Netzbetreiber mehr als 20 MHz zusammenhängendes Spektrum in einem Band besitzt, kann Intra-Band Contiguous Carrier Aggregation verwendet werden, z. B. 20+10 MHz. Der Netzbetreiber könnte hier auch eine 15+15 MHz Unterteilung konfigurieren. Dies würde jedoch Endgeräte benachteiligen, die kein Carrier Aggregation von mehreren Carriern in diesem Band unterstützen.
- In manchen Fällen besitzt ein Netzbetreiber auch mehrere Abschnitte in einem Frequenzband, die nicht zusammenhängend sind. Eine solche Kombination wird als Intra-Band Non-Contiguous Carrier Aggregation bezeichnet.
- In manchen Ländern, wie z. B. Schweden, besitzen Netzbetreiber sowohl FDD (Frequency Division Duplex) als auch TDD (Time Division Duplex) Spektrum, in dem Uplink und Downlink abwechselnd auf dem gleichen Kanal übertragen werden. Mit Carrier Aggregation ist es dann möglich, mehrere FDD und TDD Kanäle zu verwenden. Dies wird als Inter-Band FDD/TDD Carrier Aggregation bezeichnet. In der Uplink Richtung wird einer der FDD Carrier für die Datenübertragung verwendet.

Auf der Netzseite ist die gleichzeitige Übertragung von Daten in unterschiedlichen Teilen des Spektrums einfach. Hier wurden schon vor der Einführung von Carrier Aggregation Daten in mehreren Frequenzbändern gleichzeitig übertragen. Somit wurde auf Netzseite nur ein Software Update benötigt, falls ein eNodeB aus Kapazitätsgründen sowieso schon in mehreren Bändern aktiv war. Auf der Endgeräteseite ist die Nutzung von Carrier Aggregation hingegen wesentlich komplexer, da der Platz für zusätzliche Hardware wie Filter, zusätzliche Antennen, Rechenleistung, etc. Um dem Fortschritt in der Hardwareentwicklung und der Standardisierung Rechnung zu tragen, wurden Mechanismen spezifiziert, damit ein Endgerät dem Netzwerk beim Verbindungsauflauf mitteilen kann, welche CA Bandkombinationen es unterstützt. Jeder Carrier einer Bandkombination wird als Component Carrier (CC) bezeichnet. Zusätzlich wird in 3GPP TS 36.101 Tab. 5.6A-1 spezifiziert, wie viele direkt aufeinanderfolgende (Contiguous) Component Carrier ein Gerät aggregieren kann. Dies wird als Carrier Aggregation Bandwidth Class bezeichnet und in Tab. 1.2 gezeigt.

Tab. 1.2 CA Bandbreitenklassen

CA Bandwidth Class	Maximal aggregierte Bandbreite	Anzahl der zusammenhängenden Component Carrier
A	20 MHz	1
B	20 MHz	2
C	20–40 MHz	2
D	40–60 MHz	3
E	60–80 MHz	4
F	80–100 MHz	5
I	140–160 MHz	8

Aktuell werden in der Praxis die Bandbreitenklassen A bis C verwendet. Die weiteren Klassen wurden für die Kombination von LTE Carriern in den üblichen lizenzierten Bändern in Kombination mit Carriern im lizenzenfreien 5 GHz Band spezifiziert, in dem auch Wifi verwendet wird. Diese Carrier Aggregation Kombination mit lizenziertem und lizenzenfreiem Spektrum wird License Assisted Access (LAA) genannt.

Die folgenden Beispiele zeigen typische Carrier Aggregation Konfigurationen, die heute in der Praxis verwendet werden, sowie die in den Standards dazu verwendete Nomenklatur:

- CA_3A-7A: Aggregiert bis zu 20 MHz in Band 3 (1800 MHz, 3A) und bis zu 20 MHz in Band 7 (2600 MHz, 7A) für einen kombinierten 40 MHz Kanal. Falls ein Netzbetreiber weniger Spektrum zur Verfügung hat, z. B. nur 10 MHz in Band 3, beträgt die kombinierte Bandbreite 30 MHz.
- CA_3C-7A: Aggregiert bis zu 40 MHz in Band 3 und bis zu 20 MHz in Band 7 für einen kombinierten 60 MHz Kanal.
- CA_3A-3A-7A: Aggregiert bis zu 40 MHz in Band 3 und bis zu 20 MHz in Band 7 für einen kombinierten Kanal von 60 MHz. Im Unterschied zum vorigen Beispiel unterstützt ein Gerät mit dieser Kombination zwei nicht aufeinanderfolgende (non-contiguous) Kanäle in Band 3.
- CA_8A-3A-7A: Aggregiert bis zu 20 MHz in Band 8 (900 MHz), bis zu 20 MHz in Band 3 (1800 MHz) und bis zu 20 MHz in Band 7 (2600 MHz). In der Praxis ist es unwahrscheinlich, dass ein Netzbetreiber 20 MHz Spektrum im 900 MHz Band besitzt, da dieses noch immer für GSM verwendet wird. Somit ist eine Kombination von 10+20+20 MHz oder 5+20+20 MHz für 50 MHz bzw. 45 MHz Bandbreite wahrscheinlicher.
- CA_3A-7A-38A: Aggregiert bis zu 20 MHz in den Bändern 3 und 7, in denen Frequency Division Duplex (FDD) verwendet wird. Zusätzlich werden bis zu 20 MHz in Band 38 aggregiert, in dem Time Division Duplex (TD-LTE) verwendet wird.

In der Praxis unterstützen Endgeräte heute typischerweise mehrere hundert unterschiedliche Kombinationen, um die sehr diversen Spektrumszuteilungen an Netzbetreiber in vielen Teilen der Welt abzudecken. Aus diesem Grund wurden die CA Kombinationslisten in den UE Capability Information Nachrichten beim Verbindungsaufbau immer länger. Um die Listen und somit die Nachrichtengröße möglichst kurz zu halten, wurden Mechanismen in den Standard aufgenommen, die es dem Netzwerk erlauben, nur nach Bandkombinationen für die im Netzwerk verwendeten Frequenzbereiche zu fragen. Somit übertragen Endgeräte in Europa z. B. keine Bandkombinationen an das Netzwerk, die nur in Nordamerika verwendet werden.

1.9.2 CA Konfigurationen, Aktivierung und Deaktivierung

In der Praxis wird Carrier Aggregation wie folgt aktiviert: Wenn ein Endgerät vom RRC-IDLE in den RRC-Connected Zustand wechselt, wird zunächst nur ein einzelner Carrier verwendet. Falls das Netzwerk die Endgerätekonfiguration beim letzten Kommunikationsaufbau nicht gespeichert hat, sendet es eine UE Capability Enquiry Nachricht, die das Endgerät dann mit einer UE Capability Information Nachricht beantwortet. Diese Nachricht enthält unter anderem Informationen über die maximal unterstützte Datenrate, unterstützte Frequenzbänder und Carrier Aggregation Kombinationen.

Wann und ob zusätzliche Carrier konfiguriert werden ist implementationsabhängig. In einer heute typischen Herangehensweise konfiguriert das Netzwerk die Zellparameter so, dass das Endgerät im RRC-IDLE Zustand immer eine Zelle auf einem möglichst hohen Frequenzband bevorzugt, auch wenn dort die Signalqualität nicht so gut ist, wie auf Carriern in niedrigeren Frequenzbändern. Wenn das Endgerät sich dann mit dem eNodeB auf dieser hohen Frequenz verbindet, geht dieser dann davon aus, dass das Endgerät auch die Carrier in niedrigeren Frequenzbändern empfangen kann und konfiguriert Carrier Aggregation sofort während des Verbindungsaufbaus. Der für Downlink und Uplink verwendete Carrier wird auch als Primary Component Carrier (PCC) oder auch als Primary-Cell (PCell) bezeichnet. Weitere Component Carrier werden als Secondary Component Carrier (SCC) oder auch Secondary-Cells (SCell) bezeichnet. Das Hinzufügen von SCells während des Verbindungsaufbaus ist sehr schnell und dauert typischerweise weniger als 100 ms.

Manche Netzwerke verwenden eine kompliziertere Methode um zu ermitteln, ob es möglich und sinnvoll ist, zusätzliche Carrier einzubinden. Weitere Component Carrier werden dort nur dann zugeschaltet, wenn größere Datenmengen zu übertragen sind. Falls nicht klar ist, ob ein Endgerät alle Component Carrier, die der eNodeB aggregieren möchte, empfangen kann, z. B. weil das Endgerät nicht auf dem höchsten Frequenzband aktiv ist, weißt der eNodeB das Endgerät an, eine Inter-Band Signalstärkemessung durchzuführen. Wenn die zurückgemeldete Signalstärke von potenziellen SCells gut genug ist, rekonfiguriert der eNodeB dann entsprechend die Verbindung. Der Vorteil

dieser komplizierteren Prozedur ist, dass höhere Frequenzbänder für Carrier Aggregation kompatible Geräte freigehalten werden. Nicht CA-fähige Endgeräte sind im RRC-IDLE Zustand in niedrigeren Frequenzbändern, da die Zellen so konfiguriert sind, dass diese bevorzugt werden. Außerdem ist eine vorherige Messung vorteilhaft, damit nur SCells, die gut empfangen werden können, dem Kanal hinzugefügt werden. Dieser Ansatz hat jedoch auch Nachteile. Viele Endgeräte ändern den Netzwerkindikator z. B. von ‚LTE‘ nach ‚LTE+‘, sobald Carrier Aggregation verwendet wird. Dies geschieht bei diesem Ansatz wesentlich seltener als beim ersten und einfacheren Ansatz. Ein weiterer Nachteil ist der Umstand, dass bei der Übertragung von größeren Datenmengen die maximale Geschwindigkeit nicht gleich nach dem Verbindungsaufbau erreicht werden kann, sondern erst nach einigen Sekunden.

Typischerweise werden während der Konfiguration von SCells auch zusätzliche Signalqualitätsmessungen konfiguriert, damit der eNodeB dann dynamisch je nach Signalqualität Datenpakete über die einzelnen Component Carrier übertragen kann. Abb. 1.28 zeigt, wie CA in der Praxis konfiguriert wird.

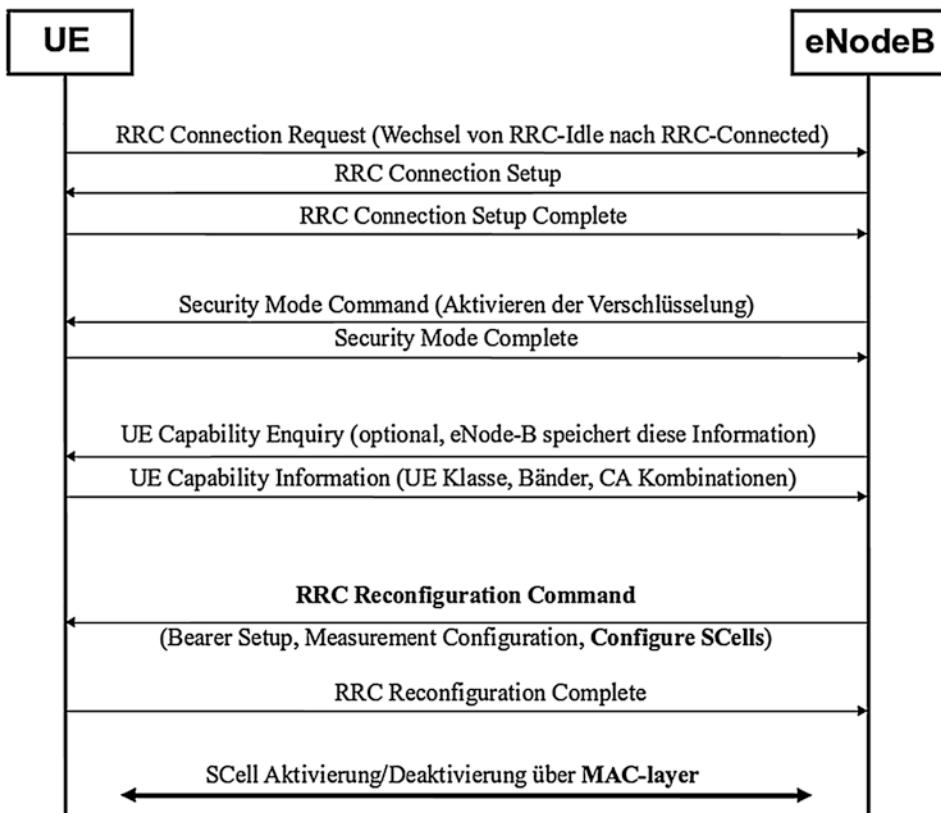


Abb. 1.28 CA Konfiguration während des RRC Verbindungsbaus

Nachdem die SCells über RRC Nachrichten konfiguriert wurden, kann der eNodeB diese jederzeit über MAC-Layer Kontrollelemente aktivieren. Die SCells können dann 8 Subframes später verwendet werden, also nach 8 ms und bleiben, falls konfiguriert, bis zum Ablauf des SCellDeactivationTimer aktiv. Ein anderer Weg, SCells zu deaktivieren ist mit einem weiteren MAC-Layer Kontrollelement.

Nachdem die SCells aktiviert sind, werden die Resource Blocks der Component Carrier separat auf jedem CC zugeteilt. Das bedeutet, dass das Endgerät den Physical Downlink Control Channel (PDCCH) jedes Component Carriers nach Zuweisungen abhören muss. Optional wurde auch das Cross-Carrier Scheduling spezifiziert. Dieses Verfahren ermöglicht die Zuteilung von SCell Ressourcen auf dem PDCCH der PCell. Die Aufteilung von Carrier Aggregation in eine langsame Konfigurationsphase, in der mehrere Parameter in einer RRConfiguration Nachricht übertragen werden und einer schnellen Aktivierung/Deaktivierung mit nur wenigen und kleinen Parametern in MAC-Layer Kontrollelementen ermöglicht die schnelle Reduktion der Leistungsaufnahme des Endgerätes, wenn keine oder nur wenige Daten zur Übertragung anstehen. Sind wieder mehr Daten zu übertragen, können die SCells dann wieder schnell zugeschaltet werden.

Abb. 1.29 zeigt, wie der Energieverbrauch reduziert werden kann, wenn wenig oder keine Daten am eNodeB auf die Übertragung warten: Zunächst gibt es genug Daten, um über die PCell und die zwei konfigurierten SCells übertragen zu werden. Nachdem der Transmit Data Buffer auf Netzwerkseite leer ist, entscheidet sich der eNodeB die SCells zu deaktivieren. Wenig später treffen wieder Daten ein und die SCells werden wieder aktiviert. Danach leert sich der Transmit Data Buffer erneut und die SCells werden wieder deaktiviert, um auf der Endgeräteseite die Leistungsaufnahme zu reduzieren. Da weiterhin keine Daten eintreffen, wird DRX (Discontinuous Reception) für die Verbindung konfiguriert, um noch mehr Energie einzusparen. Nachdem weitere Zeit vergangen ist, ohne dass erneut Daten beim eNodeB für die Übertragung eingehen, gibt der

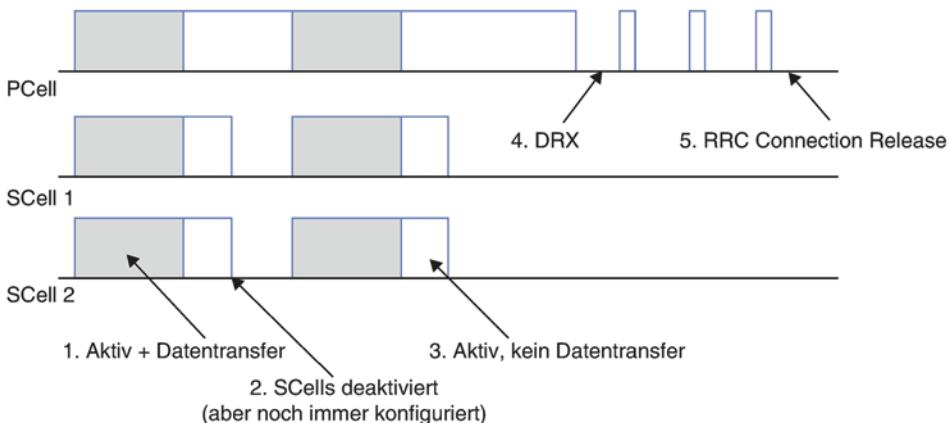


Abb. 1.29 Nutzung von PCell und SCell Ressourcen, DRX und Idle, um Energie zu sparen

eNodeB die RRC Verbindung dann komplett auf. Das Endgerät überprüft dann nur noch periodisch den Paging Kanal und ist im energiesparendsten Zustand, in dem es weiterhin die IP Adresse behält und für eingehende IP Pakete erreichbar ist.

1.9.3 Uplink Carrier Aggregation

Auch im Uplink unterstützen heute viele höherwertigen Geräte und Netzwerke Carrier Aggregation. Üblicherweise werden heute maximal zwei Carrier im Uplink gebündelt, wobei es zusätzliche Limitationen gibt:

Durch die begrenzte Sendeleistung des Endgeräts von -23 dBm ($0,2 \text{ W}$) kann Uplink Carrier Aggregation nicht am Zellrand verwendet werden. Dort ist es sogar notwendig, die Sendeleistung auf einen Teil des Kanals zu bündeln, um die Distanz zum Sendeort und die Interferenz von Nachbarzellen zu überwinden. Außerdem wird ein zusätzlicher Sender für jeden Kanal im Uplink benötigt. Typischerweise haben Endgeräte zwei oder drei Sendeeinheiten, um in allen Bereichen des Low-, Mid- und High Band Spektrums zwischen 700 MHz und 2600 MHz senden zu können. Das bedeutet, dass manche Endgeräte nicht gleichzeitig im 1800 und 2600 MHz Band senden können, da hierfür die gleiche Sendeeinheit im Gerät verwendet wird. Somit können auch teure Endgeräte oft nur eine oder beide der folgenden Carrier Aggregation Kombinationen im Uplink verwenden:

- Kombinationen aus einem Low Band (z. B. 800 MHz) und einem High Band (z. B. 1800 oder 2600 MHz) Kanal.
- Eine Kombination von zusammenhängenden Kanälen im gleichen Band, z. B. zwei Kanäle in Band 3, die unmittelbar aufeinander folgen.

Als Konsequenz daraus kann Uplink CA manchmal nicht verwendet werden, obwohl dies auf der Netzseite möglich wäre:

- Dem Endgerät wurde Band 7 (2600 MHz) als Primary Component Carrier zugewiesen, das Endgerät unterstützt aber nur Carrier Aggregation von zwei nebeneinander liegenden Kanälen in Band 3.
- Band 3 (1800 MHz) und Band 1 (2100 MHz) wurden dem Gerät für Downlink Carrier Aggregation zugewiesen. Uplink CA mit $2 \times$ Band 3 ist somit nicht möglich, da der dazu benötigte Kanal im Downlink fehlt.

In diesen Szenarien werden dem Endgerät nur High-Band Kanäle für Downlink CA zugeteilt, die auf Endgeräteseite im Uplink mit nur einem Sender bedient werden. Die Sendeeinheit kann jedoch nur ein zusammenhängendes Signal senden. Somit ist es nicht möglich, UL CA zu aktivieren.

Welche CA Kombinationen ein Endgerät im Uplink unterstützt, wird während des RRC Verbindungsaufbaus signalisiert. Theoretisch könnte das Netzwerk darauf entsprechend reagieren und eine geeignete Primary Cell und eine CA Kombination auswählen, die den CA Fähigkeiten des Endgerätes am Besten entspricht. Dies würde aber in vielen Fällen bedeuten, dass das Netzwerk zuerst Messungen konfigurieren und dann die Primary Cell wechseln müsste. Ein Nachteil wäre jedoch, dass die Datenrate während des Messens und der Rekonfiguration geringer ist. Aus diesem Grund führen nicht alle Netzbetreiber eine solche Prozedur durch.

1.10 Von Dipolen zu Aktiven Antennen und Gigabit Backhaul

In den letzten zwei Jahrzehnten gab es im Bereich des Radionetzwerkes enorme Weiterentwicklungen. Konnten ursprünglich nur Datenströme von wenigen Kilobit pro Sekunde für Telefonie pro Teilnehmer übertragen werden, sind heutige Basisstationen in der Lage, Endgeräte mit Datenraten von mehreren hundert Megabit pro Sekunde mit dem Internet zu verbinden. In den Anfangszeiten von GSM war eine Basisstation im Vergleich zu heute sehr einfach aufgebaut und unterstützte typischerweise nur ein Frequenzband, z. B. das 900 MHz GSM Band. Besonders im ländlichen Raum waren große aber einfache Stabantennen (Dipole, Kugelstrahler) recht oft anzutreffen. In Städten wurden hingegen oft drei Sektorantennen verwendet, um die Kapazität, also die gleichzeitig möglichen Telefongespräche pro Basisstation zu erhöhen. Die digitale Basebandeinheit, sowie das Radio Frontend einer Basisstation, waren in dieser Zeit gewöhnlich in einem großen und klimatisierten Schrank oder Raum untergebracht, üblicherweise am Fuße des Antennenmastes. Mit Koaxialkabeln wurde dann das Antennensignal von und zur passiven Antenne am oberen Ende des Antennenmast geleitet.

Mit Einführung von UMTS im Jahre 2003 erhöhte sich dann die Komplexität der Installation an einem Standort, da nun gleichzeitig zwei Frequenzbänder unterstützt werden mussten, üblicherweise 900 bzw. 1800 MHz für GSM und 2100 MHz für UMTS. Typischerweise installierten Netzbetreiber Dualband Sektorantennen, die aus zwei getrennten Antennen in einem Antennengehäuse bestanden. Dies erhöhte zwar die Anzahl der nötigen Koaxialkabel, konnte aber in der Praxis noch gut umgesetzt werden. Mit Start von LTE im Jahre 2009 musste dann wiederum ein zusätzliches Frequenzband pro Standort eingeführt werden. Die Technologie hatte sich jedoch weiterentwickelt und viele Netzbetreiber nutzten die Einführung der neuen Technologie, um sogenannte „Remote Radio Heads“ (RRHs) zu installieren, die sich in unmittelbarer Nähe der Antennen befinden, statt wie bisher zusammen mit den weiteren Komponenten am Fuße des Antennenmastes. In einer typischen 3-Sektor Konfiguration werden je ein oder mehrere Remote Radio Heads pro Sektor verwendet. RRHs reduzierten die Kosten erheblich, da teures Koaxialkabel eingespart werden kann. Außerdem wurde die Energieeffizienz der Basisstation erhöht, da nur noch Koaxialkabel für die kurze Strecke zwischen RRH und der passiven Antenne nötig sind. Außerdem konnten Größe und

Gewicht der digitalen Komponenten und der Backhaul Übertragungskomponenten deutlich reduziert werden, da die digitale Signalverarbeitung für GSM, UMTS, LTE und 5G NR in einer gemeinsamen Basebandkomponente zusammengefasst werden konnte. Dies wird auch als ‚Single-RAN‘ Technologie bezeichnet. Diese Größenreduktion bewirkte auch, dass die Basebandeinheit näher an den Antennenmast, oder in vielen Fällen auch direkt an den Antennenmast unterhalb der Antenne, angebracht werden konnte. Außerdem können diese Komponenten passiv gekühlt werden, es ist also keine Klimaanlage mehr nötig. Diese Konfiguration wird heute von den meisten Netzbetreibern in Kombination mit zusätzlichen Frequenzbändern verwendet. Eine Basisstation sendet heute oftmals LTE Träger im 800, 900, 1800, 2100 und 2600 MHz Band aus, sowie zusätzlich GSM Träger im 900 MHz Band und einen sehr breitbandigen Träger für 5G NR im Bereich von 3600 MHz. In Zukunft ist zu erwarten, dass noch weitere Bänder hinzukommen. Glücklicherweise hat sich auch die Antennentechnik weiterentwickelt und eine flache Sektorantenne enthält heute 3 Antennen, z. B. eine Antenne für den Frequenzbereich von 700–900 MHz, eine Antenne für den Frequenzbereich von 1800–2100 MHz und eine weitere Antenne für den Frequenzbereich um 2600 MHz. Es gibt auch Sektorantennen mit nur zwei internen Antennen, die die gleiche Anzahl an Frequenzbändern unterstützen, da eine interne Antenne den kompletten Frequenzbereich von 1800–2600 MHz abdecken kann.

Die nächste Evolutionsstufe im Basisstations- und Antennendesign sind heute aktive Antennen, bei denen der bisher separate Remote Radio Head auf der Rückseite der Antenne angebracht ist und somit zu einem integrierten Bauteil wurde. Außerdem wird eine einzelne Antenne in viele kleine Antennen aufgeteilt, die vom integrierten Remote Radio Head einzeln angesteuert werden können. Auf diese Weise ist es möglich, den vertikalen Abstrahlwinkel (Tilt) der Antenne elektronisch zu steuern und 4×4 MIMO über zwei separate Cross-Polarized Antennenstränge zu ermöglichen. Die Größe und das Gewicht der Sektorantenne erhöht sich jedoch dadurch. Dies bedeutet, dass unter Umständen der Antennenmast verstärkt werden muss, um die stärkeren Windkräfte, die durch das zusätzliche Gewicht und Größe entstehen, ausgleichen zu können. Ein Vorteil ist jedoch, dass nur noch ein einzelnes Glasfaserkabel nötig ist, um die aktive Antennenanordnung mit der digitalen Basebandeinheit des Standortes zu verbinden. Einer der ersten Netzbetreiber, der aktive LTE Antennen mit 4×4 MIMO in 2016 installierte, war T-Mobile in Nordamerika. Für 5G NR bei 3600 MHz setzen heute die meisten Netzbetreiber ebenfalls auf aktive Antennen.

Auch bei den Backhaulkomponenten, mit denen Basisstationen mit dem Netzwerk verbunden werden, hat sich über die Jahre viel geändert, da heute Übertragungskanäle mit sehr hohen Datenraten für die Anbindung des sehr breitbandigen LTE und 5G NR von entscheidender Bedeutung sind. Eine 3-Sektor Basisstation mit einer LTE Kanalbandbreite von 20 MHz kann durchaus in der Praxis eine Spitzendatenrate von 300 Mbit/s erzeugen. Durch Carrier Aggregation werden heute oftmals mehr als nur ein 20 MHz Träger pro Sektor verwendet, was wiederum zu noch höherem Bandbreiten-

bedarf führt. In der Praxis sind heute GSM, LTE und 5G NR in einer Basisstation zu finden, was ebenfalls zur weiteren Steigerung der benötigten Bandbreite führt.

Traditionell wurden Basisstationen zunächst mit Kupferkabel an das Netzwerk angeschlossen. Für GSM und UMTS wurden dafür ursprünglich 2 Mbit/s E-1 Verbindungen verwendet und für eine gewisse Zeit war für UMTS das Bündeln mehrerer solcher Verbindungen ausreichend. Mit der Einführung von LTE wurden jedoch 2 Mbit/s Leitungen zum Flaschenhals und konnten nicht mehr weiterverwendet werden. Aus diesem Grund wurden dann die Kupferleitungen durch Glasfaserkabel ersetzt. Während der Einsatz von Glasfasern optimal für LTE Basisstationen ist, ist es in der Praxis manchmal nicht möglich, diese an und in Gebäuden zu installieren. Netzbetreiber, die neben einem Mobilfunknetzwerk auch noch Festnetzverbindungen wie VDSL für Privat- und Firmenkunden anbieten, sind hier deutlich im Vorteil, da die Glasfasern nicht nur für den Mobilfunk, sondern auch für die Festnetzanbindung von Kunden verwendet werden können. Für reine Mobilfunkbetreiber sind deshalb unter Umständen Highspeed Ethernet basierte Mikrowellenanbindungen eine gute Alternative. Gängige Mikrowellensysteme bieten Datenübertragungsraten im Gbit/s Bereich.

Aktuelle Highspeed Backhaulverbindungen verwenden ausschließlich IP als Übertragungsprotokoll. Dies ist für LTE ideal, da alle Schnittstellen des Systems auf dem IP Protokoll basieren. Auch die Anbindung älterer Mobilfunkgenerationen nutzen heute diesen Backhaul, deren Übertragungsprotokolle wie z. B. E-1, werden über IP simuliert.

1.11 Self-Organizing Networks

In heutigen Mobilfunknetzwerken gibt es eine Vielzahl an Aufgaben, die während des Netzaufbaus und später während des Betriebs manuell durchgeführt werden müssen, um eine möglichst hohe Verfügbarkeit und Kapazität des Netzwerkes zu gewährleisten. Da dies mit großem Aufwand verbunden ist, wurde in 3GPP ein Arbeitspaket mit dem Namen „Self-Organizing Networks“ (SON) erstellt, um einige dieser Aufgaben zu automatisieren. 3GPP Technical Recommendation (TR) 36.902¹⁸ gibt einen Überblick, welche Funktionen teilweise oder ganz automatisiert werden könnten. Von der nachfolgenden Auswahl werden in der Praxis heute viele eingesetzt:

- Self-Configuration: Während des ersten Systemstarts eines eNodeBs verbindet sich dieser mit einem zentralen Server und holt sich von dort seine Konfiguration.
- Automatic Neighbor Relation (ANR): Mobile Endgeräte können Informationen über Nachbarzellen an die gerade aktuelle Zelle liefern. Hier kann diese Information dann verwendet werden, um evtl. fehlende Nachbarschaftsbeziehungen zu vervollständigen und diese dann später für Handover zu verwenden.
- Netzabdeckungs- und Kapazitätsoptimierung: Interferenz aufgrund von sich zu stark überlappenden Zellen und Löcher in der Abdeckung sind einige der Hauptprobleme in Netzwerken. Heute wird hauptsächlich durch Testfahrten sichergestellt, dass diese

Probleme auf ein Minimum reduziert werden. Mit SON-Funktionalität soll die Zahl der notwendigen Testfahrten reduziert werden, in dem Signalstärkemessungen von Endgeräten ausgewertet werden.

- Energieeinsparungen: Nachts und bei geringer Systemlast können einzelne Carrier in Sektoren abgeschaltet werden. Werden wieder mehr Daten übertragen, werden die Carrier dann wieder automatisch aktiviert.
- Automatische Konfiguration der Physical Cell-ID (PCI): Wie zuvor beschrieben, wird während des Suchprozesses in LTE eine Physical Cell-ID verwendet, um unterschiedliche Zellen auf der gleichen Frequenz zu unterscheiden. Insgesamt gibt es nur 504 verschiedene IDs und benachbarte Basisstationen können nur bestimmte Kombinationen verwenden. Da die Verwendung von PCIs nicht immer vorhersehbar ist, wäre eine automatische Konfiguration sehr wünschenswert. Auch diese Funktion benötigt die Endgeräte, um die PCIs der Nachbarzellen zu melden.
- Handover-Optimierung: Durch die Analyse von Fehlermeldungen können Löcher in der Funkabdeckung und falsche Handoverentscheidungen entdeckt und behoben werden.
- Lastoptimierung: Ist die aktuelle Last einer Zelle schon sehr hoch, können Endgeräte, die sich am Rand der Zelle befinden, an Nachbarzellen weitergegeben werden.
- Optimierung des Random Access Channels: Der Random Access Channel (RACH) wird während der ersten Kommunikation zwischen Endgerät und eNodeB benötigt, wenn diese noch nicht synchronisiert sind. In Abhängigkeit der Netzwerklast könnten die Ressourcen, die für den RACH benötigt werden, automatisch angepasst werden.

1.12 Kapazität eines Standortes und Anzahl gleichzeitiger Nutzer

Durch die Nutzung von LTE Carrier Aggregation kann die theoretische Kapazität eines Mobilfunkstandortes mit 3 Sektoren mehrere Gigabit pro Sekunde betragen. Dies ist der Fall, da selbst die theoretische Datenrate eines einzelnen 20 MHz Kanals mit 256 QAM Modulation und 4×4 MIMO auf dem Physical Layer schon 375 Mbit/s beträgt. Auf dem IP Layer ist der theoretische Durchsatz etwa 15 % geringer, also in etwa 320 Mbit/s. In der Praxis haben jedoch nicht alle Nutzer ideale Kanalbedingungen und Standorte in der Nachbarschaft erzeugen Interferenz. Daraus folgt, dass die tatsächlich nutzbare Kapazität deutlich niedriger ist. Ein Netzausrüster geht z. B. davon aus, dass der Durchsatz in einem 20 MHz Kanal in der Praxis etwa 10 % des theoretischen Maximums beträgt¹⁹. Typischerweise verwenden Netzbetreiber heute 50–60 MHz an Spektrum für LTE und ein Standort verfügt üblicherweise über 3 unabhängige Sektoren. Die durchschnittliche Datenrate eines solchen Standortes beträgt somit etwa 300 Mbit/s. Nutzer mit sehr guten Signalbedingungen können jedoch auch sehr viel höhere Datenraten erreichen, besonders dann, wenn ein Sektor eines Standortes nicht sehr ausgelastet ist.

Weitere Faktoren, die für die maximal erreichbare Datenrate eines Nutzers eine große Rolle spielen, sind die Anzahl der Nutzer die gerade mit der Zelle verbunden sind, wie viele Nutzer gerade Daten übertragen und deren durchschnittlich übertragenes Datenvolumen. Wie zuvor beschrieben, kann sich ein Endgerät in unterschiedlichen Zuständen befinden. Ist es angeschaltet, aber nicht aktiv genutzt, befindet sich das Endgerät hauptsächlich im RRC-IDLE Zustand. In diesem Zustand hat es eine IP Adresse, aber es existiert kein „Kontext“ im eNodeB. Das Endgerät überprüft passiv und nur periodisch die System Information Nachrichten, führt Signalstärke- und Qualitätsmessungen durch, und wechselt selbstständig zwischen den Zellen. Außerdem überprüft es periodisch den Paging Kanal, um auf ankommende Anrufe, SMS Nachrichten und auf eingehende IP Pakete reagieren zu können. All dies wird passiv durchgeführt, es wird also keine Bandbreite für die Kommunikation benötigt und somit keine Last im eNodeB erzeugt.

Sollen Daten übertragen werden, verlässt das Endgerät den RRC-IDLE Zustand und baut eine aktive Verbindung zu einem eNodeB auf. Um die Verbindung über die Luftschnittstelle aufrecht zu halten, auch wenn gerade keine IP Pakete übertragen werden, müssen Endgeräte und eNodeB im Abstand von wenigen Millisekunden ständig Signalisierungsinformationen übertragen. Dies betrifft besonders den Uplink. Dies ist nötig, damit der eNodeB die aktuelle Empfangssituation in Sende- und Empfangsrichtung abschätzen kann. Somit ist es dann möglich, dass ankommende IP Pakete sofort mit einer dem Kanal entsprechenden Kodierung und Modulation gesendet werden können. Nachdem der Sendepuffer auf beiden Seiten der Verbindung leer ist, wird der RRC-Connected Zustand für einige Zeit beibehalten, da oftmals innerhalb kurzer Zeit weitere IP Pakete zur Übertragung eintreffen, die dann auch sofort weitergeleitet werden sollen. Erst wenn für einige Zeit keine IP Pakete mehr im Sendepuffer auf beiden Seiten eingehen, wird die Verbindung dann wieder vom RRC-Connected in den RRC-IDLE Zustand gesetzt. Die Zeit bis zu dieser Zustandsänderung ist im Netzwerk konfigurierbar und beträgt typischerweise 10–30 s. Das bedeutet, dass auch bei sehr sporadischer Übertragung von kleinen IP Paketen, z. B. Messenger Status Updates, das Endgerät trotzdem während der ganzen Zeit im RRC-Connected Zustand bleibt und periodisch Status Informationen übertragen werden. Während dieser Zeit kann das Netzwerk für das Endgerät jedoch den Connected-DRX (Discontinuous Reception) Modus aktivieren. Dieser reduziert die Signalisierung in Uplink Richtung, erhöht jedoch die Verzögerungszeit, bevor neue IP Pakete wieder übertragen werden können.

Smartphones stellen heute die größte Anzahl an Endgeräten in einem Mobilfunknetzwerk dar. Das bedeutet, dass ein signifikanter Anteil der Netzkapazität für die Übertragung von Hintergrunddaten verwendet wird, die von Applikationen generiert werden, die ständig eine Verbindung zu einem Server aufrechterhalten. Dies sind z. B. Messenger Apps, die ständig erreichbar sein müssen und deshalb periodisch kleine Datenpakete schicken, um die Verbindung aufrechtzuerhalten. Unter der Annahme, dass dies alle 4 min passiert, erzeugt ein Endgerät 15 solcher Datentransfers pro Stunde im Hintergrund, für die jedes mal eine Verbindung über die Luftschnittstelle aufgebaut werden muss. Trotzdem werden nur sehr wenig Daten übertragen und es müssen dazu,

in Abhängigkeit der Connected-DRX Konfiguration, mehr oder weniger Signalisierungsdaten übertragen werden.

Um die Anzahl der Nutzer abzuschätzen, die sich typischerweise an einem eNodeB im RRC-Connected Zustand befinden, sind zwei weitere Parameter notwendig: Die Anzahl der eNodeBs im Netzwerk und die Anzahl der Geräte, die sich im Netzwerk befinden. In Deutschland hat ein Netzbetreiber beispielsweise etwa 20.000 Standorte und etwa 40 Mio. Kunden. Es wird weiterhin für dieses Beispiel angenommen, dass in dieser Zahl etwa 30 Mio. Smartphones enthalten sind. Das bedeutet, dass 30 Mio. Smartphones auf 20.000 Standorte aufgeteilt sind, also etwa 1500 Geräte pro Standort. Da ein Standort üblicherweise 3 Sektoren hat, werden etwa 500 Geräte pro Sektor versorgt. Ist jedes Gerät 7 min pro Stunde aktiv, sind somit immer etwa 60 Geräte im RRC-Connected Zustand, um Hintergrunddaten zu übertragen. Zusätzlich dazu befinden sich noch weitere Endgeräte im RRC-Connected Zustand, die aktiv von ihrem Besitzer verwendet werden und Daten übertragen. Daraus ergibt sich eine typische Anzahl von 50 gleichzeitig verbundenen Endgeräten in einem Sektor an wenig genutzten Standorten und 100–200 aktive Endgeräte in stark frequentierten Sektoren.

1.13 CS-Fallback für Sprache und SMS mit LTE

Eine entscheidende Neuerung von LTE war, dass sowohl das Kernnetzwerk, als auch das Radionetzwerk komplett auf IP basieren. Dies bedeutete, dass das in GSM und UMTS genutzte leitungsvermittelnde Kernnetzwerk und das Mobile Switching Center (MSC) für Sprachtelefonie nicht weiterverwendet werden konnte. Auch das Short Message Service Center (SMSC) konnte nicht direkt an das LTE Kernnetz angeschlossen werden. Dies vereinfacht natürlich das Netzwerkdesign und folgte dem allgemeinen Trend im Festnetz. Auch Festnetzbetreiber bieten heute ihren Sprachdienst nur noch über IP an. Beim Kunden steht dazu ein multifunktionales Endgerät, das ein integriertes DSL Modem mit Wi-Fi Access Point, DECT-Telefon Basisstation und Anschlüsse für normale Analogtelefone enthält. In der Box wird dann der analoge Sprachkanal und die Signalisierung zum Gesprächsaufbau in das Session Initiation Protocol (SIP) gewandelt, das vollständig auf IP basiert. Sprach- und Signalisierungsdaten werden dann über DSL und IP weitergeleitet, statt wie bisher über einen analogen Kanal zur Vermittlungsstelle.

Auch für LTE wurde eine rein IP basierte Lösung mit dem IP Multimedia Subsystem (IMS) angedacht, die unter dem Begriff Voice over LTE (VoLTE) firmiert. Trotz der ersten Standardisierungsbemühungen ab 3GPP Release 5 war jedoch zum Start der ersten LTE Netzwerke noch kein fertig spezifiziertes und implementiertes System für LTE verfügbar. Man entschloss sich deshalb, die vorhandenen Systeme trotz Inkompatibilität zu LTE zunächst weiterzuverwenden. Diese Lösung wird als CS-Fallback bezeichnet und wird in den folgenden Unterkapiteln beschrieben. Dem eigentlichen IMS Sprachsystem und VoLTE ist dann später ein eigenes Kapitel gewidmet. Auch wenn heute die meisten Netzwerke und Endgeräte VoLTE unterstützen, wird der

CS-Fallback Mechanismus trotzdem noch gebraucht. Zum einen gibt es durchaus noch viele nicht VoLTE-fähige Endgeräte im Netzwerk und zum anderen ist der VoLTE Dienst im internationalen Roaming bisher noch nicht sehr verbreitet.

1.13.1 SMS über SGs

Ein Dienst, der unbedingt auch über LTE funktionieren sollte, ist der Short Message Service (SMS). In GSM verwendet SMS die Signalisierungskanäle des leitungsvermittelnden Teils des Netzes. Um SMS auch ohne IMS über LTE anbieten zu können, einigte man sich in 3GPP zunächst auf eine Zwischenlösung, die sich SMS über SGs nennt und nicht direkt auf IP basiert. Standardisiert ist diese Lösung in 3GPP TS 23.272²⁰. Wie in Abb. 1.30 gezeigt, werden SMS Nachrichten mit dieser Lösung über das SGs Interface zwischen einer MSC im GSM/UMTS leitungsvermittelnden Netzwerk und der Mobility Management Entity (MME) des LTE Netzwerks ausgetauscht. Der Name der Schnittstelle ist vom Gs Interface zwischen MSC und SGSN abgeleitet, da diese sehr ähnlich funktioniert. Weitere Details zu diesen 2G/3G Netzwerkkomponenten befinden sich am Ende des Buches in den Kapiteln über GSM und GPRS.

Von der MME wird die SMS Nachricht dann über NAS Signalisierungsnachrichten an das Endgerät weitergeleitet. Vom Endgerät ausgehende SMS Nachrichten

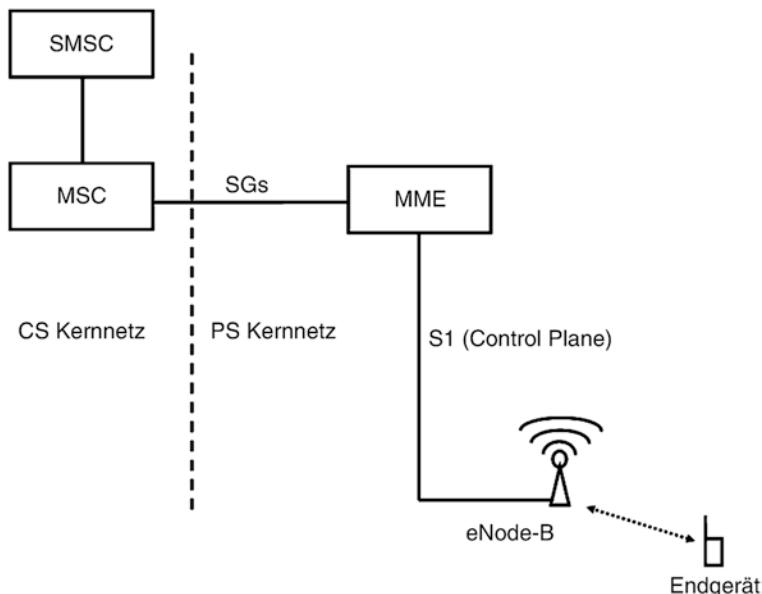


Abb. 1.30 Das SGs Interface zwischen MSC und SMS für die Signalisierung von SMS und eingehenden Gesprächen

werden entsprechend in umgekehrter Reihenfolge übertragen. Aus Ende-zu-Ende Sicht ist SMS über SGs kein IP basierter Dienst, da dieser, wie auch bei GSM und UMTS, Signalisierungsnachrichten verwendet. Indirekt wird aber dennoch das Internet Protokoll für die Übertragung verwendet, da sowohl die SGs Schnittstelle zwischen MSC und MME, als auch die S1 Schnittstelle zwischen MME und eNodeB, auf IP basieren.

Um im LTE Netzwerk SMS Nachrichten übertragen zu können, informiert das Endgerät während des Attach Prozesses das Netzwerk, dass es SMS über die SGs Schnittstelle senden und empfangen kann. Dies wird auch als „Combined Attach“ bezeichnet, da bei diesem Anmeldevorgang das Endgerät sowohl im LTE Netzwerk, als auch im leitungsvermittelten GSM/UMTS Netzwerk angemeldet wird, damit eingehende SMS Nachrichten später in das LTE Netzwerk weitergeleitet werden können.

Wenn die MME den Versand von SMS Nachrichten über das SGs Interface unterstützt, registriert sie den Teilnehmer während der Attach Prozedur im HLR, sowie in einer MSC/VLR des leitungsvermittelnden GSM/UMTS Netzwerkes. Sendet ein anderer Teilnehmer dann später eine SMS an den Teilnehmer im LTE Netzwerk, wird diese dann zunächst wie gewöhnlich an das SMS Service Center gesendet. Das SMS Service Center befragt dann das HLR nach der aktuellen MSC des Teilnehmers und leitet die SMS dann an dieses weiter. Da die MSC durch die vorherige Registrierung weiß, dass sich der Teilnehmer im LTE Netzwerk befindet, wird die SMS Nachricht dann über das SGs Interface weiterleitet.

Ist das Endgerät des Teilnehmers im RRC Connected State, kann die SMS Nachricht sofort über eine NAS Nachricht an den eNodeB und von dort an das Endgerät zugestellt werden. Ist das Endgerät im RRC Idle State, muss die MME zunächst das Endgerät über eine Paging Nachricht suchen.

1.13.2 CS Fallback

Das SGs Interface wurde nicht nur für die Übertragung von SMS Nachrichten spezifiziert, sondern auch, um Endgeräte, die sich im LTE Netzwerk befinden, über eingehende Sprachanrufe zu informieren. Da das SGs Interface nur für die Übertragung von Paging Nachrichten über eingehende Anrufe verwendet wird, muss das LTE Endgerät zur Annahme des Gesprächs, bzw. für den Aufbau eines eigenen Gesprächs nach GSM oder UMTS zurückfallen. Das Zurückfallen des Endgeräts nach GSM oder UMTS gab der CS Fallback Funktionalität ihren Namen und wurde ebenfalls in 3GPP TS 23.272 spezifiziert. Um ein Gespräch anzunehmen oder ein ausgehendes Gespräch mit dem CS Fallback Mechanismus zu führen, wird der Vorgang in zwei Phasen aufgeteilt:

Die Vorbereitungsphase (Preparation Phase)

- Beim ersten Verbindungsaufbau eines GSM/UMTS/LTE fähigen Endgerätes mit dem EPS (dem Evolved Packet System, also dem LTE Netzwerk) teilt es dem Netzwerk

mit, dass es einen „combined attach“ durchführen möchte. Dies wurde im letzten Unterkapitel zu SMS over SGs schon beschreiben.

- Die Registrierung des Endgerätes wird nicht vom Gerät selber, sondern von der MME vorgenommen, die mit dem 2G/3G MSC und dem HLR kommuniziert. Aus Sicht des MSCs verhält sich die MME wie ein SGSN. Ebenfalls aus MSC Sicht scheint sich das Endgerät bei einem 2G/3G SGSN mit einem kombinierten circuit switched/packet switched Routing Area Update zu registrieren (siehe Kapitel zu GPRS).
- Um das Endgerät im GSM/UMTS Netzwerk zu registrieren, muss die MME der MSC eine Location Area ID (LAI) übergeben, in der das Endgerät sich „theoretisch“ gerade aufhält. Da dies nur eine theoretische LAI ist, kann diese z. B. aus der Tracking Area ID abgeleitet werden, die der korrespondierende Parameter im LTE Netzwerk ist. Dies bedeutet, dass zwischen der TAI und der LAI eine Abhängigkeit gebildet wird und hat zur Folge, dass die geografischen Gebiete, die die LTE TAI und die GSM/UMTS LAI abdecken, identisch sein sollten. Nur so ist gewährleistet, dass der CS Fallback im Bedarfsfall möglichst schnell durchgeführt werden kann.

Die Ausführungsphase für ankommende Gespräche (Execution Phase – Mobile Terminated Call)

- Wenn ein Gespräch für einen Teilnehmer an der MSC eintrifft, signalisiert diese das Gespräch mit einer Paging Nachricht über das SGs Interface zur MME. Von hier wird die Paging Nachricht dann an das Endgerät weitergeleitet.
- Sollte sich das Endgerät im RRC Connected Zustand befinden, kann die MME die Paging Nachricht unmittelbar weiterleiten. Das Endgerät signalisiert dann der MME, dass es in das GSM oder UMTS Netzwerk wechseln möchte, um das Gespräch dort anzunehmen. Die MME instruiert daraufhin den eNodeB, das Endgerät entsprechend in ein GSM oder UMTS Netzwerk zu übergeben.
- Da zu diesem Zeitpunkt auch eine IP Datenübertragung im Gange sein kann, gibt es im Standard zwei Möglichkeiten, in eine andere Radionetzwerktechnologie zu wechseln. Entweder wird der Datentransfer unterbrochen oder es wird ein Handover der Paketverbindung in das UMTS Netzwerk angestoßen. Dies ist nur beim Wechsel in ein noch vorhandenes UMTS Netzwerk möglich.
- Sollte sich beim Eingang der Paging Nachricht bei der MME das Endgerät im RRC Idle Zustand befinden, muss das Endgerät von der MME zunächst mit einer eigenen LTE Paging Nachricht gesucht werden, bevor die Informationen über das eingehende Gespräch weitergeleitet werden können.
- Vor dem Wechsel von LTE nach GSM oder UMTS hat der eNodeB die Möglichkeit, Signalstärkemessungen vom Endgerät durchführen zu lassen, der Wechsel kann jedoch auch blind erfolgen.
- Nachdem das Endgerät eine GSM oder UMTS Zelle gefunden hat, nimmt es mit dieser Kontakt auf und sendet eine Paging Response Nachricht.

Die Ausführungsphase: Abgehendes Gespräch (Mobile Originated Call)

Diese Prozedur entspricht weitgehend dem zuvor beschriebenen Vorgehen bei einem eingehenden Gespräch. Der einzige Unterschied besteht darin, dass keine Paging Response Nachricht an die MSC geschickt wird. Das Endgerät muss jedoch, wie bei einem eingehenden Anruf auch, die MME zuvor informieren, dass es den Netzwerktyp wegen eines Sprachanrufs wechseln möchte. Die MME informiert dann den eNodeB, der dann entweder ein Release mit Redirect oder einen Handover veranlasst.

1.14 Network Sharing – MOCN und MORAN

Wegen des enormen finanziellen Aufwandes ein Mobilfunknetzwerk aufzubauen, erlauben die meisten nationalen Regulierungsbehörden, dass Netzbetreiber Teile ihres Netzwerks mit anderen Betreibern teilen dürfen. Üblicherweise dürfen heute Netzbetreiber ihre Standorte und Masten teilen, die eigentliche Mobilfunkausstattung und die Backhaulverbindung dürfen aber nicht geteilt werden. Am anderen Ende des Spektrums gibt es jedoch auch Länder, in denen sich Mobilfunkbetreiber das komplette Zugangsnetzwerk teilen und nur noch ein eigenes Kernnetz betreiben. Dies hat jedoch große Auswirkungen auf den Wettbewerb, da ein gemeinsam genutztes Radionetzwerk nur noch wenig Möglichkeiten zur Differenzierung bietet. So können sich Netzbetreiber gerade in wichtigen Punkten wie Abdeckung, Netzausbau und verfügbare Kapazität nicht mehr unterscheiden. In Deutschland und vielen anderen Ländern ist diese Art des Network Sharings deshalb nicht gewünscht oder zumindest auf ländliche Gebiete beschränkt. Ist vom Regulierer Netzwerk Sharing in einer oder anderen Form erlaubt, gibt es in der Praxis folgende Möglichkeiten, dies technisch umzusetzen:

1.14.1 National Roaming

National Roaming ist die erste und älteste Möglichkeit in den Standards, ein Netzwerk zu teilen. Es wird typischerweise verwendet, wenn ein neuer Netzbetreiber noch im Begriff ist, sein eigenes Netzwerk auszubauen und noch nicht über genug Flächenabdeckung verfügt. In einem solchen Szenario schließt der neue Netzbetreiber dann zeitlich begrenzt einen Vertrag mit einem Betreiber eines bereits bestehenden Netzwerkes. In dieser Zeit können dann Kunden des neuen Netzbetreibers das eigene und das des nationalen Roamingpartners verwenden. Außerdem wird National Roaming für einige Zeit nach der Fusion zweier Netzbetreibern in einem Land verwendet.

In der Praxis kommt National Roaming zur Anwendung, wenn zwei Netzbetreiber eine Übereinkunft schließen, dass Kunden eines Netzwerkes eine Kombination aus 2G, 3G, 4G oder 5G eines anderen Netzbetreibers in Teilen oder im ganzen Land verwenden dürfen. Nachdem ein solcher Vertrag unterzeichnet wurde, ist es oft notwendig, die SIM Karten der Teilnehmer entsprechend anzupassen, da z. B. üblicherweise Netze anderer

Betreiber im Heimatland in der „Forbidden PLMN“ Liste auf der SIM Karte stehen. Daten auf der SIM Karte können über einen „Over-The-Air“ (OTA) Updateprozess geändert werden, der auf nicht sichtbaren SMS Nachrichten basiert. Dieser Prozess wird z. B. auch verwendet, um bei der Ankunft im Ausland die Liste der präferierten Netzbetreiber auf der SIM Karte auf den neuesten Stand zu bringen.

In manchen Fällen teilen Netzbetreiber nicht alle Radionetzwerke. Beispielsweise können zwei Netzwerkbetreiber ein gemeinsames LTE Zugangsnetzwerk in manchen Teilen des Landes betreiben, während sie im gleichen Gebiet getrennte 2G oder 3G Netzwerk betreiben. In einem solchen Szenario würden dann die Endgeräte im 2G oder 3G Netzwerk des Heimnetzbetreibers bleiben, auch wenn ein LTE Netzwerk eines Mitbewerbers am Standort verfügbar ist, das verwendet werden könnte. Der Grund dafür ist, dass Endgeräte immer das Netzwerk des Heimnetzbetreibers bevorzugen. Da ein solches Verhalten nicht gewünscht ist, kann ein Heimnetzbetreiber eine Liste äquivalenter Netzwerke (Equivalent PLMNs) in Attach, Location Update, Routing Area Update und Tracking Area Update Nachrichten schicken. Alle Netze, die in dieser Liste mit ihrem Mobile Country Code (MCC) und Mobile Network Code (MNC) eingetragen sind, werden dann als gleichwertig zum Heimnetz des Teilnehmers vom Endgerät behandelt. Wird nur das LTE Netzwerk geteilt, müssen die Netzbetreiber in GSM SIB 2-quater und UMTS SIB 19 Nachrichten dann entsprechende LTE Reselection Parameter senden. Damit kann das Endgerät dann die Zellen des geteilten LTE Netzwerkes finden, während es in einem 2G oder 3G Netzwerk mit einem anderen MNC eingebucht ist.

1.14.2 MOCN (Multi-Operator Core Network)

Während National Roaming üblicherweise nur temporär beim Start eines neuen Netzbetreibers oder einer Fusion verwendet wird, gibt es mit der Multi-Operator Core Network (MOCN) Erweiterung in den 3GPP Standards eine Möglichkeit für das langfristige Teilen eines Radionetzwerkes. In einer normalen Netzwerkkonfiguration sendet der eNodeB den MCC und MNC eines Netzbetreibers in periodischen System Information Nachrichten an alle Endgeräte. Mit MOCN ist es möglich, dass sich mehrere Netzbetreiber alle Teile eines Mobilfunkstandortes teilen. Der eNodeB sendet dann die MCC/MNCs aller Netzbetreiber aus, die sich das Radionetzwerk teilen. Das bedeutet, dass sich alle beteiligten Netzbetreiber den Radiokanal teilen. Es ist jedoch möglich, dass die beteiligten Netzbetreiber ihr Spektrum bündeln und auch dieses gemeinsam verwenden. Welcher Netzbetreiber dann wie viel der vorhanden Kapazität verwenden darf, ist nicht spezifiziert und wird den beteiligten Partien überlassen.

In LTE wird die Liste der Kernnetze, an die der eNodeB angeschlossen ist, in System Information Block (SIB) 1 übertragen. Dies wurde schon in der ersten Version des Standards spezifiziert und somit unterstützten dies heute auch alle verwendeten Endgeräte. Für GSM wurde MOCN jedoch erst sehr spät mit 3GPP Release 11 spezifiziert. Somit unterstützten heute nicht alle verwendeten Geräte diesen Ansatz in GSM

und finden somit diese Information dort nicht. In UMTS wurde MOCN mit Release 6 eingeführt und fast alle heute noch in Benutzung befindlichen 3G Endgeräte sind somit ebenfalls MOCN fähig. Der folgende Auszug aus einer LTE SIB-1 Nachricht zeigt, wie eine MOCN Netzwerkliste von einem eNodeB an die Endgeräte übertragen wird:

```
SystemInformationBlockType1:  
[...]  
CellAccessRelatedInfo:  
    PLMN Identity List  
        PLMN-Identity  
            MCC 310  
            MNC 260  
            cellReservedForOperatorUse notReserved  
        PLMN-Identity  
            MCC 311  
            MNC 660  
            cellReservedForOperatorUse notReserved  
        trackingAreaCode 42241  
        cellIdentity 5159682  
        cellBarred notBarred  
[...]
```

Wenn sich ein Endgerät dann mit einem eNodeB verbindet, sendet es gleich beim Verbindungsaufbau in der RRC Connection Setup Complete Nachricht, mit welchem Kernnetzwerk es kommunizieren möchte:

```
rrcConnectionSetupComplete:  
rrc-TransactionIdentifier 1  
selectedPLMN-Identity 2  
registeredMME  
    mmegi 350  
    mmec 23  
dedicatedInfoNAS 1704E61DA36 [...]
```

Der eNodeB entscheidet dann anhand des „selected PLMN-Identity“ Feld, mit welchem Kernnetz, also mit welcher MME, eine Signalisierungsverbindung für diesen Teilnehmer aufgebaut werden soll. Das bedeutet, dass der eNodeB mit mehreren Kernnetzwerken gleichzeitig verbunden ist. Üblicherweise werden diese Verbindungen jedoch über die gleiche Backhaulverbindung hergestellt. An einem Router am Rand des Radionetzwerkes sind dann entsprechend beide Kernnetzwerke, also die jeweiligen MMEs und Serving Gateways der Netzbetreiber angeschlossen. Weitere Details zu diesem Konzept sind in 3GPP TS 23.251²¹ zu finden.

An dieser Stelle sei noch erwähnt, dass in UMTS das Endgerät das Kernnetzwerk in der Initial Direct Transfer Nachricht identifiziert, in der auch die erste NAS Message (Location Update Request) enthalten ist. Details hierzu finden sich in 3GPP TS 25.331, 10.2.16c²².

1.14.3 MORAN (Mobile Operator Radio Access Network)

MORAN ist ein weiterer Ansatz, mit dem sich mehrere Netzbetreiber ein Radionetzwerk teilen können. Hier teilen sich Netzbetreiber den digitalen Teil des eNodeB, also die Base Band Unit, die passiven Antennen und die Backhaulverbindung. Jeder Netzbetreiber verwendet an einem MORAN Standort jedoch sein eigenes Spektrum, das nur für seine eigenen Kunden zur Verfügung steht. Während für diesen Ansatz entsprechende Software in der BBU nötig ist, können Endgeräte nicht zwischen diesem Ansatz und zwei völlig voneinander getrennten Netzen unterscheiden, da jeder Netzbetreiber seinen MCC/MNC auf seinen eigenen Kanälen aussendet.

1.15 IPv6 im Mobilfunk

Wie auch das Festnetz sind Mobilfunknetzbetreiber von der mangelnden Verfügbarkeit an freien IPv4 Adressbereichen betroffen. Die meisten Netzbetreiber teilen heute private und somit nicht routbare IPv4 Adressen an Endgeräte zu, und verwenden Network Address Translation (NAT), um ausgehende TCP und UDP Verbindungen über wenige öffentliche IPv4 Adressen zu leiten. Diese Methode wird auch in DSL-, Kabel- und Glasfasernetzwerken verwendet. Ein großer Nachteil für Geräte hinter einem NAT ist, dass Endgeräte zwar ausgehende Verbindungen zu einem Server im Internet aufbauen können, jedoch für eingehende Verbindungen nicht erreichbar sind. Um einen Server zu betreiben, der vom Internet aus erreichbar ist, muss auf dem DSL-, Kabel- oder Fiber Router ein TCP oder UDP Port Mapping konfiguriert werden. In Mobilfunknetzwerken ist dies jedoch meist nicht möglich. Ein weiteres Problem der großen Mobilfunknetzbetreiber ist, dass es auch mit privaten IP Adressen nicht möglich ist, jedem Kunden eine eindeutige IP Adresse zuzuteilen. Selbst der private Class A Adressraum umfasst nur 16.7 Mio. eindeutige IP Adressen. Somit müssen Netzbetreiber auch im privaten Adressraum IP Adressen mehrfach zuweisen. Somit ist die direkte Kommunikation zwischen Endgeräten nicht möglich. Während dies für Smartphones und andere Endgeräte kein Problem ist, stellt dies ein logistisches Problem für die Netzbetreiber dar. Manche Netzbetreiber sind deshalb dazu übergegangen, IPv6 Adressen an Endgeräte zuzuweisen. Auch für den VoLTE Sprachdienst verwenden heute viele Netzbetreiber IPv6.

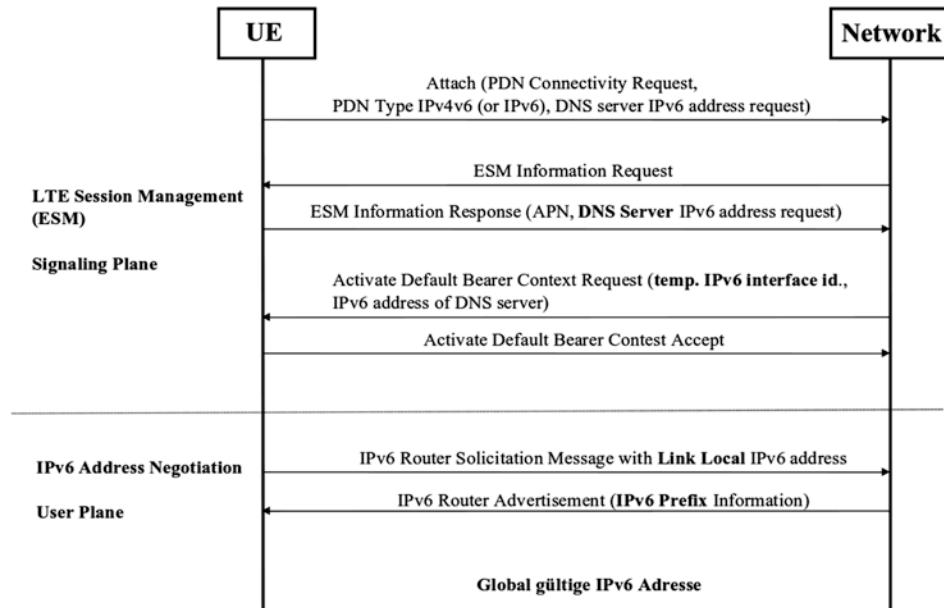


Abb. 1.31 Aufbau eines IPv6 Default Bearer

1.15.1 Das IPv6 Prefix und Interface IDs

Während beim Verbindungsauflauf das Mobilfunknetzwerk eine 32 Bit lange IPv4 zuteilt, ist dies bei IPv6 nicht der Fall. Bei IPv6 informieren Router in Netzwerk die Endgeräte über das IPv6 Prefix, das in diesem Netzwerk gültig ist. Diese periodisch oder auf Anfrage gesendeten Pakete werden als Router Advertisement Pakete bezeichnet. Jedes Endgerät erstellt dann einen festen oder zufälligen Interface Identifier und fügt diesen an das Präfix an. Dies ist dann die 128 Bit lange IPv6 Adresse des Endgeräts. Auch in Mobilfunknetzwerken wird diese Art der IPv6 Adressgenerierung verwendet.

Um ein IPv6 Prefix vom Netzwerk zu bekommen, fordert das Endgerät zunächst einen IPv6 oder IPv4v6 Default Bearer wie in Abb. 1.31 gezeigt, an. Dies geschieht mit einer PDN Connectivity Request Nachricht, die als Teil der LTE Attach Prozedur gesendet wird. In dieser Nachricht sind folgende Information Elemente (IE) enthalten:

```

PDN type: IPv4v6
[...]
Protocol or Container ID: DNS Server IPv6 Address Request (0x0003)
  
```

Das erste Information Element ist der PDN Typ, der auf IPv4, IPv6 oder IPv4v6 gesetzt werden kann. Im dritten Fall möchte das Endgerät den Bearer mit einer IPv4 und einer IPv6 Adresse verwenden. Das zweite IE ist notwendig, um eine IPv6 Adresse eines DNS Servers zu bekommen, der später für die Übersetzung von Domainnamen in IP Adressen verwendet werden kann.

Die meisten Mobilfunknetzwerke sind so konfiguriert, dass sie das Endgerät während des Attach Prozesses nach einem Access Point Name (APN) für den ersten Default Bearer fragen. Dies erfolgt mit der ESM (Session Management) Information Request Nachricht, nachdem die Authentifizierung durchgeführt und die Verschlüsselung aktiviert wurde. Das Endgerät antwortet darauf mit einer ESM Information Response Nachricht. Diese enthält unter anderem den APN und nochmals die Anfrage nach einer IPv6 DNS Server Adresse. Das Netzwerk teilt dem Endgerät dann eine IPv4 Adresse und eine IPv6 Interface ID in einer Activate Default Bearer Request Nachricht zu. Diese enthält unter anderem folgende IPv6 spezifische Information Elemente:

```
PDN Address
PDN Type: IPv4v6
PDN IPv6 Interface ID: 0002:0001:c731:f114
[...]
Container ID: DNS Server IPv6 Address
IPv6: 2a12:577:733:0:10:74:312:312
```

Die IPv6 PDN Adresse im gezeigten Ausschnitt ist nicht die IPv6 Adresse, sondern nur der Interface Identifier. Diesen verwendet das Endgerät dann zusammen mit dem Link Local Prefix im nun folgenden Stateless (zustandslosen) IPv6 Address Auto-Configuration (SLACC) Verfahren. Während das Netzwerk also in einer solchen Nachricht direkt eine IPv4 Adresse zuteilt, vergibt es hier keine IPv6 Adresse und auch kein IPv6 Prefix.

Zu diesem Zeitpunkt hat das Endgerät eine IPv4 Adresse und einen IPv6 Interface Identifier, der nun zusammen mit dem Link Local Prefix eine IPv6 Link Local Adresse bildet, mit der das Endgerät dann eine IPv6 Router Solicitation Nachricht über den nun aufgebauten Bearer schicken kann. Während die IPv6 Adressvergabe also noch nicht abgeschlossen ist, können schon IPv4 Pakete übertragen werden. Im folgenden Beispiel lautet die IPv6 Source Adresse des Endgerätes, die nur lokal, aber nicht im Internet verwendet werden kann, fe80::2:1:c731:f114. Die Zieladresse der Anfrage ist die Multicast Link Local Adresse „All Routers“, fe80::2. Das Netzwerk antwortet auf diese Anfrage von der Link Local Router Multicast Adresse fe80::5 mit einer Router Advertisement Message, die das IPv6 Prefix enthält:

```
Internet Protocol Version 6, Src: fe80::5 (fe80::5), Dst: ff02::1  
(ff02::1)  
Internet Control Message Protocol v6  
Type: Router Advertisement (134)  
Code: 0  
Checksum: 0x760f [correct]  
Current hop limit: 0  
Flags: 0x40  
Router lifetime (s): 65535  
Reachable time (ms): 0  
Retransmission timer (ms): 0  
ICMPv6 Option (Prefix information : 2a12:577:9941:f99c::/64)
```

Zu diesem Zeitpunkt hat das Endgerät dann alle Information, um eine IPv6 Adresse für sich zu generieren. Dazu verwendet es das gerade erhaltene IPv6 Prefix und den Interface Identifier, den es zuvor in der Activate Default Bearer Nachricht erhalten hat. Die IPv6 Adresse sieht somit wie folgt aus:

2a12:577:9941:f99c:0002:0001:c731:f114

In 3GPP wurde auch spezifiziert, dass das Endgerät zu diesem Zeitpunkt auch einen beliebigen Interface Identifier wählen kann. Zusammen mit der IPv6 DNS Adresse hat das Endgerät dann alles, um über IPv6 mit Servern im Internet zu kommunizieren.

Aus einer Ende-zu-Ende Sicht ist es interessant, wie die mobilfunkspezifische 3GPP Signalisierung zusammen mit dem standardisierten IPv6 Stateless Auto-Configuration Mechanismus verwendet wird. Dies macht den Prozess in der Praxis etwas komplizierter, erlaubt es jedoch, jederzeit das IPv6 Prefix und auch den Interface Identifier, z. B. aus Gründen der Privatsphäre, zu wechseln.

Wie zuvor erwähnt, wird für IPv6 kein Network Address Translation (NAT) benötigt und Endgeräte könnten somit aus dem Internet auch mit eingehenden Verbindungsanfragen erreicht werden. In der Praxis blockieren Mobilfunknetzbetreiber jedoch eingehende IPv6 Verbindungsanfragen, z. B. aus Energiespar- und Sicherheitsgründen. In den meisten Fällen ist dies für den Kunden vorteilhaft, limitiert aber die möglichen Anwendungen, wenn der Netzbetreiber keine Möglichkeit bietet, diesen Filter zu deaktivieren.

1.15.2 IPv6 und Roaming

In der Praxis ist es in manchen Fällen nicht möglich, einen IPv4v6 Bearer im internationalen Roaming aufzubauen. Wenn ein 3GPP Netzwerk alte Software nutzt oder nicht richtig konfiguriert ist, kann es die für IPv6 relevanten Parameter nicht richtig erkennen und der Verbindungsaufbau kann fehlschlagen. Das bedeutet, dass dann auch kein IPv4 Bearer vorhanden ist und der Nutzer somit keinen Internetzugang hat. Aus diesem Grund schalten Netzbetreiber die für IPv6 relevanten Parameter auf dem Roaming Interface nur frei, wenn sie sicher sind, dass das andere Netzwerk diese korrekt verarbeiten kann, oder zumindest einfach ignorieren. Auf der Endgeräteseite schalten manche Betriebssysteme IPv6 im Roaming Fall einfach ab und geben dem Nutzer auch keine Möglichkeit, dies manuell zu aktivieren. Andere Betriebssysteme, wie z. B. Android, sind oft so konfiguriert, dass sie nur einen IPv4 Bearer im Ausland aufbauen. Der Nutzer kann diese Einstellung jedoch in den Netzwerkeinstellungen ändern.

1.15.3 IPv6 und Wi-Fi Tethering

Eine heute oft genutzte Funktion eines Smartphones ist Wi-Fi Tethering, also die Nutzung des Smartphones als Wi-Fi Access Point, um anderen Geräten wie Notebooks den Zugang zum Internet zu ermöglichen. Um IPv6 zusätzlich zu IPv4 im Wi-Fi Netzwerk des Smartphones zu unterstützen, haben die meisten Hersteller die Funktion entsprechend angepasst.

Verbindet sich ein Notebook oder ein anderes Gerät über Wi-Fi Tethering mit dem Smartphone, fordert es zunächst eine IPv4 Adresse, sowie die IPv4 Adresse eines DNS Servers mit dem DHCP (Dynamic Host Configuration Protocol) an. IPv6 fähige Endgeräte erzeugen zudem mit dem folgenden Prozess eine IPv6 Adresse:

Zunächst überprüft das Notebook, ob die gewählte IPv6 Link Local Adresse, deren Interface Identifier aus der MAC Adresse des Wi-Fi Interfaces generiert wurde, im Netzwerk schon vergeben ist. Dies erfolgt über eine Neighbor Solicitation Nachricht. Wird keine Antwort empfangen, wird die Adresse als gültig angesehen und dann für die nachfolgenden Aktionen verwendet.

Im nächsten Schritt sendet dann das Notebook eine Router Solicitation Nachricht, um einen eventuell vorhandenen IPv6 Router im Wi-Fi Netzwerk zu finden. Für diese Nachricht wird die zuvor generierte IPv6 Link Local Adresse verwendet. Die Router Solicitation Nachricht wird an die IPv6 „All Routers“ Multicast Adresse gesendet. Wenn das Smartphone IPv6 Tethering implementiert hat, antwortet es an die Link Local Adresse des Notebooks mit einer Router Advertisement Nachricht, die, wie in Abb. 1.32, gezeigt, folgende Informationen enthält:

- IPv6 Prefix
- MTU Size (Maximale Paketgröße)

No.	Time	Source	Destination	Protocol	Src Prt	Dst Prt	Length	Info
21	21:01:02:00:226 ::		ff02::1:ffcb:c2f8	ICMPv6			78	Neighbor Solicitation for fe80::6e88:14ff:febc
24	21:01:03:500358 ::	fe80::6e88:14ff:febc:c2f8 ff02::2		ICMPv6			79	Router Solicitation from 6c:88:14:cb:c2:f8
28	21:01:03:639492 ::	fe80::e50:8bff:fe1:47cb	fe80::6e88:14ff:febc:c2f8	ICMPv6			142	Router Advertisement from e8:50:8b:f1:47:cb
33	21:01:03:988229 ::		ff02::1:ffcb:c2f8	ICMPv6			78	Neighbor Solicitation for 2a01:598:9941:4:ca0:6
44	21:01:04:526224 ::		ff02::1:ff03:471	ICMPv6			78	Neighbor Solicitation for 2a01:598:9941:4:ca0:9
77	21:01:05:019318 ::	2a01:598:9941:4:ca0:95e0:7	2001:67c:1560:8003::c7	NTP	123	123	110	NTP Version 4, client
81	21:01:06:025014 ::	fe80::e50:8bff:fe1:47cb ff02::1:ff03:471		ICMPv6			86	Neighbor Solicitation for 2a01:598:9941:4:ca0:9
82	21:01:06:025112 ::	2a01:598:9941:4:ca0:95e0:7	fe80::e50:8bff:fe1:47cb	ICMPv6			86	Neighbor Advertisement 2a01:598:9941:4:ca0:95e0
83	21:01:06:027008 ::	2001:67c:1560:8003::c7	2a01:598:9941:4:ca0:95e0:710f	NTP	123	123	110	NTP Version 4, server
89	21:01:06:133283 ::	fe80::6e88:14ff:febc:c2f8 ff02::1:2		DHCPv6	546	547	119	Information-request XID: 0xca601d CID: 0001000
90	21:01:06:138941 ::	fe80::e50:8bff:fe1:47cb	fe80::6e88:14ff:febc:c2f8	DHCPv6	547	546	133	Reply XID: 0xca601d CID: 0001000174e5c0465209
107	21:01:06:193668 ::	2a01:598:9941:4:ca0:95e0:7	2001:67c:1560:8003::c7	NTP	123	123	110	NTP Version 4, client
108	21:01:07:0998908 ::	2001:67c:1560:8003::c7	2a01:598:9941:4:ca0:95e0:710f	NTP	123	123	110	NTP Version 4, server
116	21:01:08:08158478 ::	fe80::e50:8bff:fe1:47cb fe80::6e88:14ff:febc:c2f8		ICMPv6			86	Neighbor Solicitation for fe80::6e88:14ff:febc
117	21:01:08:08168511 ::	fe80::6e88:14ff:febc:c2f8 fe80::e50:8bff:fe1:47cb		ICMPv6			78	Neighbor Advertisement fe80::6e88:14ff:febc:c2
123	21:01:08:0919323 ::	2a01:598:9941:4:ca0:95e0:7	2001:67c:1560:8003::c7	NTP	123	123	110	NTP Version 4, client
124	21:01:09:0999463 ::	2001:67c:1560:8003::c7	2a01:598:9941:4:ca0:95e0:710f	NTP	123	123	110	NTP Version 4, server


```
Type: Router Advertisement (134)
Code: 0
Checksum: 0x5344 [correct]
Cur hop limit: 128
▶ Flags: 0x40
Router lifetime (s): 9000
Reachable time (ms): 0
Retrans timer (ms): 0
▼ ICPNv6 Option (Prefix information : 2a01:598:9941:4:ca0::/64)
  Type: Prefix Information (3)
  Length: 4 (32 bytes)
  Prefix Length: 64
  ▶ Flag: 0x40
    Valid Lifetime: 4294967295 (Infinity)
    Preferred Lifetime: 4294967295 (Infinity)
    Reserved
    Prefix: 2a01:598:9941:4:ca0:: (2a01:598:9941:4:ca0::)
  ▶ ICPNv6 Option (MTU : 1500)
  ▶ ICPNv6 Option (Source link-layer address : e8:50:8b:f1:47:cb)
  ▶ ICPNv6 Option (Recursive DNS Server 2a01:598:7fff:0:10:74:210:210)
    Type: Recursive DNS Server (25)
    Length: 3 (24 bytes)
    Reserved
    Lifetime: 3600
    Recursive DNS Servers: 2a01:598:7fff:0:10:74:210:210 (2a01:598:7fff:0:10:74:210:210)
```

Abb. 1.32 Ein IPv6 Router Advertisement mit Wi-Fi Tethering

- Link Layer Adresse des Routers
- DNS Server Informationen

Die Übermittlung der IPv6 Adressen der DNS Server in der Router Advertisement Nachricht wird nicht von allen Endgeräten unterstützt. Solche Endgeräte setzen bei der Anfrage das „Other“ Flag und senden dann später für die DNS Server IP Adressen eine separate Anfrage.

Mit dem IPv6 Prefix, das in der Router Advertisement Nachricht enthalten ist, erstellt dann das Endgerät seine eigene IPv6 Adresse, typischerweise eine Kombination des Prefix mit der MAC Adresse seines Wi-Fi Interfaces. Danach wird eine weitere Neighbor Solicitation Nachricht mit der IPv6 Link Local Adresse gesendet. Somit wird sicher gestellt, dass kein anderes Gerät die gleiche globale IPv6 Adresse verwendet.

Um anonym zu sein, erstellen viele Endgeräte dann noch zusätzlich eine weitere globale IPv6 Adresse, die aus dem Prefix und einem zufälligen Interface Identifier besteht. Danach wird dann erneut eine Neighbor Solicitation Nachricht gesendet, um auszuschließen, dass diese Adresse schon von einem anderen Gerät im Netzwerk verwendet wird. Erhält das Endgerät keine Antwort, kann es dann mit dieser IPv6 Adresse mit Servern im Internet kommunizieren.

In der Praxis benötigt dieser Prozess etwa zwei Sekunden und dauert somit etwas länger, als über DHCP eine IPv4 Adresse zu beziehen. Man könnte hier argumentieren, dass dies eine sehr lange Zeit ist, da in anderen Bereichen viele Optimierungen spezifiziert wurden, um den LTE Verbindungsprozess um einige Millisekunden zu verkürzen. Ein Grund für diese Ineffizienz könnte sein, dass die IPv6 Konfigurationsprozedur aus einer Zeit stammt, in der Endgeräte noch hauptsächlich stationär verwendet wurden, permanent über eine Netzwerkverbindung verfügten und nur mit einem einzelnen Netzwerk verbunden waren. In einem solchen Szenario spielt natürlich die Konfigurationszeit keine Rolle. Aus Kompatibilitätsgründen wurde diese Prozedur dann auch für das IPv6 Wi-Fi Tethering übernommen.

1.15.4 IPv6-Only Kommunikation

Aktuell bieten viele Netzbetreiber ihren Mobilfunkkunden, wie in den vorigen Unterkapiteln beschrieben, eine IPv4+IPv6 Verbindung zum Internet an. Ziel ist es jedoch, die Probleme durch die begrenzte Anzahl von IPv4 Adressen durch die ausschließliche Verwendung von IPv6 zu lösen. Das bedeutet, dass Endgeräte beim LTE Attach Prozess einen IPv6-only Bearer aufbauen. Da viele Server im Internet weiterhin jedoch nur über IPv4 Adressen erreichbar sind, muss für einen IPv6-only Zugang eine Umsetzung zwischen IPv6 und IPv4 stattfinden. Dies wird als Network Address Translation 6 to 4 (NAT64) bezeichnet und ist in IETF RFC 6052²³ beschrieben. In der Praxis funktioniert diese Methode wie folgt:

Wenn ein Endgerät den DNS Server im Mobilfunknetzwerk hinter dem SGi Interface nach der IP Adresse für einen Domain Namen fragt, und dieser Server nur über eine IPv4 Adresse verfügt, erzeugt der DNS Server eine Umsetzung zu einer IPv6 Adresse und sendet diese an das mobile Endgerät zurück. Diese Prozedur wird als DNS64 bezeichnet. Das Endgerät spricht somit den Server über eine IPv6 Adresse an. Alle Datenpakete an die IPv6 Adresse werden vom Netzwerk dann an ein IPv4v6 Gateway geleitet, das dann den IPv6 Header durch einen IPv4 Header austauscht und das so erzeugte IPv4 Paket an den IPv4 Server im Internet weiterleitet. Einerseits wird für diese Prozedur ein spezieller DNS Server benötigt, der IPv4 zu IPv6 Adressen umsetzen kann, sowie ein IPv4v6 Gateway. Andererseits ist die Prozedur für das Endgerät völlig transparent, es sind also keine Änderungen im Protokollstack notwendig.

Besonders auf Geräten, die über ein Smartphone und Wi-Fi Tethering eine Verbindung ins Internet herstellen, kann IPv6 jedoch deaktiviert sein, oder manche Applikation verlangen explizit nach einer IPv4 Adresse. Für dieser Fälle wurde der 464XLAT Service in RFC 6877²⁴ standardisiert. Zusätzlich zu NAT64 im Netzwerk muss hier der 464XLAT Dienst als Teil des Netzwerkstacks von Smartphones oder anderen mobilen Geräten vorhanden sein. Dieser terminiert IPv4 Verbindungen direkt auf dem mobilen Endgerät und leitet alle IPv4 Pakete als IPv6 Pakete über die Mobilfunkverbindung zum NAT64 Gateway weiter. Dort werden dann aus den IPv6 Paketen

wieder IPv4 Pakete. In der Gegenrichtung wird entsprechend in umgekehrter Weise vorgegangen. Somit ist es möglich, dass Applikation oder Geräte, die auf IPv4 limitiert sind, trotzdem über einen IPv6-only Bearer eine Verbindung zu Servern im Internet aufbauen können. Alle aktuellen Android Endgeräte unterstützen heute 464XLAT und auch die Anzahl der Netzbetreiber, die IPv6-only Bearer anbieten nimmt ständig zu. Beispiele sind T-Mobile in den USA, die Deutsche Telekom und der Mobilfunkbetreiber Bouygues Telecom in Frankreich²⁵.

1.16 Network Function Virtualization

Ein wichtiges Thema in der Weiterentwicklung des mobilen Kernnetzwerkes, also der Komponenten wie der MME, das S-GW, P-GW und der IMS Komponenten ist die „Virtualisierung“. Unter dem Begriff ‚Network Function Virtualization‘ (NFV) werden mehrere Ansätze kombiniert, die nun in diesem Unterkapitel beschrieben werden. Um die Konzepte möglichst konkret zu beschreiben, folgt zunächst eine Beschreibung, wie die Virtualisierung auf privaten und geschäftlichen PCs verwendet werden kann. Danach wird gezeigt, wie die Virtualisierung in Rechenzentren und der Cloud heute verwendet wird. Von dort ist es dann nur ein kleiner Schritt Richtung NFV im Mobilfunknetz. Und schließlich wird in diesem Unterkapitel auch besprochen, was es Software-Defined Networking auf sich hat.

1.16.1 Virtualisierung auf dem Desktop

In den vergangenen Jahren erlebten Desktop und Notebook PCs einen sehr starken Leistungsschub. Heute verfügen diese Geräte über eine Hauptspeichergröße, die oftmals weit über das hinausgeht, was ein einzelner Anwender für die meisten seiner Aufgaben benötigt. Auch bei der SSD Kapazität sind nach oben heute nur wenig Grenzen gesetzt. Bei vielen Anwendern sind somit Prozessor und Speicher nur wenig ausgelastet. Auch im Server Bereich fand eine ähnliche Entwicklung statt. Wird ein Server z. B. nur als Fileserver verwendet, wird die Hardware oftmals nicht optimal genutzt.

Um die Hardware besser zu nutzen, wurde das Konzept der Virtualisierung eingeführt. Ziel ist es, eine virtuelle Umgebung auf einem PC, oder generell einem Computer, zu schaffen, die sich wie ein ‚richtiger‘ PC für jegliche Software, die darin läuft, verhält. Diese Simulation ist so perfekt, dass auch ganze Betriebssysteme in einer solchen Umgebung ausgeführt werden können und diese keinen Unterschied zwischen physischer Hardware (Bare Metal) und simulierter (virtueller) Hardware erkennen. Das Programm, das diese virtuelle Maschine erzeugt und kontrolliert wird Hypervisor genannt. Die Idee hinter einem Hypervisor ist einfach: Er kann als eine Software angesehen werden, die alle Komponenten eines PCs simuliert. Programmen, die in dieser virtuellen Maschine ausgeführt werden, erlaubt der Hypervisor keinen direkten

Zugriff auf die physisch vorhandene Hardware. Umgesetzt wird dieser Ansatz, in dem der Hypervisor der Software in der virtuellen Umgebung nur Zugang zu bestimmten CPU Instruktionen gibt. Wann immer ein Programm in der virtuellen Umgebung dann über die CPU mit einer Input/Output Instruktion auf ein physisches Gerät zugreifen möchte, unterbricht die CPU das laufende Programm und übergibt die Kontrolle an den Hypervisor. Dies soll mit einem abstrahierten Beispiel weiter verdeutlicht werden. Wird eine CPU Instruktion aufgerufen, mit der ein Datenblock vom Speicher zu einem physischen Gerät wie z. B. einer Festplatte kopiert werden soll, übernimmt der Hypervisor die Kontrolle und schreibt den Datenblock in eine Datei auf einer echten Festplatte, die die virtuelle Festplatte repräsentiert. Danach gibt der Hypervisor die Kontrolle an die Software in der virtuellen Maschine zurück. Somit bekommt diese gar nicht mit, dass die Daten gar nicht auf eine echte Festplatte geschrieben werden. Auch die Interaktion mit anderer physischer Hardware, wie z. B. mit der Grafikkarte und USB Geräten wie Tastatur und Maus, werden in gleicher Weise simuliert, also durch Unterbrechung des Programmablaufs durch die CPU, sobald I/O Instruktionen verwendet werden, die lesend oder schreibend auf physische Geräte zugreifen wollen.

1.16.2 Nutzung eines Betriebssystems in einer virtuellen Maschine

Auf dem PC gibt es eine Reihe unterschiedlicher Anwendungen für eine solche virtuelle Umgebung. Zum einen kann in einer solchen Umgebung ein komplettes Betriebssystem ohne Modifikationen ausgeführt werden, das sonst auf echter Hardware läuft. Ein Anwendungsbeispiel dafür wäre z. B. ein Nutzer, der auf seinem Windows PC gelegentlich neue Software testen möchte. Diese Software möchte er jedoch nicht auf dem ‚echten‘ Betriebssystem installieren, da das Programm vielleicht nicht dauerhaft installiert bleiben soll. In diesem Szenario wird in der virtuellen Maschine ein weiteres Windows Betriebssystem ausgeführt, das als Gastbetriebssystem (Guest Operating System) bezeichnet wird. In diesem können die Installation, Konfiguration und Betrieb neuer Software getestet werden, ohne dass dies Auswirkungen auf das ‚echte‘ Betriebssystem hat, das als Hostbetriebssystem (Host Operating System) bezeichnet wird.

In einem anderen Szenario verwendet ein Nutzer eine Linux Distribution wie Ubuntu als sein normales Betriebssystem, also als Host Operating System. Ab und an möchte der Nutzer jedoch auch ein Programm verwenden, das es nur für Windows gibt. Während es durchaus Möglichkeiten gibt, Windowsprogramme direkt unter Linux auszuführen, bietet die Nutzung einer virtuellen Maschine, in der Windows als Betriebssystem installiert ist, jedoch einige Vorteile: In diesem und auch im ersten Beispiel läuft das Gastbetriebssystem in einem Fenster auf dem Hostbetriebssystem. Abb. 1.33 zeigt, wie Windows als Gast in einem Fenster auf einem Linux Host läuft. Das Gastbetriebssystem weiß nicht, dass sein Bildschirm in einem Fenster dargestellt wird, da es aus seiner Sicht die grafische Oberfläche über die (simulierte) Grafikkarte auf einem Bildschirm anzeigt. Das Gastbetriebssystem sieht keinen Unterschied zwischen echter und virtueller

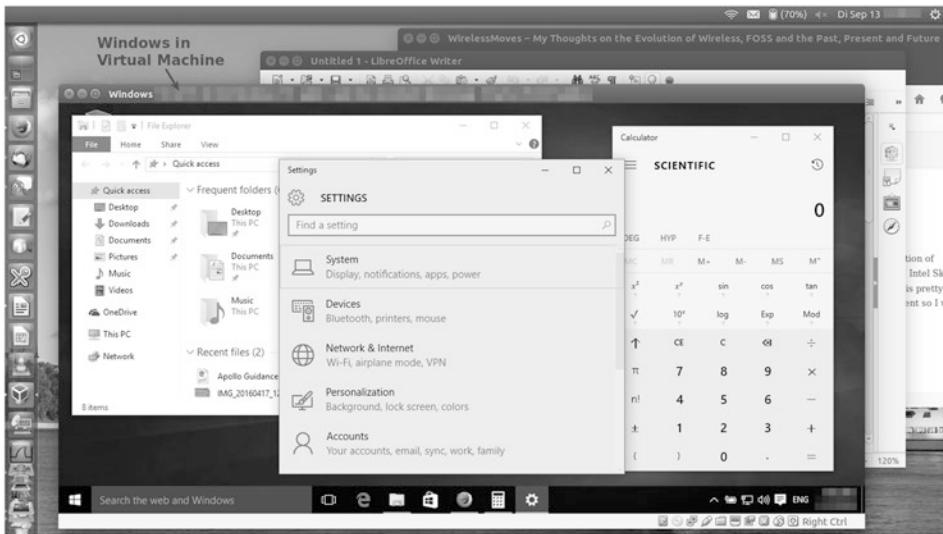


Abb. 1.33 Ein Linux Host mit einem Fenster, in dem Windows als Gast in einer virtuellen Maschine ausgeführt wird

Hardware. Schreibt das Gastbetriebssystem auf die (simulierte) Festplatte, übersetzt der Hypervisor den Schreibvorgang und modifiziert eine große Datei auf der physischen Festplatte. Auch hiervon bekommt das Gastbetriebssystem nichts mit.

1.16.3 Das gleichzeitige Ausführen mehrerer virtueller Maschinen

Die zweite interessante Funktionalität eines Hypervisors ist, mehrere virtuelle Maschinen gleichzeitig auf einem physischen Gerät ausführen zu können. Dies soll an folgendem Beispiel verdeutlicht werden: Ein Nutzer verwendet Linux als Host Betriebssystem auf seinem Notebook und nutzt dieses für die meisten seiner täglichen Aufgaben. Ab und an möchte er jedoch neue Dinge ausprobieren, die unter Umständen die Konfiguration und Stabilität des Host Betriebssystems beeinträchtigen könnten. Deswegen nutzt er eine virtuelle Maschine mit Ubuntu als Gast Betriebssystem und kann dort ohne Gefahr für den Host experimentieren. Zusätzlich läuft gleichzeitig eine weitere VM mit Windows als Gast OS. Somit laufen auf dem Notebook drei Betriebssysteme gleichzeitig: Ubuntu als Host OS, ein weiteres Ubuntu als Gast OS in der ersten VM und Windows als Gast OS in der zweiten VM. Dies benötigt natürlich entsprechend viel Hauptspeicher (RAM) sowie Speicherplatz auf der Festplatte, und alle Betriebssysteme teilen sich die Leistung der CPU. Wird eine VM für eine Weile nicht benötigt, kann das dazugehörige Fenster einfach minimiert werden, das Gast OS muss nicht beendet

werden. Wird im Gast OS gerade kein Programm ausgeführt, benötigt die VM auch nur wenig CPU Leistung. Aus Sicht des Gast OS ist das minimieren des Fensters nicht sichtbar.

1.16.4 Snapshots von Virtuellen Maschinen

Ein weiterer Vorteil von virtuellen Maschinen ist, dass mit der VM Management Software Abbilder (Snapshots) erstellt werden können. Im einfachsten Fall kann ein Snapshot erzeugt werden, während die virtuelle Maschine nicht in Betrieb ist. In diesem Fall wird die Datei, die die Festplatte der VM repräsentiert, in ihrem Zustand eingefroren und alle zukünftigen Änderungen werden in eine neue Datei geschrieben. Um dann zum ursprünglichen Zustand zurückzukehren, ist es lediglich notwendig, die neue Datei, in der die Änderungen vermerkt werden, zu löschen. Es ist sogar möglich, einen Snapshot zu erzeugen, während eine virtuelle Maschine ausgeführt wird. Zusätzlich zur neuen Datei, in der die Änderungen aufgezeichnet werden, wird der Zustand aller simulierten Geräte inklusive der CPU und aller Register aufgezeichnet und ein Abbild des Hauptspeichers auf der Festplatte des Hosts gespeichert. Nachdem dies abgeschlossen ist, läuft die virtuelle Maschine weiter. Aus Sicht der VM hat sich durch das Erstellen des Snapshots nichts geändert, da dieser vom Hypervisor außerhalb der virtuellen Umgebung erstellt wurde. Um später zum ursprünglichen Zustand zurückzukehren, wird einfach die Datei, die alle Änderungen enthält, verworfen. Danach wird das Abbild des Hauptspeichers wieder in den Hauptspeicher geladen und der Zustand aller simulierter Geräte inklusive CPU und Register werden wiederhergestellt. Schließlich wird die virtuelle Maschine wieder aktiviert und alle Programme, die zum Zeitpunkt der Erstellung des Snapshots ausgeführt wurden, laufen an exakt der Stelle weiter, an der der Snapshot erstellt wurde. Beispielsweise ist der Mauszeiger zurück an der Stelle, an der dieser zum Zeitpunkt der Snapshot Erstellung war, Musik die abgespielt wurde, spielt an der ursprünglichen Stelle weiter, etc. Aus Sicht des Gast OS hat sich nichts geändert, es kann nicht einmal erkennen, dass der Snapshot gerade wiederhergestellt wurde. Einzig das Datum und die Uhrzeit sind natürlich nicht mehr korrekt, da dieses ebenfalls vom Zeitpunkt der Snapshoterstellung stammt.

1.16.5 Klonen einer Virtuellen Maschine

Da virtuelle Maschinen auf dem Host im ausgeschalteten Zustand nur aus einer Konfigurationsdatei und einem Festplattenimage bestehen, ist es sehr einfach, Klone anzufertigen, also exakte 1:1 Kopien. Dazu muss lediglich die Konfigurationsdatei und das Festplattenimage kopiert werden. Die Kopie kann dann auf dem gleichen Host, oder auch auf einem anderen PC oder Server im Hypervisor in einem neuen Profil gestartet werden. Verwendet der andere PC oder Server den gleichen Prozessor, wird das

Gastbetriebssystem keinen Unterschied bemerken. Wird eine andere CPU verwendet, z. B. ein neueres Modell, wird dies vom Gastsystem bemerkt, da der Hypervisor ja den direkten Zugriff auf die CPU zulässt, solange keine Input/Output Instruktionen ausgeführt werden, die direkt auf die Hardware zugreifen.

1.16.6 Virtualisierung in Rechenzentren in der Cloud

Die Hauptanwendung der Virtualisierung ist heute zweifellos in privaten und öffentlichen Rechenzentren zu finden. In diesem Zusammenhang wird oft von ‚Cloud Computing‘ gesprochen. Eine Private Cloud sind Rechenzentren einer Firma, die für eigene Zwecke verwendet werden. Eine Public Cloud hingegen sind Rechenzentren einer Firma, die physische und virtuelle Server vermietet, sowie andere Dienstleistungen für externe Firmen und auch Privatpersonen zur Verfügung stellt. Firmen, die Public Clouds betreiben, sind zum Beispiel Amazon Web Services (AWS), Rackspace, Google Cloud, Microsoft Azure, Hetzner in Deutschland, und viele mehr. Server, die von Public Cloud Unternehmen betrieben werden, finden heute überall Anwendung, egal ob es sich nur um eine kleine Webseite handelt oder für große Videostreamingdienste mit hunderten Millionen Kunden. Dafür gibt es verschiedene Gründe. Kleine und mittelständische Betriebe können ihre Produkte und Dienstleistungen von einem Rechenzentrum aus anbieten, das eine sehr breitbandige und ausfallsichere Verbindung mit dem Internet hat. Dies ist oft am Standort einer Firma oder bei einer Privatperson nicht gegeben. Auch muss keine Hardware angeschafft werden, da sämtliche Server und auch Speicherplatz für Daten nur gemietet werden. Abgesehen von Sicherheits- und Datenschutzaspekten, die hier nicht weiter betrachtet werden sollen, können diese virtuellen Server und Speicherplatz in gleicher Weise wie lokale physische Server genutzt und administriert werden. Die meisten Server werden heute nicht über eine grafische Benutzeroberfläche administriert, sondern über die Kommandozeile und das Netzwerk. Somit macht es keinen Unterschied, ob ein Systemadministrator einen Ubuntu Server im lokalen Netzwerk verwaltet, oder einen Ubuntu Server, der in einem Public Cloud Rechenzentrum steht und über das Internet angebunden ist. Üblicherweise werden cloud-basierte Server nicht direkt auf der physischen Hardware ausgeführt, sondern in virtuellen Maschinen. Leistungsstarke Server in Rechenzentren haben heute speziell für diesen Einsatzzweck optimierte Hardware und CPUs mit vielen dutzend Rechenkernen und einer Hauptspeichergröße von mehreren hundert Gigabyte. Somit können auf einem einzigen Server dutzende virtuelle Maschinen gleichzeitig ausgeführt werden. Ressourcen werden somit effizient genutzt, da nicht alle Server jederzeit alle ihnen zugeteilten Ressourcen wie Rechenzeit und Speicherplatz benötigen. Es wird in diesem Zusammenhang auch von Überprovisionierung gesprochen, da z. B. CPU Kerne nicht dediziert einer einzelnen virtuellen Maschine zugeteilt werden.

Virtuelle Maschinen können auch zwischen physischen Servern während ihres Betriebs verschoben werden. Das kann z. B. notwendig sein, wenn ein physischer Server

in Überlast gerät, wenn z. B. viele virtuelle Maschinen gleichzeitig viel Rechenleistung benötigen. Somit steht dann zu wenig CPU Leistung pro virtueller Maschine zur Verfügung. Um eine aktive VM von einem physischen Server auf einen anderen zu transferieren, wird zur Laufzeit der komplette Inhalt des RAMs der virtuellen Maschine von einem physischen Server zu einem anderen physischen Server übertragen. Da die VM während dieses Transfers weiterläuft, ändern sich natürlich auch weiterhin Speicherbereiche, die schon kopiert wurden. Der Hypervisor muss sich also merken, welche RAM Bereiche sich während dieses Prozesses ändern. An einem gewissen Punkt hält der Hypervisor dann die virtuelle Maschine an und überträgt den Teil des Hauptspeichers, der sich noch geändert hat. Danach wird der virtuelle Zustand der VM noch an das Zielsystem übertragen, also z. B. der Inhalt der CPU Register, sowie der Zustand der simulierten Hardware. Danach befindet sich auf dem Zielserver eine exakte Kopie der virtuellen Maschine. Der Hypervisor auf dem Zielserver kann dann die virtuelle Maschine wieder aktivieren. Diese läuft daraufhin an exakt der Stelle weiter, an der sie auf dem ursprünglichen Server unterbrochen wurde. Natürlich ist es wichtig, die Umschaltzeit so kurz wie möglich zu halten. In der Praxis können Zeiten von nur wenigen Sekunden erreicht werden. Der Transfer von virtuellen Maschinen zwischen physischen Servern kann auch für das sogenannte Load-Balancing verwendet werden, sowie für den Transfer von VMs, wenn ein physischer Server für Wartungszwecke oder für einen Austausch heruntergefahren werden soll.

1.16.7 Administration von Virtuellen Maschinen in der Cloud

Ein weiterer wichtiger Aspekt der Virtualisierung ist, wie VMs in der Cloud administriert werden. Auf einem Desktop oder Notebook PC werden Hypervisor wie z. B. Virtualbox verwendet. Diese haben ihr eigenes grafisches Benutzerinterface, über das virtuelle Maschinen gestartet, angehalten, erzeugt und konfiguriert werden können. Für virtuelle Maschinen in Rechenzentren wird hingegen ein etwas anderer Ansatz gewählt. AWS, Google, Microsoft, Hetzner, etc., stellen ihren Kunden für die Administration eine webbasierte Oberfläche zur Verfügung. Um virtuelle Maschinen dort zu mieten, reicht es meist, einen Benutzeraccount anzulegen. Danach kann dann mit vorkonfigurierten VM Abbildern, in denen ein Betriebssystem wie Ubuntu Linux oder auch Windows bereits installiert ist, eine virtuelle Maschine angelegt werden. Weitere wichtige Parameter bei der Konfiguration der VM sind die gewünschte Größe des Hauptspeichers und wie viel Festplattenplatz der VM zur Verfügung stehen soll. Mit diesen Informationen wird dann eine neue VM Instanz erzeugt, die nach wenigen Sekunden einsatzbereit ist. Über die SSH Remote Shell kann der Administrator sich dann über das Netzwerk in die virtuelle Maschine einloggen. Während viele Cloud Provider proprietäre Software verwenden, gibt es auch Firmen, die die Open Source Alternative OpenStack verwenden. Beispiele sind die Firmen Rackspace und OVH. OpenStack ist auch ideal für Firmen, um ihre

virtuellen Server und Ressourcen in ihrem eigenen Rechenzentrum, also in ihrer eigenen private Cloud, zu verwalten.

1.16.8 Network Function Virtualization

Nachdem sich die Virtualisierung im Internet durchgesetzt hatte, begannen auch Netz-ausrüster und Netzbetreiber damit, ihre Netzkomponenten zu virtualisieren. Der dafür eingeführte Begriff lautet Network Function Virtualization (NFV) und das Konzept soll an folgendem praktischen Beispiel verdeutlicht werden: Das Voice over LTE (VoLTE) System, das später in einem eigenen Kapitel beschrieben wird, benötigt eine beträchtliche Anzahl an logischen Netzwerkkomponenten. Diese werden als Call Session Control Functions (CSCFs) bezeichnet und sind Teil des IP Multimedia Subsystems (IMS). Diese Netzwerkfunktionen werden üblicherweise vom Hersteller mit eigener Server Hardware geliefert. Alle CSCFs sind jedoch Netzwerkfunktionen, die keine spezifische Hardware benötigen. Somit ist es auch nicht notwendig, diese auf speziellen und herstellerspezifischen Servern auszuführen. Die Idee der Network Function Virtualization (NFV) ist nun, die Software von der Hardware zu trennen und die CSCF Software in virtuellen Maschinen auszuführen. Wie bereits oben erwähnt, ergeben sich daraus einige Vorteile. Durch die Trennung von Hardware und Software können Netzbetreiber nun Hardware und Software von unterschiedlichen Herstellern erwerben. Die CSCF Software wird also von einem Hersteller erworben, der sich auf die Entwicklung dieser Software spezialisiert hat, während standardisierte $\times 86$ basierte Server von einem anderen Hersteller eingekauft werden. Der große Vorteil hierbei ist, dass standardisierte Server in großen Stückzahlen produziert werden und es einen großen Wettbewerb in diesem Bereich zwischen Firmen wie HP, Dell und vielen anderen gibt. Dies senkt den Preis und steigert die Leistungsfähigkeit der Server. Ein weiterer positiver Effekt ist, dass Kapazitätserweiterungen des IMS Systems nun einfach durch die Installation und Inbetriebnahme weiterer Server möglich wird, auf denen dann neue CSCF VM Instanzen gestartet werden. Auf einem leistungsfähigen Server kann außerdem eine große Anzahl an CSCF Instanzen gleichzeitig betrieben werden und es ist ebenso möglich, VMs mit Software für unterschiedliche Aufgaben zu betreiben. Und schließlich vereinfacht die Nutzung einheitlicher Hardware für unterschiedliche Funktionen im Netzwerk auch die Administration des Netzwerkes.

Eine weitere Netzfunktion, die sehr einfach virtualisiert werden kann, ist die LTE Mobility Management Entity (MME). Wie bereits in diesem Kapitel beschrieben, kommuniziert diese Komponente mit den mobilen Endgeräten über die LTE Basisstationen und ist für das Mobilitäts- und Session Management zuständig. Schaltet ein Nutzer sein Endgerät ein, kommuniziert die MME nach der Authentifizierung auch mit anderen Netzwerkkomponenten, um für das Endgerät einen Nutzdatentunnel einzurichten. Auch im weiteren Verlauf tauscht die MME mit dem Endgerät und anderen Netzkomponenten Signalisierungsnachrichten aus, um die Verbindung des Teilnehmers

auch bei Wechseln zwischen den Mobilfunkzellen mit dem Internet aufrechtzuerhalten. Sämtliche Signalisierung verwendet dabei das Internet Protokoll und es ist somit keine spezielle Hardware oder ein spezielles Netzwerkinterface nötig, um mit der Außenwelt zu kommunizieren. Wichtig dabei ist, dass die MME nur Managementaufgaben hat. Alle Nutzdatenpakete, die zwischen dem Endgerät und dem Internet ausgetauscht werden, werden nur durch das S-GW und das P-GW geleitet. Die MME ist nicht Teil des Übertragungswegs und hat somit keine Routing Funktionalität. Während die für eine MME benötigte Bandbreite deshalb recht gering ist, wird für jede Nachricht ein hoher Rechenaufwand für die Verarbeitung benötigt. Eine MME eignet sich somit besonders für die Virtualisierung auf standardisierter Hardware.

1.16.9 Router Virtualisierung

Zusätzlich zu Netzkomponenten wie den MMEs und CSCFs, die für die Signalisierung verwendet werden, enthält das Netzwerk auch Komponenten, die für das Routing von Paketen zuständig sind. Diese untersuchen jedes eingehende Paket, entscheiden auf welchem Netzwerkinterface es weitergeschickt werden soll und modifizieren in manchen Fällen Pakete vor der Weiterleitung. Dies sind z. B. das LTE Serving Gateway (SGW) und das Packet Data Network Gateway (PDN-GW). Diese werden von den MMEs konfiguriert, Daten zwischen Nutzern und dem Internet in Nutzdatentunneln weiterzuleiten und somit die Mobilität des Teilnehmers zu ermöglichen. Aus Gründen der Effizienz wird in traditionellen Routern ein Teil des Routing Algorithmus nicht in Software, sondern in speziell dafür entwickelter Hardware (Application-specific Integrated Circuits, ASICs) implementiert. Dieser Ansatz kann jedoch nur schwer virtualisiert werden, da die ASICs herstellerspezifisch sind und nicht für eine Virtualisierung, also gleichzeitige Verwendung aus mehreren virtuellen Maschinen entwickelt wurden. Es wird deshalb ein Ansatz benötigt, mit dem das Routing von IP Paketen auch aus virtuellen Maschinen effizient durchgeführt werden kann. In der Praxis werden dafür heute Netzwerkkarten verwendet, die speziell für die Virtualisierung entwickelt wurden und sich in mehrere logische Netzwerkkarten unterteilen lassen. Aus einer physischen Netzwerkkarte werden somit viele logische Netzwerkkarten, die vom Hypervisor an einzelne virtuelle Maschinen zugeteilt werden können. Dieses Verfahren wird auch als Single-Root IO-Virtualization (SR-IOV)²⁶ bezeichnet. Auf höheren Schichten des Protokollstacks wird dann z. B. das Data Plane Development Kit (DPDK) verwendet, um die IP Paketanalyse und Modifikation auch effizient in Software durchführen zu können²⁷. Auf diese Weise ist es möglich, auch IP Routing ohne dedizierte ASICs mit normalen Servern durchzuführen.

1.16.10 Software-Defined Networking

Der Begriff Software-Defined Networking (SDN) ist ein Begriff, der oft im Zusammenhang mit Network Function Virtualization verwendet wird, jedoch davon völlig unabhängig ist. Um IP Pakete im Internet zu übertragen, werden Router benötigt. Jeder Router zwischen Quelle und Ziel eines IP Paketes entscheidet anhand des IP Headers, an welches ausgehende Netzwerkinterface ein Paket weitergeleitet werden soll. Dieser Prozess beginnt im DSL/Wi-Fi/Kabelrouter, der eingehende Pakete entweder im lokalen Netzwerk verteilt, oder über das WAN (Wide Area Network) Interface ins Internet weiterleitet. Router im Internet haben üblicherweise weit mehr Netzwerkschnittstellen und die Entscheidung, wohin ein IP Paket weitergeleitet werden soll, ist komplexer. Hier kommen IP Routing Tabellen zum Einsatz, die eine Beziehung von IP Adressbereichen und ausgehenden Netzinterfaces abbilden. Routingtabellen sind nicht statisch, sondern ändern sich, wenn ein Interface plötzlich nicht mehr verfügbar ist. Dies ist der Fall, falls z. B. wegen eines Hardwareproblems eine Netzwerkkarte oder Verbindung ausfällt, oder weil neue Routen zu einem Ziel bekannt werden. Noch häufiger ist der Fall, dass Netzbereiche über die Welt verteilt neu hinzukommen oder wegfallen. Für die globale Verwaltung der Netze, die zusammen das Internet darstellen, wird das Border Gateway Protocol (BGP) verwendet. Mit diesem Protokoll informieren sich die Router an den Übergängen von Netzwerken gegenseitig, welche Netze über eine Schnittstelle erreicht werden können. Empfängt ein Router ein BGP Update, muss dieser dann entscheiden, ob er seine Routing Tabelle anpassen muss. Falls dies geschieht, sendet auch dieser Router dann ein BGP Update an alle anderen erreichbaren Router. Jeder Router entscheidet dann anhand der eingehenden BGP Updates, wie seine Routing Tabelle angepasst werden muss. Für Netzwerkadministratoren bedeutet dies, dass sie ein sehr gutes Verständnis des Zustands jedes Routers in ihrem Netzwerk haben müssen. Erschwerend kommt hinzu, dass Router von unterschiedlichen Herstellern natürlich auch unterschiedliche Administrationstools verwenden. Ziel des Software-Defined Networking (SDN) ist es nun, die proprietären Konfigurationstools und individuellen Änderungen der Routingtabellen in jedem Router nur noch an einer zentralen Stelle durchzuführen und dann an die Router im Netzwerk zu verteilen. Die Aufgabe, Routingtabellen zu erstellen wird somit zentralisiert. Die Router erhalten dann Updates direkt vom SDN Server und kümmern sich nur noch um das Weiterleiten von IP Paketen. Router, die SDN unterstützen, implementieren dazu eine standardisierte Schnittstelle, die in der OpenFlow Spezifikation der Open Network Foundation (ONF) beschrieben ist. Das zentralisierte Management der Router in einem Netzwerk wird somit unabhängig vom Hersteller der Router Hardware.

1.17 Machine Type Communication (MTC) und das Internet der Dinge

Das ursprüngliche Ziel der LTE Entwicklung war es, eine neue Mobilfunktechnologie für den schnellen Internetzugang zu entwickeln, deren Möglichkeiten weit über die Fähigkeiten von UMTS hinausgehen sollte. Während dieses Ziel mittlerweile erreicht wurde, entwickelten sich die Anforderungen weiter und es wurde klar, dass die LTE Netzwerkarchitektur für andere Anwendungsgebiete wie ‚Wearables‘, Sensoren, Heimvernetzung, Automatisierung, etc., also für das Internet of Things (IoT) nicht optimal war. IoT Geräte sind oft sehr klein, übertragen nur sehr wenige Daten und sind mit sehr kleinen Batterien ausgestattet und sollen trotzdem für Wochen, Monate oder sogar Jahre in Betrieb bleiben. Außerdem werden solche Geräte oft an Orten verwendet, die nur schwer erreicht werden können, wie z. B. in Kellern oder im industriellen Umfeld. Im Heimbereich können IoT Geräte auf das lokale Netzwerk oder einen zentralen IoT Hub zugreifen, über den die Geräte dann über Wifi, Kabel, DSL oder Glasfaser mit dem Internet kommunizieren können. Es gibt jedoch viele Szenarien, in denen es keine lokalen Netzwerke gibt und somit direkt ein Mobilfunknetzwerk für die Verbindung zum Internet die einzige Alternative ist.

Zwar wird GSM heute noch immer für viele Zwecke genutzt, es ist jedoch absehbar, dass viele Netzbetreiber diese nun schon seit den frühen 1990er-Jahren eingesetzte Technologie abschalten wollen. Somit kommen in Zukunft für Anwendungsfälle, in denen kein lokales Netzwerk verwendet werden kann, entweder proprietäre Technologien oder LTE als Zugangsnetzwerk für IoT Geräte in Frage. LTE war jedoch nie für extrem stromsparende Endgeräte konzipiert und auch nicht für eine sehr große Anzahl sehr günstiger Endgeräte pro Zelle, die nur sehr selten Daten versenden und empfangen. Aus diesem Grund hat 3GPP für das Internet der Dinge eine Anzahl LTE Erweiterungen spezifiziert, die von einfachen Modifikationen bestehender Funktionalitäten bis zu einer neuen Luftschnittstelle (Air Interface) reichen. Folgende Ziele wurden bei der Spezifikation verfolgt:

- Günstige Radiomodule, die \$5 oder weniger kosten.
- Tausende Endgeräte pro Zelle, die nur ein paar Bytes pro Tag übertragen.
- Sehr energiesparende Module mit einer Batterilaufzeit von bis zu 10 Jahren für Endgeräte, die nur sehr wenig Daten pro Tag übertragen.
- Effektive Unterstützung von Geräten mit sehr niedrigen Datenraten von wenigen hundert Kilobits pro Sekunde, im Ausgleich für eine sehr einfache Technologie, die billig produziert werden kann und einer sehr hohen Empfangsempfindlichkeit für den Einsatz an Orten mit sehr schlechtem Empfang.

In der Praxis kann ein Radionetzwerk nicht alle potenziellen Nutzungsszenarien abdecken. Manche IoT Applikationen übertragen Daten recht oft und mit einer Geschwindigkeit von mehreren hundert Kilobits pro Sekunde, können dafür jedoch

Kompromisse bei Stromverbrauch und Netzwerkabdeckung machen. Andere IoT Endgeräte, wie z. B. Stromzähler oder Heizungssensoren, übertragen nur wenige Bytes pro Tag, müssen aber extrem stromsparend sein und auch an Orten funktionieren, an denen 10 oder 20 MHz breite Kanäle nicht vordringen können. Um diese sehr unterschiedlichen Anforderungen abzudecken, wurden mehrere, voneinander unabhängige, Erweiterungen spezifiziert und es wurden folgende neue Endgerätekategorien eingeführt:

- LTE Category 1: Für Endgeräte mit Geschwindigkeiten bis 10 Mbit/s.
- LTE Category 0: Für Endgeräte mit Geschwindigkeiten bis zu 1 Mbit/s.
- LTE Category M1: Für Endgeräte mit Geschwindigkeiten bis zu 1 Mbit/s mit zusätzlichen Stromsparerweiterungen.
- LTE Category NB1: Für Narrow-Band IoT (NB-IoT) Anwendungen und Geräte und einer maximalen Geschwindigkeit von wenigen Kilobit pro Sekunde, möglichst geringer Stromverbrauch und erweiterte Netzabdeckung durch ‚Indoor Coverage Extensions‘.

Alle Erweiterungen, die im Laufe der Zeit spezifiziert wurden, haben gemein, dass keine neuen Netzwerkinfrastruktur benötigt wird, d. h. die existierende LTE Infrastruktur eines Netzbetreibers kann mit einem Softwareupdate aufgerüstet werden. Auch sind keine neuen Basisstationen notwendig, da die existierenden eNodeBs nach dem Softwareupdate wie bisher mit LTE Endgeräten kommunizieren können, die eine schnelle Internetverbindung benötigen, sowie mit Geräten aus den neuen Kategorien. In 3GPP werden diese als Verbesserungen für ‚Machine Type Communication‘ bezeichnet, sowie als ‚Cellular Internet of Things‘ (CIoT).

1.17.1 LTE CAT-1 Endgeräte

Etwas erstaunlich ist es vielleicht, dass sogar schon die erste 3GPP LTE Spezifikation (Release 8) eine Endgerätekategorie (Device Category CAT-1) für relativ einfache und stromsparende Geräte mit einem maximalen Durchsatz von 10 Mbit/s beinhaltet. Um die Komplexität zu verringern, können CAT-1 Endgeräte mit nur einer Antenne gebaut werden, also ohne MIMO (Multiple Input Multiple Output) Unterstützung.

1.17.2 LTE CAT-0 Endgeräte und PSM

Einige Jahre später spezifizierte 3GPP dann die Endgerätekategorie CAT-0 in Release 12. Diese Geräte müssen nur eine Datenrate von 1 Mbit/s unterstützen. Außerdem ist es CAT-0 Endgeräten optional erlaubt, nur Halbduplexübertragung zu unterstützen, senden und empfangen ist also nur abwechselnd möglich. Dies spart zusätzlich Kosten

und reduziert den Stromverbrauch, da keine Duplexfilter sondern nur ein elektronischer Sende- und Empfangsumschalter benötigt wird.

Zusätzlich wurde der Power Save Mode (PSM) spezifiziert. Dieser erweitert den RRC Idle Zustand um einen Modus, in dem der Paging Kanal je nach Konfiguration für Stunden, Tage oder Wochen nicht überwacht werden muss. Dies wird nachfolgend noch genauer beschrieben. Für PSM sind keine Änderungen auf der Luftschnittstelle nötig, sämtliche Änderungen betreffen nur das Kernnetz. Zum einen ist es für PSM nötig, Paging Timer pro Gerät aushandeln zu können. Zum anderen muss sich das Netzwerk merken, für welche Endgeräte eingehende Daten zwischengespeichert werden müssen, und diese Endgeräte erst dann zu benachrichtigen, wenn sie den Paging Kanal für kurze Zeit überwachen. Während CAT-1 Endgeräte heute von allen Netzwerken unterstützt werden, benötigen LTE Netzwerke einen Software Upgrade, um auch CAT-0 Endgeräte zu unterstützen, die mit 3GPP Release 12 eingeführt wurden.

1.17.3 LTE CAT-M1 Endgeräte

Die bis zu 20 MHz breiten LTE Kanäle bringen in der Praxis neben einer hohen Endgerätekomplexität auch einen hohen Stromverbrauch mit sich, da ein Endgerät ständig einen 20 MHz breiten Kanal nach Daten und Kontrollinformationen abhören muss. Für IoT Anwendungen mit geringem Datenaufkommen wurde deshalb die Endgerätekategorie CAT-M1 in 3GPP Release 13 spezifiziert. Solche Endgeräte müssen nur einen 1,4 MHz breiten Kanal und eine maximale Datenrate von 1 Mbit/s unterstützen. Eine Änderung der maximalen Kanalbandbreite bedeutet natürlich eine Änderung des Layer 1 der Luftschnittstelle, da die normalen LTE Kontrollkanäle die komplette Kanalbandbreite von bis zu 20 MHz verwenden. Für CAT-M1 Endgeräte wurden deshalb zusätzliche Kontrollkanäle spezifiziert, die nur eine Bandbreite von 1,4 MHz verwenden und für normale LTE Endgeräte nicht sichtbar sind. An dieser Stelle sei angemerkt, dass der gesamte LTE Carrier noch immer bis zu 20 MHz breit sein kann, CAT-M1 Geräte sehen davon jedoch nur 1,4 MHz. Um die Reichweite und Versorgung in Gebäuden zu verbessern, ist es möglich, Signalisierungs- und Nutzdaten zu wiederholen. Dies fügt dem Datenstrom mehr Redundanz hinzu und verbessert somit die Fehlerkorrekturmöglichkeiten.

Wie auch für die Unterstützung von CAT-0 Geräten ist für CAT-M1 Geräte ein Software Update auf der Netzwerkseite erforderlich. Ohne neue Netzsoftware können CAT-M1 Endgeräte keinen Kanal erkennen, da die neuen Signalisierungskanäle nicht ausgestrahlt werden. In vielen Quellen werden auch CAT-M Endgeräte erwähnt, die jedoch mit CAT-M1 Geräten identisch sind, da die Bezeichnung während des Standardisierungsprozesses geändert wurde.

1.17.4 LTE NB1 (NB-IoT) Endgeräte

Für die zuvor genannten neuen Endgerätekategorien wurden hauptsächlich neue Funktionalitäten zur bereits existierenden LTE Luftschnittstelle hinzugefügt. Mit dem NB-IoT Work Item ging 3GPP in Release 13 jedoch weit darüber hinaus, denn das Ziel war, die Leistungsaufnahme in den Endgeräten nochmals signifikant zu verringern. Mehrere Möglichkeiten wurden eingehend analysiert und das Resultat wurde im Technical Report TR 45.820²⁸ auf über 500 Seiten dokumentiert. Im September 2015 wurde ein Kompromiss gefunden²⁹, der schlussendlich als Teil von LTE standardisiert wurde.

NB-IoT soll sehr kostengünstige Endgeräte ermöglichen, deren Modems höchstens 5 US\$ kosten sollen. Zwar bietet NB-IoT nur sehr geringe Datenraten von wenigen Kilobits pro Sekunde, im Ausgleich dafür wurde der Stromverbrauch jedoch nochmals gesenkt und die Versorgung in Gebäuden durch mehr Redundanz nochmals wesentlich verbessert. Somit unterscheidet sich die neue NB-IoT Luftschnittstelle sehr deutlich vom klassischen LTE Air Interface.

Die Kanalbreite eines NB-IoT Kanals beträgt nur 180 kHz, im Vergleich zu einem 20 MHz LTE Kanal oder gar LTE Carrier Aggregation, sehr wenig. Trotzdem verwendet NB-IoT, in gleicher Weise wie der LTE Physical Layer, Orthogonal Frequency Division Multiplexing (OFDM), den gleichen 15 kHz Subcarrier Abstand, die gleiche OFDM Symbolzeit, das gleiche Frame Format und die gleiche Subframe Länge. Ebenso werden die in LTE verwendeten RLC, RRC und MAC Protokolle auch bei NB-IoT verwendet. Im Bereich der Sicherheit werden ebenfalls die LTE Authentifizierungs- und Verschlüsselungsfunktionen verwendet und auch das Konzept der SIM-Karten wird bei NB-IoT verwendet. In kleinen Endgeräten können statt normalen SIM Karten auch Embedded-SIMs (eSIMs) verwendet werden. Diese verhalten sich wie normale SIM Karten, sind aber sehr viel kleiner und werden direkt auf die Platine gelötet.

Handover in RRC-Connected Zustand, ein wichtiges Feature bei LTE, wird bei NB-IoT jedoch bewusst nicht verwendet. Wenn ein NB-IoT Endgerät bemerkt, dass eine andere Zelle besser empfangen werden kann, muss es in den RRC-IDLE State zurückfallen und dann eine Cell Reselection Prozedur durchführen. Außerdem unterstützt NB-IoT bewusst keine Signalstärkemessungen mit Meldung der Ergebnisse zum Netzwerk. Diese Funktionen werden für NB-IoT als kontraproduktiv angesehen, da das System für die Übertragung von nur kleinen Datenmengen pro Zeiteinheit optimiert wurde. Aus diesem Grund ist es nicht nötig, bei sich ändernden Kanalbedingungen einen laufenden Datentransfer in einer anderen Zelle weiterzuführen. Mehr dazu im Laufe dieses Kapitels. Und schließlich wird auch die Rückwärtskompatibilität zu LTE, GSM und UMTS nicht unterstützt, damit ein NB-IoT Modem nur den NB-IoT Teil der Spezifikation unterstützen muss und somit entsprechend einfach, stromsparend und billig aufgebaut sein kann.

1.17.5 NB-IoT Konfigurationsoptionen

Die Kanalbandbreite von 180 kHz wurde gewählt, da dies eine Anzahl unterschiedlicher Konfigurationen in der Praxis ermöglicht. Eine Option ist die Nutzung von einem oder mehreren NB-IoT Kanälen innerhalb eines größeren LTE Kanals. Ebenso ist es möglich, einen NB-IoT Kanal als Teil des Guard Bands eines LTE Trägers zu konfigurieren. Und schließlich kann ein NB-IoT Kanal auch an Stelle eines GSM Kanals übertragen werden, ohne dass dabei GSM Nachbarkanäle gestört werden.

Alle drei Konfigurationsoptionen sind für nicht-NB-IoT fähige Endgeräte transparent. Smartphones, Tablets, etc., können also weder den NB-IoT Kanal im LTE Träger erkennen, noch in dessen Guard Band, und GSM Endgeräte sehen statt eines NB-IoT Kanals nur Rauschen auf einem sonst für GSM genutzten Kanal.

1.17.6 Die NB-IoT Luftschnittstelle

In der Downlink Richtung verwendet NB-IoT Orthogonal Frequency Division Multiplexing (OFDM) mit Quadrature Phase Shift Keying Modulation (QPSK, 2 Bits pro Übertragungsschritt) oder Binary Phase Shift Keying (BPSK, 1 Bit pro Übertragungsschritt). MIMO (Multiple Input Multiple Output) wird auf einem NB-IoT Träger nicht verwendet. Außerdem werden zwölf 15 kHz Subcarrier verwendet, die als ‚Tones‘ (Töne) bezeichnet werden, also die gleiche Anzahl an Subcarriern wie in einem LTE Resource Block (RB). Die Kanalzuweisung (Scheduling) von Endgeräten in der Downlink Richtung unterscheidet sich jedoch von LTE mit seiner sehr großen Kanalbandbreite. Bei LTE können viele Endgeräte gleichzeitig Daten empfangen, da hier 50 RBs in einem 10 MHz Kanal zur Verfügung stehen und 100 RBs in einem 20 MHz Kanal. In einem 180 kHz NB-IoT Kanal gibt es jedoch nur einen RB und es wurde entschieden, dass nur ein Endgerät Daten auf allen 12 Subcarriern pro 1 ms Subframe empfangen darf. Abb. 1.34 zeigt diese Anordnung.

In der Uplink Richtung wird für die Datenübertragung ebenfalls wie bei LTE ein 15 kHz Subcarrier Raster und Single Carrier Frequency Division Multiple Access (SC-FDMA) mit BPSK und QPSK Modulation verwendet. Optional können auch 3,75 kHz Subcarrier (Tones) verwendet werden. Diese Option wurde für Fälle spezifiziert, in denen das Endgerät zwar Daten vom Netzwerk empfangen kann, aber nicht in der Lage ist, mit einer kleinen Antenne und geringer Sendeleistung von seinem Standort aus das Netzwerk zu erreichen. Wird statt einem 15 kHz Kanal ein 3,75 kHz Kanal verwendet, kann die Sendeleistung auf dem nochmals engeren Kanal gebündelt werden. Dies erhöht das ‚Link-Budget‘ und somit die Chance, auf der Netzwerksseite empfangen zu werden. Solche sehr schlechten Empfangsbedingungen werden als ‚Extreme Coverage‘ bezeichnet. NB-IoT ist dafür ausgelegt, auch noch unter Bedingungen Daten übertragen zu können, die 20 dB unterhalb der GSM Spezifikation liegen. Endgeräte, die statt für ‚Extreme Coverage‘ für möglichst geringe Leistungsaufnahme optimiert sind, bietet

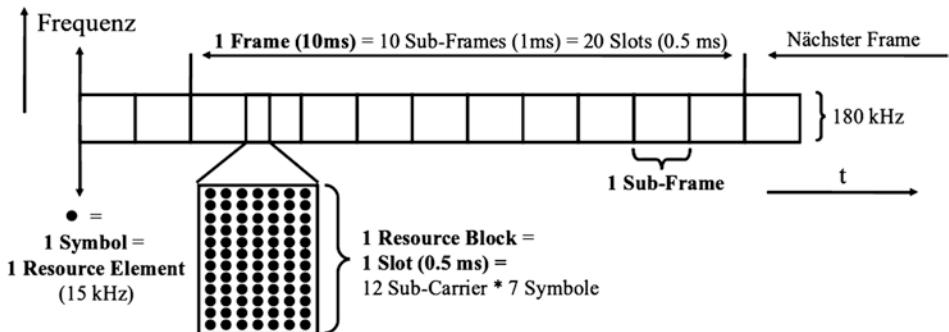


Abb. 1.34 Das NB-IoT Resource Grid

Power Class 5 eine Sendeleistung von nur 20 dBm (0,1 W). Im Unterschied zum Downlink Kanal, auf dem nur ein Gerät alle 12 Subcarrier zugewiesen bekommt, kann in Uplink Richtung einem Gerät 1, 3, 6 oder alle 12 Töne zugewiesen werden. Wird mehr als ein Ton verwendet, wird dies als Multi-Tone Übertragung bezeichnet.

1.17.7 NB-IoT Control Channels und Scheduling

Aufgrund der zahlreichen Modifikationen sind die traditionellen Signalisierungskanäle nicht für NB-IoT nutzbar. Zwar werden die grundsätzlichen Ideen wie Random Access und Zuweisung von Transmission Opportunities beibehalten, das Format und Lage der Kanäle ist jedoch neu. Wie in LTE werden 7 Töne auf der Zeitachse in einem 0,5 ms Slot zusammengefasst und zwei Slots bilden einen 1 ms Subframe, der die kleinste Übertragungseinheit darstellt. 10 Subframes werden in einem 10 ms Radio Frame gebündelt. Wie in LTE gibt es Narrowband Varianten der Referenzsignale, sowie Primary und Secondary Synchronisationssignale, damit Endgeräte den Kanal finden und sich synchronisieren können. User Daten und Systeminformationen werden über den Narrowband Physical Downlink Shared Channel (NPDSCH) übertragen. Im Uplink gibt es den gleichen Kanal für Uplink Daten und Acknowledgements für Downlink Daten.

Der Narrowband Physical Downlink Control Channel (NPDCCH) wird für folgende Zwecke verwendet:

- **Downlink Zuweisungen (Assignments):** Wenn Daten für ein NB-IoT Gerät beim eNodeB eintreffen, wird der Downlink Control Channel verwendet, um Downlink Übertragungen für das Gerät anzukündigen. Jedes Assignment enthält die Position und die Anzahl der Subframes, die auf dem Downlink Shared Channel dem Endgerät zugeteilt werden und wie oft die Daten wiederholt werden, um das Link Budget zu verbessern, und ob eine Bestätigung (Acknowledgement) vom Endgerät

erwartet wird, nachdem es die Daten erhalten hat. Im Unterschied zu LTE Downlink Assignments, die sich nur auf den aktuellen Subframe beziehen, kann bei NB-IoT Downlink Assignments eine Verzögerung zwischen 5 und 10 Subframes zwischen Assignment und Nutzung des Downlink Kanals angegeben werden. Weiterhin kann sich ein Assignment wegen der sehr schmalen Kanalbandbreite auf mehr als nur einen Subframe beziehen. Ein weiterer Grund für dieses Inter-Subframe Scheduling ist, dass nur zwei Downlink Assignments pro Subframe übertragen werden können. Da nur eine Datenübertragung in einem nachfolgenden Subframe enthalten sein kann, müssen die Zuweisungen unterschiedliche Verzögerungswerte (Delays) beinhalten.

- **Uplink Grants:** Das Endgerät kann den eNodeB über Buffer Status Informationen und Random Access Requests mitteilen, dass Daten für die Übertragung in Uplink Richtung bereitstehen. Der eNodeB teilt dem Endgerät dann Ressourcen in Uplink Richtung zu. Der Empfang von Uplink Daten wird immer vom eNodeB im Downlink bestätigt.
- **Paging:** Wenn Downlink Daten für ein Endgerät eingehen, das sich im RRC-IDle Zustand befindet, muss es auf dem Paging Kanal benachrichtigt werden. Ein typisches Paging Intervall dauert zwischen einer und zwei Sekunden.

Wegen der engen Kanalbandbreite wird der NPDCCH nicht in jedem Subframe in den ersten Symbolen übertragen, sondern nur einmal über viele Subframes und verwendet dann den kompletten Subframe. Die Periode dieses und anderer Kanäle (z. B. die Random Access Kanal Periode im Uplink) kann 40 ms bis 2,56 s betragen und wird in den System Information Messages bekannt gegeben, damit alle Endgeräte wissen, wann sie im Downlink nach Downlink Zuweisungen, Uplink Grants und Paging Nachrichten suchen müssen. Wie in LTE werden System Information Messages, die in NB-IoT als SIB-NB bezeichnet werden, über den Downlink Shared Channel übertragen. Der Master Information Block (MIB) wird jedoch separat übertragen. Der MIB enthält nur 34 Bits und wird über eine Periode von 640 ms mit vielen Wiederholungen gesendet, um die Redundanz zu erhöhen. Er enthält unter anderem die 4 höchstwertigen Bits der System Frame Nummer, 4 Bits für Größe und Lage des SIB1-NB, 5 Bits, die einen System Information Value Tag beschreiben und 1 Bit um anzugeben, ob Access Class Barring angewendet wird, um den Zugang zu dieser Zelle auf eine Untermenge aller Endgeräte im Überlastfall zu limitieren. Alle anderen Informationen sind in den System Information Messages enthalten.

1.17.8 NB-IoT Multi-Carrier Operation

Um die Systemkapazität zu erhöhen, können mehrere NB-IoT Kanäle (Carrier) pro Sektor eines eNodeBs konfiguriert werden. Einer der Kanäle wird dann als Anchor (Anker) Carrier verwendet und strahlt alle Systeminformationen, sowie zusätzlich den Control Channel und den Shared Channel, aus. Alle Endgeräte, die sich im

RRC-IDLE Zustand befinden, sind auf diesen Carrier synchronisiert. Endgeräte im RRC-Connected Zustand können dann ihre Daten entweder über den Anchor Carrier senden und empfangen oder können angewiesen werden, einen anderen Carrier zu verwenden. Uplink und Downlinkdaten eines Endgeräts können aufgrund der Halbduplex Übertragung in NB-IoT getrennt auf unterschiedlichen Carrieren gesendet werden. Wenn ein Endgerät wieder in den RRC-IDLE Zustand zurückfällt, kehrt es zum Anchor Carrier zurück. Auf diese Weise kann die Kapazität in einem Sektor vergrößert werden, ohne die Komplexität der Endgerätehardware zu erhöhen oder eine höhere Stromaufnahme zu verursachen. Zu jeder Zeit sendet oder empfängt ein Endgerät nur auf einem 180 kHz Träger.

1.17.9 NB-IoT Durchsatz und Anzahl der Geräte pro Zelle

Basierend auf den physikalischen Eigenschaften eines NB-IoT Trägers kann folgende Datenrate pro Träger errechnet werden: Auf der Frequenzachse gibt es 12 Subcarrier und 7 Symbole pro 0,5 ms Slot. Multipliziert mit 2000 (1/s) und 2 Bits pro Übertragungsschritt (QPSK Modulation) beträgt die Rohdatenrate 336 kbit/s. Nicht alle Symbole können jedoch für die Datenübertragung verwendet werden. Wenn ein NB-IoT Kanal in einem LTE Träger eingebettet ist, können die ersten 1–3 Symbole pro Subframe nicht verwendet werden, da hier LTE Kontrollinformationen gesendet werden. Dies reduziert die Kanalkapazität in Abhängigkeit der Anzahl der Symbole für LTE Kontrollinformationen um 10–20 %. Zusätzlich wird auch Kapazität für die Narrowband Reference Signale benötigt, für LTE Reference Signale, falls der NB-IoT Kanal ‚In-Band‘ verwendet wird, für den MIB, für den Narrowband Control Channel und für die System Information Nachrichten, die periodisch ausgestrahlt werden. In der Praxis beträgt deswegen die nutzbare Datenrate eines NB-IoT Kanals etwa 200 kbit/s, falls Nutzdaten nur einmal und nicht mehrmals aus Redundanzgründen übertragen werden. Trotz dieses sehr geringen Durchsatzes wurde das System so spezifiziert, dass jeder eNodeB Sektor bis zu 50.000 Endgeräte verwalten kann. Dies macht deutlich, dass NB-IoT speziell für Endgeräte ausgelegt wurde, die nur sehr selten und sehr wenig Daten übertragen. 3GPP TR 45.820 Abschn. 5.3.6 enthält eine Netzwerksimulation die bestätigt, dass eine solch hohe Gerätanzahl mit einem 180 kHz NB-IoT Kanal unterstützt werden kann. In der Simulation senden die Endgeräte 105 Bytes Daten pro Übertragung, die 20 Nutzdatenbytes enthalten, sowie den kompletten IP- und Radio Layer Overhead. Zusätzlich flossen in die Simulation viele Systemparameter ein, wie die räumliche Verteilung von eNodeBs, Netzwerkzugangssignalisierung, Uplink Grants, Downlink Assignment und unterschiedliche Signalstärken in Abhängigkeit, wie weit Endgeräte vom Zellmittelpunkt entfernt sind.

Mit einer einfachen Näherungsrechnung kann dieser Wert grob überprüft werden: 50.000 Endgeräte die 105 Bytes fünfmal pro Stunde senden, produzieren 26.250.000 Bytes pro Stunde. Geteilt durch 60 min und 60 s sowie multipliziert mit 8 Bits per Byte

resultiert daraus eine Datenrate von 58.333 Bits pro Sekunde. Das bedeutet, dass das Datenvolumen gut in den NB-IoT Kanal passt, auch wenn der komplette Overhead von der Datenrate subtrahiert wird.

Eine andere interessante Zahl, die sich näherungsweise einfach errechnen lässt, ist die Anzahl der Random Access Requests, die 50.000 Endgeräte produzieren, die 5 mal pro Stunde Daten übertragen. Multipliziert man die Anzahl der Endgeräte mit 5 Anfragen pro Stunde und dividiert das Ergebnis dann durch 60 min und 60 s, erhält man 70 RACH (Random Access Channel) Anfragen pro Sekunde, also einen Request alle 14 ms.

1.17.10 NB-IoT Stromsparmechanismen

Eine weitere Anforderung an die NB-IoT Luftschnittstelle ist eine Batterilaufzeit von bis zu 10 Jahren bei sehr seltenen Datentransfers. 3GPP kam in TR 45.820 Abschn. [5.3.6.4](#) zum Schluss, dass dies unter folgenden Bedingungen möglich ist:

Für die Berechnung wurde eine Batteriekapazität von 5 Wattstunden zugrunde gelegt, was etwa einem Drittel der Batteriekapazität einer Smartphonebatterie entspricht. Weiterhin wurden im TR folgende Annahmen über den Energieverbrauch des Endgeräts in unterschiedlichen Aktivitätszuständen gemacht: Im ‚Idle‘ Zustand, in dem sich das Endgerät die meiste Zeit befindet, wird ein Verbrauch von 0,015 mW angenommen. Wenn das Endgerät immer in diesem Zustand wäre, würde ein 5 Wh Batterie Energie für 38 Jahre liefern können. Die Selbstendladungsrate der Batterie ist hierbei jedoch noch nicht berücksichtigt. Das Varta Handbuch für ‚Primary Lithium Cells‘³⁰ gibt eine Selbstendladungsrate für nicht wieder aufladbare Lithium Batterien von weniger als 1 % pro Jahr an. Dies ist signifikant weniger als die Selbstendladungsrate von wieder aufladbaren Akkus.

Natürlich befindet sich ein Endgerät nicht immer im ‚Idle‘ Zustand. Während der Datenübertragung wird von einem Stromverbrauch von 500 mW ausgegangen. Würde sich ein Endgerät immer in diesem Zustand befinden, wäre die Batterie schon nach 10 h verbraucht. Da sich diese Werte signifikant unterscheiden, beschäftigte sich die Studie dann mit unterschiedlichen Übertragungsverhalten von Endgeräten. Wenn 200 Bytes alle 2 h übertragen werden, würde ein Endgerät mit einer 5 Wh Batterie für 1,7 Jahre betrieben werden können. Wenn das Endgerät nur einmal pro Tag 50 Bytes versendet, würde die Batteriekapazität für 18,1 Jahre ausreichen.

1.17.11 NB-IoT – High Latency Kommunikation

Eine wichtige Anforderung in vielen IoT Szenarien ist die Nutzbarkeit von Geräten, die sich tief innerhalb von Gebäuden befinden (Deep In-house Coverage), sowie an Orten mit sehr schlechter Netzabdeckung. NB-IoT Netzwerke bieten für solche Szenarien die Möglichkeit, den Datentransfer sehr oft zu wiederholen, um dem Empfänger die

Möglichkeit zu geben, die Signalenergie, die bei jedem Sendeversuch verwendet wird, zu kombinieren. Nachteil dieser Methode ist jedoch, dass auch der Versand eines kleinen IP Paketes eine lange Zeit benötigt. Laut einer interessanten Berechnung in einem Ericsson White Paper zufolge kann es bei sehr schlechtem Empfang bis zu 7 s dauern, bis ein kleines UDP Paket empfangen wird, da das System jeden Übertragungsschritt wie die Anforderung eines Übertragungskanals, die Bandbreitenzuteilung, den Datentransfer und die Bestätigung (Acknowledgement) viele dutzend Male wiederholen muss³¹.

Eine andere Art der Verzögerung tritt durch den Stromsparmechanismus auf. Wenn es für ein Gerät ausreicht, nur jede halbe Stunde zu überprüfen, ob eingehende IP Pakete auf die Übertragung warten, kann es sein Radiomodul in dieser Zeit komplett deaktivieren. Der Nachteil hierbei ist jedoch, dass es im schlechtesten Fall bis zu 30 min dauert, bis das Endgerät auf ein ankommendes IP Paket reagieren kann. Für solche Szenarien wurde in den 3GPP Spezifikationen eine Reihe von Erweiterungen unter dem Begriff „High Latency Communication“ spezifiziert, die nun nachfolgend beschrieben werden.

Extended Idle Mode Discontinuous Reception (eDRX)

Solange ein Endgerät im Idle Zustand ist, muss es den LTE Paging Kanal überwachen, damit es auf eingehende Paging Nachrichten reagieren kann. Diese werden vom Netzwerk gesendet, wenn keine aktive Radioverbindung besteht und IP Pakete vom Internet für das Endgerät eingehen. Das Endgerät antwortet dann auf das Paging und es wird eine Air Interface Verbindung aufgebaut, über die dann das IP Paket weitergeleitet wird. Ein typisches Paging Intervall in LTE Netzwerken ist heute 1,28 s, der Radiochip muss sich also einmal alle 1,28 s aktivieren und den Paging Kanal abhören. Während die dazu zusätzliche benötigte Energie für Geräte wie Smartphones im Verhältnis zum restlichen Energieverbrauch nicht relevant ist, kann ein solches Intervall jedoch zu einem signifikanten und nicht akzeptablen Energieverbrauch für ein IoT Gerät führen.

Falls ein längeres Paging Intervall für ein IoT Gerät notwendig ist, kann es dies dem Netzwerk während der Attach und der Tracking Area Update Prozedur mitteilen und ein längeres Paging Intervall zwischen 5,12 s und 2621,44 (43,69 min) anfordern. Das Netzwerk kann das vorgeschlagene Intervall akzeptieren, modifizieren oder ablehnen. Nachdem die Attach oder Tracking Area Update Prozedur beendet ist und das Netzwerk den Radio Access Bearer abgebaut hat, kann das Endgerät das Radiomodem während der „Extended DRX Time“ deaktivieren, ohne den Bearer Kontext, also die IP Adresse, zu verlieren.

Extended Buffering von Daten

Wenn Daten für ein Endgerät vom Internet eingehen, bittet der S-GW die MME, das Endgerät zu page und einen Radiokanal aufzubauen. Wenn die MME erkennt, dass ein Endgerät im Extended Idle Mode DRX Zustand ist, teilt sie dies dem S-GW mit. Dieser puffert dann die IP Pakete, bis das Gerät wieder angesprochen werden kann. Die MME wartet dann für den Rest des DRX Zyklus und schickt dann eine Paging Nachricht an

das Endgerät. Dies kann bis zu 43 min dauern. Erst danach, oder falls das Endgerät zu jedem Zeitpunkt davor neue Daten in Uplink Richtung übertragen will, wird ein Radio-kanal aufgebaut und die wartenden Pakete übertragen.

Power Save Mode

Eine weitere Option, das Radiomodul für eine lange Zeit abzuschalten, ist die Power Save Mode (PSM) Funktionalität. Um diese zu aktivieren, verhandelt das Endgerät mit dem Netzwerk eine ‚Active Time‘, in der es den Paging Kanal abhört, nachdem es in den RRC-Idle Zustand wechselt. Nachdem die ausgehandelte Zeit abgelaufen ist, ist das Gerät dann nicht mehr erreichbar und kann sein Radiomodul deaktivieren, bis es wieder eine Tracking Area Update Prozedur durchführen muss oder es neue Daten im Uplink zu übertragen hat. Zusätzlich kann das Endgerät einen längeren Tracking Area Update Timer mit dem Netzwerk aushandeln. Im Normalfall ist dieser Timer auf einige Stunden konfiguriert. Falls das Netzwerk zustimmt, kann das Endgerät einen Tracking Area Update Timer (T3412) mit einer Periode von mehreren Tagen oder sogar Wochen bekommen. Dies macht besonders für Geräte und Anwendungen Sinn, die nur Daten zu einem Server schicken wollen und nur eine Antwort während oder kurz nach einem solchen Ereignis erwarten.

1.17.12 NB-IoT – Optimierung von IP und nicht-IP basierend Datenübertragungen

Zusätzlich zu den bisher beschriebenen Stromsparmechanismen für Zeiten ohne Datenübertragung spezifizierte 3GPP auch Stromsparmechanismen für Fälle, in denen das Endgerät nur wenige Daten über die Luftschnittstelle übertragen will. Neben des Energiebedarfs für die Datenübertragung reduzieren die nachfolgend beschriebenen Funktionen auch den Overhead für die Datenübertragung. Wenn ein einzelner eNodeB mit mehreren hundert oder mehreren tausend NB-IoT Endgeräten kommuniziert, ist es sehr wichtig, den dafür nötigen Overhead für die Signalisierung so klein wie möglich zu halten. 3GPP TS 23.401 Abschn. 1.1 enthält dazu drei Funktionen:

Wiederverwenden von RRC Verbindungen

Die erste Funktion wurde als Teil der „User Plane CIoT (Cellular IoT) EPS Optimizations“ standardisiert und umfasst die Suspendierung und Wiederaufnahme einer RRC Verbindung. In LTE Netzwerken wird eine RRC Verbindung üblicherweise nach 10–20 s Inaktivität deaktiviert und ein neuer RRC Kontext wird aufgebaut, wenn neue IP Datenpakete zur Übertragung anstehen. Dies ist in LTE kein Problem, da dieser Prozess nur etwa 100 ms dauert und die Anzahl der Nutzdaten, die dann übertragen werden, den Overhead um Größenordnungen übersteigt. Wird der ganze Prozess jedoch durchlaufen, um dann nur ein paar Bytes in einem oder wenigen IP Paketen zu übertragen, ist dies sehr ineffizient. Als Konsequenz daraus wurde eine Möglichkeit

spezifiziert, den RRC Kontext in Übertragungspausen zu erhalten, ihn also auf Endgeräteseite und im Netzwerk zu suspendieren. Auf diese Weise ist keine erneute Authentifizierung und Aktivierung der Verschlüsselung notwendig und es müssen keine RRC Connection Reconfiguration Nachrichten für einen erneuten Beareraufbau gesendet werden, wenn neue Signalisierungs- und Nutzdaten zur Übertragung anstehen.

Nutzdaten über Signalisierungskanäle

Eine sehr viel weitergehende Möglichkeit, den Overhead zu reduzieren, ist das Aufheben der Trennung zwischen User Plane und Control Plane. In LTE wird die Control Plane für Management Tasks wie den Aufbau einer Verbindung auf der Luftschnittstelle, für Radio Link Control, Authentifizierung, Aktivieren der Verschlüsselung, sowie für das Mobility- und Session Management verwendet. Aus Sicht des Radionetwerks sind der eNodeB und die Mobility Management Engine (MME) die Endpunkte für Signalisierungsnachrichten, die über einen logischen Signaling Radio Bearer (SRB) über die Luftschnittstelle erreicht werden. Userdaten, also IP Pakete, werden über einen logischen Data Radio Bearer (DRB) transparent über den eNodeB zu und vom Serving Gateway (S-GW) und von dort über das Packet Gateway (P-GW) von und zum Internet übertragen. Aus logischer Sicht ist diese Trennung sehr wichtig, erzeugt aber zusätzlichen Overhead, besonders auf der Luftschnittstelle, da hier Signalisierungsnachrichten notwendig sind, um zusätzlich zu einem SRB einen DRB aufzubauen. Um diesen Overhead zu reduzieren, wurde eine Funktion namens „Control Plane CIoT EPS Optimization“ standardisiert, mit der IP Pakete in einem transparenten Container in EPS Session Management Nachrichten über die MME übertragen werden können. Die MME packt die Nutzdaten aus und schickt sie dann an den S-GW weiter, der die Datenpakete dann an den P-GW weiterleitet. Von hier werden die IP Datenpakete zum Internet weitergeleitet. In der Gegenrichtung findet der umgekehrte Prozess statt. Das Netzwerk kann sich entscheiden, vom Internet eingehende IP Pakete am S-GW über einen User Data Bearer zu übertragen oder aber die Daten an die MME weiterzusenden, damit diese dann die Daten in einen transparenten Container in Signaling Messages verpackt und dann weiterleitet.

Non-IP Data Delivery (NIDD)

Um den Datentransfer noch weiter zu optimieren, sehen die Standards auch eine Funktion vor, die als Non-IP Data Delivery (NIDD) bezeichnet wird. Details sind in 3GPP TS 23.682³², Abschn. 1.5.14 zu finden. Mit dieser Funktion kann das Endgerät zu versendende Daten in einem transparenten Container verpacken, es wird jedoch kein IP Stack dazu verwendet. Die MME auf Netzwerkseite leitet solche Datenpakete an die Service Capability Exposure Function (SCEF), wie in Abb. 1.35 gezeigt, weiter. Für Server im Internet stellt die SCEF diese Daten dann über ein IP basiertes API (Application Programming Interface) zur Verfügung. Um Daten an ein solches Endgerät aus dem Internet zu senden, ist ebenfalls die SCEF der Ansprechpartner. Das bedeutet,

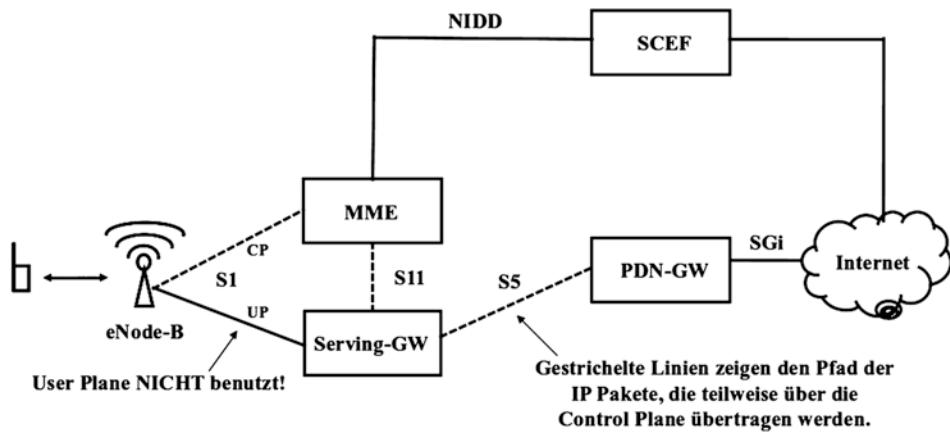


Abb. 1.35 CIoT Control Plane Optimization und Non-IP Data Delivery

dass keine direkte Kommunikation zwischen Anwendungsserver im Internet und einem IoT Gerät besteht, es ist immer der Server des Netzbetreibers dazwischengeschaltet.

Alle beschriebenen Methoden sind voneinander unabhängig und ergänzen sich gegenseitig. Ein Netzbetreiber kann sich deshalb entscheiden, ob er IP-basierte oder non-IP basierte NB-IoT Datenübertragungen, oder beide Varianten unterstützen will.

1.17.13 NB-IoT Zusammenfassung

Verglichen mit allen 3GPP Machine Type Communication (MTC) und Internet of Things (IoT) Erweiterungen der vergangen Jahre ist NB-IoT die bei weitem umfangreichste. Optimierte für stromsparenden Betrieb, niedrige Kosten und niedrige Datenraten benötigt NB-IoT ein neues Modem- und Baseband Design. Wenn die geplanten Stromsparmechanismen und Deep-Inhouse Szenarien in der Praxis umgesetzt werden können, wird die NB-IoT Luftschnittstelle viele neue Möglichkeiten bieten, Endgeräte mit dem Internet zu verbinden, ohne dabei einen lokalen Hub zu benötigen.

1.18 Fragen und Aufgaben

1. Wie viele Subcarrier werden in einem 10 MHz FDD LTE-Kanal verwendet?
2. Was ist der Unterschied zwischen einem S1- und einem X2-Handover?
3. Welche Aufgaben hat die MME und welche das Serving-Gateway?
4. Was ist ein Resource Block (RB)?
5. Wie kann ein Endgerät Zugriff auf den Physical Uplink Shared Channel erhalten?
6. Was ist der Unterschied zwischen ARQ und HARQ?

7. Was ist der Unterschied zwischen einem Default- und einem Dedicated-Bearer?
8. Für was wird DRX im RRC Connected-Zustand verwendet?
9. Wie wird das Mobility Management im RRC Idle-Zustand gehandhabt?
10. Was ist der Unterschied zwischen einer Cell Change Order und einem Handover?
11. Wie kann das LTE-Netzwerk mit dem GSM/UMTS-Netzwerk verbunden werden und warum sollte man dies tun?
12. Was bedeutet der Begriff CS-Fallback?
13. Welche Nachteile haben internet-basierte Sprachdienste verglichen mit dem Sprachdienst der Netzbetreiber?
14. Welche Technologien werden heute verwendet, um Mobilfunkstandorte mit dem Kernnetz zu verbinden?

Lösungen sind auf der Website zum Buch unter <http://www.cm-networks.de> zu finden.

Anmerkungen

1. 3GPP, Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) Radio Access Capabilities Release 8, TS 36.306.
2. 3GPP, Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) Radio Transmission and Reception, version 9.2.0, TS 36.101.
3. 3GPP, Evolved Universal Terrestrial Radio Access Network (E-UTRAN); S1 Application Protocol (S1AP), TS 36.413.
4. The Internet Engineering Task Force (IETF), Stream Control Transmission Protocol, RFC 4960, <http://tools.ietf.org/html/rfc4960>.
5. M. Sauter, How File Sharing of Others Drains Your Battery, https://blog.wirelessmoves.com/2007/05/how_file_sharin.html, May 2007.
6. Calhoun et al., Diameter Base Protocol, IETF RFC 3588.
7. 3GPP, Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation, TS 36.211.
8. 3GPP, Radio Resource Control (RRC); Protocol Specification, TS 36.331.
9. 3GPP, Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures, TS 36.213.
10. 3GPP, Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Link Control (RLC) Protocol Specification, TS 36.322.
11. L.-E. Johnsson et al., The ROBust Header Compression (ROHC) Framework, IETF RFC 4995 and 5795.
12. G. Pelletier, ROBust Header Compression Version 2 (ROHCv2): Profiles for RTP, UDP, IP, ESP and UDP-Lite, IETF RFC 5225.
13. 3GPP, Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) Protocol Specification, TS 36.321.
14. 3GPP, General Packet Radio Service (GPRS) Enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) Access, TS 23.401.

15. 3GPP, Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 Application Protocol (X2AP), TS 36.423.
16. 3GPP, Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) Procedures in idle Mode, TS 36.304.
17. 3GPP, General Packet Radio Service (GPRS); Service Description; Stage 2, TS 23.060.
18. 3GPP, Self-configuring and Self-optimizing Network (SON) Use Cases and Solutions (Release 9), TS 36.902.
19. Huawei, LTE Radio Network Planning Introduction. <https://www.academia.edu/35361843/LTE-Radio-Network-Planning-Introduction.pdf>, ohne Datum
20. 3GPP, Circuit Switched (CS) Fallback in Evolved Packet System (EPS); Stage 2, TS 23.272.
21. 3GPP, Network sharing; Architecture and functional description, TS 23.251.
22. 3GPP, Radio Resource Control (RRC); Protocol specification3GPP TS 25.331, Chapter 10.2.16c.
23. Bao et al., IPv6 Addressing of IPv4/IPv6 Translators, IETF RFC 6052.
24. 464XLAT service has been standardized in RFC 6877 [33]
25. M. Sauter, IPv6-only in a French Mobile Network, <https://blog.wirelessmoves.com/2020/09/ipv6-only-in-a-french-mobile-network.html>, September 2020
26. Wikipedia, I/O Virtualization, https://en.wikipedia.org/wiki/I/O_virtualization, September 2021
27. Wikipedia, Data Plane Development Kit, https://en.wikipedia.org/wiki/Data_Plane_Development_Kit, September 2021
28. 3GPP, Cellular System Support for Ultra-Low Complexity and Low Throughput Internet of Things (CIoT), 3GPP Technical Report TR 45.820.
29. 3GPP, Narrowband IoT Work Item Description, RP-151621.
30. Varta, Primary Lithium Cells – Sales Program and Technical Handbook, http://www.varta-microbattery.com/applications/mb_data/documents/sales_literature_varta/HANDBOOK_Primary_Lithium_Cells_en.pdf, Accessed September 2016.
31. S. Landström et al., NB-IoT: A Sustainable Technology for Connecting Billions of Devices, Ericsson Technology Review Volume 93, 3/2016.
32. 3GPP, Architecture Enhancements to Facilitate Communications with Packet Data Networks and Applications, TS 23.682.



5G New Radio (NR) und das 5G Kernnetz

2

2.1 Einführung und Überblick

2015 begannen in der 3GPP Standardisierungsorganisation die ersten Diskussionen über einen Nachfolger des 4G LTE Systems. Diese Phase dauerte etwa 2 Jahre und endete mit der Erstellung einer Anzahl von Technical Report (TR) Dokumenten. Diese enthielten eine detaillierte Analyse von unterschiedlichen Optionen für alle Teile der neuen Netzarchitektur und eine Einigung, welche dieser Optionen in den darauf folgenden zwei Jahren standardisiert werden sollten. Am Ende des Jahres 2018 wurde dann in 3GPP Release 15 die erste Version des neuen 5G Standards veröffentlicht. Ein zentrales Dokument ist dabei 3GPP TS 38.401¹, das einen Überblick über das neue 5G Radio Access Network (RAN) gibt, das auch als 5G New Radio (NR) bezeichnet wird. Viele weitere TS Dokumente in der 38er Serie enthalten dann alle Details über das 5G Zugangsnetzwerk.

Grundsätzlich hat die 5G NR Luftschnittstelle sehr viele Gemeinsamkeiten mit LTE. LTE wurde jedoch nur für einen spezifischen Anwendungsfall entwickelt: Bereitstellung eines schnellen und mobilen Internetzugangs. Wie im Kapitel über LTE beschrieben, wurde die LTE Luftschnittstelle dann später ergänzt, um auch weitere Anwendungsfälle mit anderen Anforderungen an Bandbreite, maximaler Verzögerung und Leistungsaufnahme abdecken zu können. Dazu zählt z. B. die Narrow-Band Internet of Things (NB-IoT) Erweiterung. Viele Kompromisse waren jedoch notwendig, um die Rückwärtskompatibilität sicherzustellen. Das 5G Radionetzwerk wurde deswegen sehr viel flexibler standardisiert. Statt die gleichen Regeln über die komplette Bandbreite eines Kanals und in jedem Radio Frame anzuwenden, ist es nun möglich, den Kanal in verschiedene Bereiche einzuteilen, in denen jeweils eigene Regeln gelten können. Bei Erscheinen dieser Auflage wird diese neue Flexibilität der 5G Luftschnittstelle jedoch noch nicht genutzt, da die erste Anwendung des Systems eine Erhöhung der

Geschwindigkeit und die Erweiterung der Kapazität der bestehenden LTE Netzwerke war.

Hauptsächlich aufgrund des Marktdrucks wurde entschieden, das 5G System so schnell wie möglich einzuführen. Dies wurde durch die Weiterverwendung des existierenden LTE Radio- und Kernnetzes als sogenannter „Kommunikations-Anker“ für 5G Radiozellen erreicht. Solche 5G Netzwerke werden deshalb auch als 5G Non-Standalone Architecture (NSA) Netzwerke bezeichnet und dieses Kapitel wird im Folgenden nochmals eine kurze Einführung der 4G LTE Prozeduren geben, die für den Aufbau eines 5G Kanals notwendig sind.

Die Nutzung der 4G/5G Non-Standalone Architektur für den schnellen mobilen Internetzugang in Kombination mit einem 4G Kernnetzwerk ist jedoch nur der erste Schritt. Ziel ist es, über einen längeren Zeitraum das heutige 4G Netzwerk durch 5G Komponenten zu ersetzen, um dann auch andere Nutzungsszenarien wie Narrow-Band Machine Type Communication (MTC), eine niedrigere Latenz und unterschiedliche Quality-of-Service Mechanismen bereitzustellen. Hierfür wird ein 5G Kernnetz benötigt, das auch in 3GPP Release 15 spezifiziert wurde. Ein Überblick über das 5G Kernnetz wurde in 3GPP TS 23.501² spezifiziert. In der Praxis werden 5G Kernnetzwerke von manchen Netzbetreibern heute schon in der Praxis verwendet und es ist zu erwarten, dass 4G und 5G Kernnetze über viele Jahre gleichzeitig verwendet werden. Für neue Arten von Netzwerken, z. B. für sogenannte „Campus Netzwerke“, die Geräte in Fabriken vernetzen, ist keine Rückwärtskompatibilität notwendig. Somit können dort Netzwerke eingesetzt werden, die nur ein 5G Kernnetzwerk verwenden. Hier ist es dann auch einfacher, neue Funktionalitäten wie geringere Latenz und sehr hohe Übertragungssicherheit zu verwenden.

2.1.1 Gründe für den 5G Start als Hybrid Lösung

Neben der Notwendigkeit, 5G Netzwerke schnellstmöglich einzuführen, gab es eine Anzahl weiterer Gründe, weshalb Netzbetreiber zunächst auf eine 4G/5G Hybridlösung setzten. In der Praxis wurde von den meisten Netzbetreibern schon sogenanntes Low-, Mid- und High-Band Spektrum für LTE eingesetzt. Frequenzbereiche zwischen 700 und 900 MHz werden dabei als Low-Band Spektrum bezeichnet, die Frequenzbänder um 1800 und 2100 MHz werden als Mid-Band Spektrum bezeichnet. High-Band Spektrum bezeichnet den Bereich um 2600 MHz, ist jedoch aufgrund von schlechteren Ausbreitungseigenschaften nicht für die Netzbdeckung außerhalb von Städten geeignet. Somit stand in den meisten Teilen der Welt in diesen Frequenzbereichen kein freies Frequenzband mehr zur Verfügung, das eine höhere Übertragungsgeschwindigkeit für 5G gegenüber 4G geboten hätte. Stattdessen wurde in Europa und Asien ein Frequenzband im Bereich von 3500 MHz von den Regulierungsbehörden zur Verfügung gestellt, um dort 5G mit einer Bandbreite von bis zu 100 MHz zu betreiben. Verglichen mit der maximalen LTE Kanalbandbreite von 20 MHz ist dies eine deutliche Steigerung. In

anderen Teilen der Welt, z. B. Nordamerika, wurden Anfangs keine weiteren Frequenzbereiche unterhalb von 6 GHz für 5G zur Verfügung gestellt. Dort waren jedoch Frequenzen oberhalb von 24 GHz erhältlich und Netzbetreiber entschlossen sich dort, 5G im sogenannten Millimeterwellenbereich (mmWave) einzuführen. Die Datenübertragung in diesem Ultra-high-band Spektrum ist jedoch wesentlich aufwendiger als in den bisher verwendeten Frequenzbereichen unterhalb von 6 GHz.

Es wäre natürlich möglich gewesen, Spektrum im Low-Band für 5G umzuwidmen. Dies hätte jedoch große Nachteile mit sich gebracht. Da nur sehr wenig Spektrum im Low-Band für Mobilfunknetzwerke zur Verfügung steht, hätte eine Umwidmung eine Reduktion der sowieso schon niedrigen Kapazität der LTE Netzwerke in diesem Bereich geführt. Um dies zu kompensieren, wäre es möglich gewesen, zusätzlich zum 5G Kanal noch LTE Kanäle für eine Verbindung zu bündeln. Das resultierende Hybrid-Setup wäre jedoch ein ähnliches Szenario wie die zuvor beschriebene LTE-Anker Lösung und würde von der Idee abweichen, ein reines 5G Netzwerk ohne LTE Komponenten zu betreiben. Außerdem wäre eine Konfiguration mit einer 5G NR Zelle als Anker und LTE als zusätzliche Kapazitätsschicht komplizierter gewesen, als die Konfiguration mit LTE als Anker der Verbindung. Der Grund dafür ist, dass in dieser Konfiguration ein 5G Kernnetz notwendig gewesen wäre, das zum Marktstart in 2019 zwar spezifiziert, aber noch nicht fertig entwickelt war.

Ein weiterer Grund, warum 5G als hybride 4G/5G Lösung gestartet wurde zeigt sich, wenn man den Start von 5G Netzwerken mit Starts von LTE Netzwerken 10 Jahre zuvor vergleicht. Damals gab es noch deutlich mehr freies Spektrum. In Europa wurde damals GSM von den meisten Netzbetreibern in 10 MHz Spektrum in den 900 und 1800 MHz Bändern betrieben. Für UMTS nutzten Netzbetreiber weitere 10 MHz Spektrum im 2100 MHz Band. In dieser Situation war es einfach, LTE Netzwerke auszurollen, da es genug Kapazität im 800 MHz Band gab, und, trotz Nutzung für GSM, auch noch viel freies Spektrum im 1800 MHz Band vorhanden war. Zusätzlich gab es noch deutlich mehr freies Spektrum im 2600 MHz (2,6 GHz) Band. Für 5G jedoch standen diese Bänder 10 Jahre später jedoch nicht mehr zur Verfügung, da diese heute komplett von LTE genutzt werden. Deshalb musste sich 5G zunächst mit Spektrum im 3,5 GHz Band und in Nordamerika oberhalb von 24 GHz begnügen, um höhere Datenraten und Netzkapazität zu ermöglichen.

2.1.2 Frequency Range 1 und 2

Während die meisten Netzbetreiber 5G Netzwerke in Frequenzbändern unterhalb von 6 GHz in Betrieb nahmen, starteten manche Netzbetreiber in den USA zunächst in Frequenzbändern oberhalb von 24 GHz. Da die Ausbreitungsbedingungen sich in diesen Bändern deutlich von den bisher für Mobilfunk genutzten Frequenzbereichen unterscheiden, entschloss sich 3GPP, die Physical Layer Spezifikationen in zwei Teile aufzuteilen. Der erste Teil ist dabei für alle Frequenzen unterhalb von 6 GHz zuständig. Dieser

Bereich wird als Frequency Range 1 (FR1) bezeichnet. Spektrum über 6 GHz wird als Frequency Range 2 (FR2) oder auch als Millimeter Wellen (mmWave) Spektrum bezeichnet. Der große Vorteil des mmWave Spektrum sind die sehr hohen Bandbreiten, die in diesem Spektrum zur Verfügung stehen. Während die maximale Bandbreite eines 5G Trägers 100 MHz in FR1 beträgt, können Träger in FR2 bis zu 400 MHz breit sein. Der größte Nachteil von solch hohen Frequenzen ist jedoch, dass die Datenübertragung nur bei direktem Sichtkontakt zwischen Sender und Empfänger möglich ist und auf eine Distanz von wenigen 10 m beschränkt ist. Deshalb sind FR2 Zellen hauptsächlich innerhalb von Gebäuden wie Bahnhöfen, Flughäfen und Ausstellungshallen nützlich, da hier an vielen Orten eine Sichtverbindung zwischen mmWave Zellen und Endgeräten möglich ist. Da es in Europa bisher keine FR2 Netzwerke gibt, unterstützen in Europa verkaufte Endgeräte den mmWave Frequenzbereich typischerweise nicht, da ihnen die entsprechende Hardware fehlt. Ob es in Zukunft auch FR2 Netzwerke in Europa geben wird, ist zum Zeitpunkt dieser Auflage noch ungewiss.

2.1.3 Dynamic Spectrum Sharing in Low- und Mid-Bands

Der Einsatz von 5G in Low- und Mid-Band Spektrum ist in den meisten Teilen der Welt nicht ohne Weiteres möglich, da dieses Spektrum typischerweise schon von LTE verwendet wird. Gerade in einer Übergangszeit, in der viele Nutzer noch kein 5G-fähiges Endgerät besitzen, werden viele Netzbetreiber nur sehr ungern die LTE Kapazität dort verringern wollen. Aus diesem Grund wurde in 3GPP ein Verfahren namens Dynamic Spectrum Sharing (DSS) entwickelt, das es erlaubt, LTE und 5G im gleichen Übertragungskanal zu verwenden. Wie später in diesem Kapitel noch genauer gezeigt wird, haben die 4G und 5G Luftschnittstelle viele Gemeinsamkeiten. Wird die 5G Luftschnittstelle entsprechend konfiguriert und 4G sowie 5G Kontrollkanäle ausgestrahlt, kann die Basisstation dann die Kanalkapazität dynamisch je nach Verkehrsaufkommen an 4G und 5G Geräte zuweisen. Auf diese Weise kann dann z. B. der Kanal hauptsächlich für 4G Endgeräte verwendet werden, solange es nur wenige 5G Geräte im Netz gibt. Bei zunehmender Anzahl an 5G Endgeräten in einer Zelle kann dann der 4G/5G Mix entsprechend automatisch angepasst werden.

2.1.4 Netze in der Praxis und weiterer Aufbau des Kapitels

Wie auch vorangegangene 3GPP Releases enthält Release 15 eine Flut an neuen Funktionen, von denen in der Praxis nur ein kleiner Bruchteil Verwendung finden dürfte. Dieses Kapitel fokussiert sich somit auf die Funktionen, die in Netzen heute im praktischen Einsatz sind, oder in Kürze Einzug finden dürfen. Der erste Teil des Kapitels widmet sich zunächst dem 5G NR Non-Standalone (NSA) Ansatz, dem neuen Radionetzwerk, den angepassten LTE Kernnetzelementen, den Teilen der 5G

NR Luftschnittstelle, die für die Non-Standalone Nutzung notwendig sind, sowie den Management Operationen, um 5G NR Zellen zusätzlich zu einer LTE Verbindung zu nutzen. Darauf folgt dann die Beschreibung der 5G Standalone (SA) Netzarchitektur und eine Einführung in die Teile der 5G Luftschnittstelle und Signalisierungsprozeduren, die für 5G SA notwendig sind. Im letzten Teil des Kapitels werden dann Funktionen beschreiben, um mit 5G neben schnellem Internetzugang weitere Anwendungsfelder zu erschließen.

2.2 Die 5G Non-Standalone (NSA) Architektur

2.2.1 Netzwerk Architektur und Schnittstellen

Die meisten 5G Netze, die bisher international in Betrieb genommen wurden, starteten mit der 5G New Radio (NR) Non-Standalone Architektur (NSA). Abb. 2.1 gibt einen Überblick über diese Konfiguration, und die meisten Netzwerkkomponenten wurden schon im Kapitel über LTE beschrieben. Im Kernnetz leiten das Packet Data Network Gateway (PDN-GW) und das Serving-GW (S-GW) IP Pakete zwischen Endgerät (User Equipment, UE) und Internet weiter. In den Standards wird der Transport von IP Paketen des Benutzers auch als User Plane (UP) bezeichnet. Die Mobility Management Entity

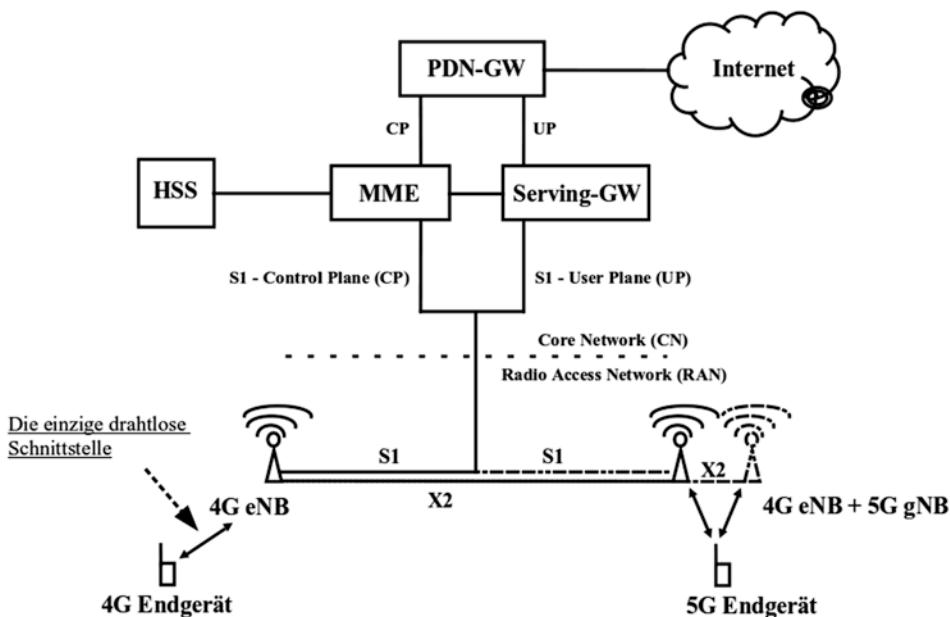


Abb. 2.1 Die 5G New Radio Non-Standalone Architektur

(MME) im Kernnetz ist dagegen für Aufgaben wie die Authentifizierung der Nutzer und für das Mobilitätsmanagement verantwortlich und somit Teil der Control Plane (CP).

Und schließlich wird im Kernnetzwerk auch die Home Subscriber Server (HSS) Datenbank benötigt, in der für jeden Teilnehmer Authentifizierungs- und Service-informationen hinterlegt sind. Für den Internetzugang enthält das HSS pro Teilnehmer Informationen wie z. B. die Quality of Service Einstellungen und die maximale Datenrate, die einem Teilnehmers gewährt wird. Für die Sprachtelefonie enthält das HSS Information wie die Telefonnummer jedes Teilnehmers und seine Einstellungen für die Anrufweiterleitung. Details hierzu werden im Kapitel über LTE beschrieben.

Sowohl die User Plane als auch die Control Plane sind IP-basiert und IP Pakete von beiden Planes werden über das gleiche physische Interface zwischen dem Radionetzwerk (Radio Access Network, RAN) und dem Kernnetzwerk übertragen. Diese Schnittstelle wird als S1-Interface bezeichnet. In LTE besteht das RAN aus den Komponenten an den Sendestandorten und den Backhaul Verbindungen zum Kernnetzwerk. Hier werden typischerweise Glasfaserleitungen oder Mikrowellenfunk verwendet.

An einem Sendestandort befinden sich typischerweise mehrere Base Stations, die im Standard als eNode-Bs (eNBs) bezeichnet werden. Diese bestehen, wie in Abb. 2.2 gezeigt, aus folgenden Komponenten:

- **Den digitalen Baseband Units (BBU).** Diese werden mit dem Kernnetz über Glasfasern oder per Mikrowelle verbunden. Die BBUs sind für das Management einer Mobilfunkzelle verantwortlich und für die Erzeugung und Dekodierung des digitalen „Baseband“ Signals für alle Radiotechnologien (siehe Kapitel über LTE).
- **Den Remote Radio Units (RRUs).** Diese wandeln den per Glasfaser erhaltenen digitalen Baseband Datenstrom von der BBU in ein analoges Signal um, das dann an die Antenne weitergeleitet wird. Die Ausgangsleistung liegt dann typischerweise zwischen 10 und 200 W. In Empfangsrichtung verstärken die RRUs zunächst das sehr schwach empfangene analoge Signal und wandeln es dann in einen digitalen Datenstrom um, der dann anschließend von der BBU bearbeitet wird.
- **Den Antennen.** Flache Panelantennen, die intern aus mehreren Antennen für unterschiedliche Frequenzbereiche bestehen, werden über Kupferkabel mit den Remote Radio Units (RRUs) verbunden.

Für 5G Kanäle im 3,5 GHz Band mit bis zu 100 MHz Bandbreite werden typischerweise weitere BBUs installiert, da zusätzliche Rechenkapazität am Standort benötigt wird. Eine BBU Einheit kann heute die digitale Signalverarbeitung aller am Standort verwendeten Radiotechnologien durchführen. Um die Kapazität zu erhöhen, wird der Standort jedoch in 3 Sektoren unterteilt. Jeder Sektor hat dann eine oder mehrere eigene BBUs, da eine einzelne Einheit nicht genug Rechenkapazität für den kompletten Datenverkehr eines Standortes bereitstellen kann.

Statt flacher Panelantennen und Remote Radio Units (RRUs) werden für 5G Kanäle im 3,5 GHz Band heute hauptsächlich aktive Antennensysteme (Active Antenna

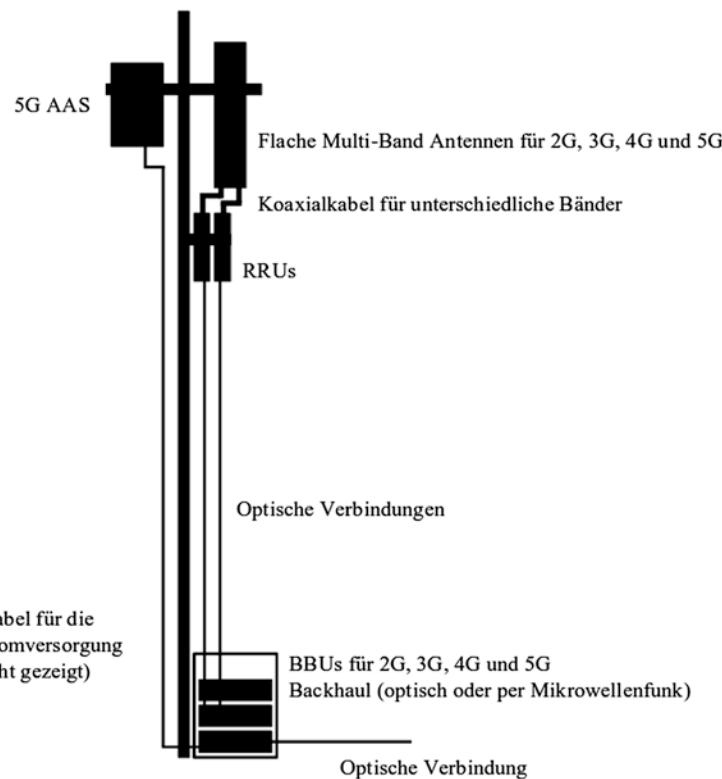


Abb. 2.2 Komponenten an einem LTE/5G Sendestandort

Systems, AAS) mit 32 oder 64 kleinen Antennen verwendet. Diese steigern durch das später noch genauer beschriebene Beamforming und mit dem Multi-User MIMO Verfahren die Kapazität eines Standortes deutlich.

Zusätzlich zu den kleinen Antennenelementen enthalten Aktive Antennen auch die Remote Radio Units. Dies wurde notwendig, da es nicht praktikabel ist, eine solch hohe Anzahl von Antennenelementen über Koaxialkabel mit einem Radiomodul zu verbinden. Dadurch sind solche Antennensysteme natürlich deutlich schwerer als passive Panelantennen. Andererseits wird nur noch ein Glasfaserkabel und eine Stromversorgung für die Antenne benötigt, statt diese mit vielen dicken Koaxialkabeln zu verbinden. Natürlich ist es auch möglich, billigere Panelantennen für 5G im 3,5 GHz Band zu verwenden. Hier werden dann typischerweise 8×8 MIMO-Antennen verwendet, mit denen zumindest eine Reichweiterhöhung erzielt werden kann. In Bändern unterhalb von 3,5 GHz werden für 5G Kanäle üblicherweise die bereits für 2G, 3G und 4G installierten Antennen mitbenutzt.

In LTE wurde das X2 Interface eingeführt, um die unterschiedlichen Sektoren eines Standortes und unterschiedliche Standorte miteinander zu verbinden und somit schnelle

Handover zu ermöglichen. Dies ist im LTE Radionetzwerk (RAN) sehr nützlich, da hier keine zentrale Instanz, sondern die eNBs selber entscheiden, wann eine Verbindung zu einem anderen eNB weitergegeben werden soll. Wie auch alle anderen Schnittstellen in LTE ist das X2 Interface IP basiert. Da 5G Zellen logisch unabhängig vom LTE System sind, wird auch hier das X2 Interface verwendet, um diese mit dem LTE Netzwerk zu kombinieren. Dazu waren nur wenige X2 Protokollerweiterungen notwendig. 5G Zellen können somit mit LTE Zellen am gleichen und auch an anderen Standorten zusammengeschaltet werden. Falls eine BBU an einem Standort für 4G und 5G Zellen zuständig ist, über die ein Nutzer gerade Daten austauscht, wird ein virtuelles X2 Interface innerhalb der BBU verwendet. Zwischen Sektoren eines Standorts und zwischen unterschiedlichen Standorten wird das X2 Interface über Glasfaserkabel oder Mikrowellenverbindungen übertragen. Benachbarte Standorte sind jedoch üblicherweise nicht direkt miteinander verbunden. Stattdessen verwendet das X2 Interface die Backhaul Glasfaserleitung des S1 Interface bis zum nächsten IP Router am Rand des RANs (Aggregation Router). Von dort werden dann die IP Pakete des X2 Interface zurück zum eigentlich benachbarten Standort geleitet.

In LTE Netzwerken wird die Kombination aus Antennen, Radio- und Baseband Modulen als eNode-B oder eNB bezeichnet. Bei 5G NR wurde die Bezeichnung gNB gewählt. Diese Bezeichnungen haben ihren Ursprung in den 3G UMTS Spezifikationen, dort wurde der Radiostandort als Node-B bezeichnet. LTE griff dann diesen Namen auf und stellte ein ‚e‘ für ‚Evolution‘ voran. Für 5G wurde das ‚e‘ durch ein ‚g‘ ersetzt, und steht dort für die ‚nächste Generation‘³.

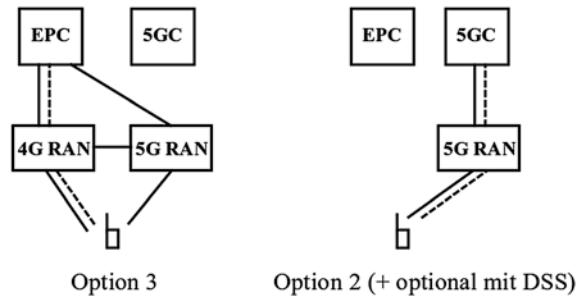
2.2.2 3GPP 5G NR Option 2 und 3 mit Dynamic Spectrum Sharing

Während die meisten 5G Netze zunächst mit der Non-Standalone (NSA) Architektur in Betrieb genommen wurden, war dies nur ein erster Schritt. Langfristiges Ziel ist es jedoch, alle Teilnehmer und Anwendungen in ein 5G Standalone (SA) Netz zu überführen. Um von einer NSA zu einer SA Architektur zu gelangen und dabei das Netzwerk abwärtskompatibel zu halten, wurden von 3GPP unterschiedliche Möglichkeiten spezifiziert, die die Nummern 1–7 tragen. Mit zusätzlichen Buchstaben hinter der Nummer werden unterschiedliche Unterimplementierungen unterschieden. In der Praxis haben sich jedoch nur folgende Optionen durchgesetzt:

- Option 3x: Die heute verwendete Non-Standalone Architektur (NSA)
- Option 2: Die heute verwendete Standalone Architektur (SA)
- Dynamic Spectrum Sharing (DSS): Verwenden eines Kanals für LTE und 5G mit Option 3x NSA.

Abb. 2.3 stellt die Optionen 3 (NSA) mit optionalem DSS und Option 2 (SA) gegenüber, die später in diesem Kapitel beschrieben wird. Die gestrichelten Linien symbolisieren

Abb. 2.3 Gegenüberstellung von Option 3x (NSA) und Option 2 (SA)



die Signaling Plane zwischen Endgerät und Kernnetz, die durchgezogenen Linien die User Plane.

In der Praxis hat sich der gleichzeitige Einsatz von 5G NSA (Option 3x) und 5G SA (Option 2) durchgesetzt, sowie Dynamic Spectrum Sharing als Zwischenschritt, also die gleichzeitige Nutzung eines Kanals für LTE und 5G NR. Alle anderen spezifizierten Optionen werden heute in der Praxis nicht eingesetzt und werden deshalb hier auch nicht weiter beschrieben. Endgeräte verwenden dann in Abhängigkeit ihrer unterstützten 5G Funktionen das Netzwerk in unterschiedlicher Weise:

- Neue und somit Option 2-fähige 5G Endgeräte die den Protokollstack des 5G Kernnetz verstehen, können 5G NR Carrier Aggregation verwenden und kommunizieren nur über 5G gNBs.
- Ältere 5G Non-Standalone Endgeräte, die nur 5G NR Option-3 unterstützen, benötigen eine 4G Zelle als Anker und ein 4G Kernnetz. Zusätzliche Kapazität wird durch 4G und optional mit 5G Carrier Aggregation hinzugenommen.
- Noch ältere LTE-only Endgeräte nutzen nur 4G Carrier Aggregation zusammen mit dem 4G Kernnetz.

Dies ermöglicht in der Praxis eine Koexistenz der unterschiedlichen 4G und 5G Endgeräte. In Kanälen, in denen Dynamic Spectrum Sharing verwendet wird, wird nicht nur Bandbreite für die 4G Signalisierung, sondern auch für die 5G Signalisierung benötigt. Dies reduziert die Kanalkapazität um etwa 10–15 %. Somit ist die Kapazität in diesen Übertragungskanälen für die Nutzdatenübertragung etwas geringer.

2.2.3 Unterschiede der Optionen 3, 3A und 3X

Viele Netzbetreiber begannen ihren 5G Rollout zunächst mit 5G Option 3. Diese Konfiguration wird auch als ‚EUTRAN New Radio – Dual Connectivity‘ (EN-DC) bezeichnet. Diese Abkürzung bezieht sich auf eine EUTRAN (LTE) Zelle als Master und Anker der Verbindung, die dann bei Bedarf 5G NR Zellen zum Übertragungskanal eines

Endgerätes hinzugeschaltet. Beschrieben ist diese Variante in 3GPP TS 37.340⁴. Netzwerk und Endgerät übertragen hier Daten gleichzeitig über 4G und 5G.

In diesem Setup wird die 4G LTE Basisstation, der eNB, auch als Master-eNB oder MeNB bezeichnet. Die 5G NR Basisstation, der gNB, wird als Secondary-gNB oder SgNB bezeichnet. Da ein Dual-Connectivity Setup üblicherweise mehr als eine LTE Zelle als Teil von Carrier Aggregation beinhaltet, wird oftmals auch der Begriff Master Cell Group (MCG) in den Spezifikationen verwendet, der dann den gesamten LTE Teil der Verbindung bezeichnet. Der 5G NR Teil der Verbindung wird dann als Secondary Cell Group (SCG) bezeichnet. Auch hier ist es möglich, mehrere Zellen, also mehrere Kanäle zu aggregieren. Ältere 5G NSA fähige Endgeräte können jedoch noch keine 5G Kanäle im Option 3 Setup kombinieren.

An dieser Stelle sei auch erwähnt, dass der Dual Connectivity Ansatz von LTE/NR schon in 3GPP Release 12 für LTE als Teil des ‚Heterogeneous Network‘ Konzepts spezifiziert wurde. Hier war die prinzipielle Idee, dass LTE Macro Zellen mit kleinen LTE Mikrozellen kombiniert werden könnten. Der Unterschied zum normalen LTE Carrier Aggregation ist, dass zwei oder mehr Zellen, die die Kanäle für eine Verbindung bereitstellen, unabhängig voneinander arbeiten können. Diese unabhängigen Zellen verwenden dann jeweils ihren eigenen Air-Interface Scheduler. Dies war notwendig, da die Makrozelle ein größeres Gebiet abdeckt, in dem dann mehrere kleine Zellen einen Unterbereich versorgen. Da die Zellen nicht am gleichen Ort stehen, ist es nicht möglich, die Ressource Verteilung auf der Luftschnittstelle zentral zu steuern, da die Laufzeit zu lange und evtl. die Bandbreite der Verbindung zwischen Makro- und Mikrozelle zu gering ist. Während in der Praxis das Konzept für LTE nie eingesetzt wurde, waren jedoch nur wenige Änderungen nötig, dieses für die Verbindung von 4G und 5G mit Option 3 zu verwenden. Auch in diesem Anwendungsfall ist somit der 5G Scheduler im gNB unabhängig vom 4G Scheduler im eNB.

Für EN-DC wird eine erweiterte Version des existierenden LTE X2 Interface verwendet, das ursprünglich für den schnellen Handover zwischen LTE Zellen vorgesehen war. Für EN-DC verbindet nun das X2 Interface den eNB und den gNB am gleichen oder auch an unterschiedlichen Standorten. Ein LTE eNB kann somit einen gNB als ‚Speed Booster‘ für eine Verbindung hinzuziehen. Die Daten werden dann über einen sogenannten ‚Split Bearer‘ übertragen, also gleichzeitig über das LTE und das 5G NR Air Interface. Dies ist notwendig, da heute sehr viel Bandbreite in unterschiedlichen Frequenzbereichen für LTE verwendet wird. Nur durch Kombination mit neuen Frequenzbereichen, die nur für 5G verwendet werden, kann eine 5G Datenrate erreicht werden, die höher als in LTE-only Netzwerken ist. Die Bezeichnung ‚Split-Bearer‘ sagt aus, dass die Nutzdatenpakete über 4G und 5G gleichzeitig und geteilt übertragen werden. In den Standards wurden 3 unterschiedliche EN-DC Varianten spezifiziert:

- **Option 3:** Der LTE eNB kommuniziert mit dem Kernnetz, empfängt Nutzdaten und leitet einen Teil davon über das X2 Interface zum 5G gNB weiter.

- **Option 3X:** Nutzdaten werden zwischen dem Kernnetzwerk und dem 5G gNB ausgetauscht. Der 5G gNB leitet dann einen Teil der Nutzdaten über das X2 Interface an den 4G eNB weiter. Dies ist die typische ‚Split-Bearer‘ Konfiguration, die heute in Netzwerken verwendet wird, da neue gNB Baseband Einheiten mit 10 Gbit/s Ethernet Ports ausgestattet sind, während schon existierende eNB BBUs typischerweise nur 1 Gbit/s Ports besitzen. Die Signalisierung und das Mobility Management werden jedoch weiterhin vom 4G eNB durchgeführt.
- **Option 3A:** Das Serving-Gateway im Kernnetzwerk kommuniziert sowohl mit dem eNB, als auch mit dem gNB. Diese Option ist zwar spezifiziert, wird aber in der Praxis nicht verwendet.

2.2.4 Das Fronthaul Interface

Abb. 2.2 zeigte bereits die wichtigsten Elemente eines Mobilfunkstandortes, also das S1 Interface über Glasfaser oder Mikrowelle zum Kernnetz, die Baseband Units (BBUs) für die digitale Signalverarbeitung, die Remote Radio Unit (RRU), die das digitale Baseband Signal in ein analoges Radio Signal umwandelt, sowie die Antennen. Die Schnittstelle zwischen der BBU, die üblicherweise in einem Schaltschrank am Fuß eines Mobilfunkstandortes eingebaut ist und den Remote Radio Units, die typischerweise direkt unterhalb der Antenne angebracht sind, wird als Fronthaul bezeichnet. Aus Redundanzgründen verwendet das Fronthaul Interface typischerweise zwei Glasfaserleitungen pro Sektor und Frequenzband. Ein typischer 3-Sektor Standort mit 3 Frequenzbändern verwendet somit 12 Glasfaserkabel. Da Glasfaserkabel preiswert und im Vergleich zu Koaxialkabel sehr dünn sind, bietet eine direkte Installation der RRUs an der Antenne einen großen Vorteil. Zusätzlich wird nur noch ein Kabel für die Spannungsversorgung der RRUs benötigt.

Das Fronthaul Interface verwendet das Evolved Common Public Radio Interface (eCPRI) Protokoll⁵, um die digitalen I/Q Informationen, die von einer BBU generiert werden, an die Remote Radio Units weiterzuleiten. Auf Layer 1 wird Glasfaser verwendet, auf Layer 2 typischerweise das Ethernet Protokoll und optional IP auf Layer 3. Details zur Übertragung eines analogen Radiosignals als digitale I/Q Information befinden sich im LTE Kapitel zum Thema Quadrature Amplitude Modulation. Die RRUs konvertieren dann den digitalen I/Q Datenstrom in ein analoges RF Signal, verstärken dieses und senden es dann weiter zur Antenne. In der umgekehrten Richtung empfangen die RRUs ein analoges Signal von der Antenne, wandeln es in einen digitalen I/Q Datenstrom um und senden diesen dann über die Glasfaserleitung zur BBU, die dann die digitale Signalverarbeitung übernimmt.

2.3 Die 5G TDD Luftschnittstelle

Ältere Mobilfunktechnologien inklusive 3G UMTS verwendeten ausschließlich eine Frequency Division Duplex (FDD) Luftschnittstelle. Hier wurde die Downlink-Übertragung von Netzwerk zu Endgerät und die Uplink-Übertragung in der Gegenrichtung auf unterschiedlichen Frequenzen durchgeführt. Ein Schutzabstand (Guard Band) trennte diese Bereiche. Ein 5 MHz 3G UMTS FDD Kanal nutzte somit einen 5 MHz breiten Kanal für den Downlink und einen separaten 5 MHz breiten Kanal für den Uplink. In 4G LTE begannen dann Netzbetreiber vor allem in den USA und Asien, auch eine Time Division Duplex (TDD) Luftschnittstelle zu verwenden. Hier werden Downlink und Uplink im selben Kanal übertragen und über Zeitschlitz voneinander getrennt. Auch in Europa verwenden heute einzelne Netzbetreiber für manche Kanäle dieses Übertragungsverfahren, und es ist sogar möglich, FDD und TDD Kanäle mit Carrier Aggregation zu kombinieren. Vor allem höherpreisige Geräte unterstützen dies heute.

In 5G NR unterschieden sich die Startbedingungen jedoch von denen früherer Netzwerke deutlich. Erste 5G Netzwerke setzten vor allem auf TDD Kanäle, um damit signifikante Geschwindigkeitsvorteile und Kapazitätserweiterungen im Vergleich zu LTE zu ermöglichen. 5G NR FDD Kanäle wurden erst später hinzugefügt, um Teile des Spektrums in niedrigeren Frequenzbereichen zu migrieren, die zuvor schon für UMTS oder LTE verwendet wurden. Dies wird im Laufe des Kapitel noch genauer beschrieben.

Die Entscheidung, bei 5G NR zunächst auf TDD zu setzen, hatte mehrere Gründe. Ein wichtiger Grund war, dass in Mobilfunknetzen hauptsächlich Daten im Downlink übertragen werden. Ein übliches Verhältnis zwischen dem Datenvolumen im Downlink und Uplink ist 10:1. Tritt eine Überlast an einem Mobilfunkstandort auf, tritt diese üblicherweise im Downlink und nicht im Uplink auf. Zwar sind die möglichen Datenraten im Uplink wegen niedriger Sendeleistung und Antennenkonfiguration geringer, es ist jedoch trotzdem nicht sehr ökonomisch, die gleiche Kanalbandbreite für den Downlink und den Uplink zu nutzen.

Ein zweiter Faktor war der Umstand, dass besonders in den USA Spektrum im 28 GHz Bereich für erste 5G Netzwerke verwendet wurde, das bisher nicht im Mobilfunk für die Luftschnittstelle verwendet wurde. Dort unterscheiden sich die Ausbreitungsbedingungen deutlich von Frequenzbereichen unterhalb von 6 GHz, die bisher schon verwendet wurden. Eine nur sehr geringe Reichweite und eine hohe Signaldämpfung wird durch bis zu 400 MHz breite Kanäle kompensiert, und es wäre hier besonders ungünstig gewesen, Downlink und Uplink im Frequenzbereich zu trennen. Frequenzbereiche oberhalb von 6 GHz werden in den 3GPP Spezifikationen als Frequency Range 2 (FR2) bezeichnet und später in diesem Kapitel noch genauer behandelt.

Spektrum unterhalb von 6 GHz wird als Frequency Range 1 (FR1) bezeichnet. Ein letzter großer Bereich, der dort noch für den Mobilfunk außerhalb den USA zugeteilt werden konnte, war, wie in Tab. 2.1 gezeigt, 400–500 MHz Spektrum zwischen 3,3 und

Tab. 2.1 Frequenzbänder für 5G TDD in FR1

Band Nummer	Frequenzbereich	Region
n78	3300–3800 MHz	Europa, Asien
(n77)	3300–4200 MHz	Band n78 ist Teilmenge, Japan
n79	4400–5000 MHz	China, Japan
n41	2496–2690 MHz	USA, nur ein Netzbetreiber

3,8 GHz. Auch hier wurde beschlossen, das Spektrum mit TDD zu verwenden. Somit ist hier kein Guard Band zwischen Uplink und Downlink notwendig und der Großteil des Spektrums kann für den Downlink verwendet werden. Auch ist heute zu beobachten, dass LTE Netzwerke 2 bis 5 LTE Kanäle in Downlink Richtung bündeln, während die Uplink Kanalbündelung meist auf zwei Kanäle auch bei hochpreisigen Endgeräten limitiert ist. Dies liegt hauptsächlich daran, dass für jeden Kanal eine eigene und teure Sendeeinheit benötigt wird.

Nachfolgend wird nun zunächst die 5G NR TDD Luftschnittstelle für den Frequenzbereich 1 (FR1) betrachtet. In Europa und Asien wird 5G NR TDD vor allem im 3,5 GHz Bereich betrieben, der auch als Band n78 bezeichnet wird. In Ländern wie Japan und China wurde vom nationalen Regulierer außerdem zusätzliches Spektrum zwischen 4,4 und 5 GHz für 5G Netzwerke bereitgestellt. In den USA waren diese Frequenzbereiche nicht für den Start für 5G verfügbar. Ein Netzbetreiber entschied sich deshalb, 5G NR TDD im 2,5 GHz Bereich (Band n41) zu starten, da dieser dort noch viel ungenutztes Spektrum zur Verfügung hatte. Die Bandnummern von 5G NR sind zu den LTE Bandnummern identisch, es wird jedoch ein „n“ vorangestellt, um die Nutzung des Bandes für 5G NR zu kennzeichnen.

2.3.1 Flexibles OFDMA für den Downlink

Grundsätzlich haben die 5G NR und die LTE Luftschnittstelle viele Gemeinsamkeiten. Wichtigste Gemeinsamkeit ist die Verwendung von Orthogonal Frequency Division Multiplexing (OFDM, siehe Kapitel über LTE). Da LTE nur für schnelles Internet und nur für Frequenzbereiche unterhalb von 3 GHz spezifiziert wurde, konnten große Teile der benötigten Parameter für den Physical Layer statisch spezifiziert werden. Da 5G NR auch Frequenzbereiche oberhalb von 3 GHz abdeckt und zusätzlich auch das mmWave Spektrum oberhalb von 24 GHz nutzt, sind sehr viele Physical Layer Parameter nun konfigurierbar. Diese Flexibilität wird in den 5G NR Spezifikationen auch als „Numerology“ bezeichnet⁶.

Einer der wichtigsten Physical Layer Parameter in LTE und NR ist die OFDM Subcarrier Bandbreite. In LTE ist dieser Parameter mit 15 kHz spezifiziert. Ein Symbol wird dabei in 66,67 µs übertragen und es gibt somit genügend Zeit, den Delay Spread des

Signals durch Reflexionen eines Teil des Signals an unterschiedlichen Gegenständen wie Hauswänden und Bäumen zu kompensieren, oder sogar für die MIMO Datenübertragung zu verwenden. Ein anderer Grund, 15 kHz Subcarrier zu verwenden, war die Limitierung der Anzahl an Subcarrier, die in einem 20 MHz Kanal zu dekodieren sind und somit eine vertretbare Komplexität des Empfängers zu ermöglichen.

Mit der Spezifikation der 5G NR Luftschnittstelle wurde die Kanalbandbreite der Subcarrier flexibler und kann jetzt Werte von 15, 30 60 und 120 kHz im FR1 Bereich annehmen. In der Praxis wird heute eine Kanalbandbreite von 30 kHz im 3,5 GHz n78 Band verwendet. Dies bedeutet, dass die Übertragung eines Symbols nur halb so lange dauert als bei LTE. Eine Erhöhung der Subcarrierbreite war möglich, da die Reichweite eines Signals bei 3,5 GHz wesentlich geringer als in Frequenzbereichen z. B. um 800 MHz ist. Somit kann auch die Übertragungszeit der einzelnen Symbole reduziert werden, da weniger Delay Spread auftreten kann. In Kombination mit anderen Maßnahmen ist es somit möglich, die Verzögerungszeit für Datenpakete gegenüber LTE etwas zu senken. An dieser Stelle ist es interessant anzumerken, dass 60 und 120 kHz Subcarrier bisher nicht im FR1 Bereich verwendet werden, während eine Bandbreite von 15 kHz pro Subcarrier für 5G FDD Datenübertragung in niedrigeren Frequenzbereichen angewandt wird.

Zusätzlich sieht der Standard auch noch vor, unterschiedliche Subcarrier-Bandbreiten in unterschiedlichen Teilen eines Kanals gleichzeitig zu verwenden. Auf diese Weise können unabhängige „Inseln“ innerhalb des Kanals konfiguriert werden. Somit ist es möglich, 15 kHz Subcarrier in manchen Teilen des Kanals z. B. für schnelles Internet zu nutzen und 60 kHz Subcarrier für Geräte und Anwendungen in einem anderen Teil zu verwenden, die eine geringe Latenz benötigen. Und wieder andere Bereiche des Kanals können für Dienste konfiguriert werden, die eine hohe Übertragungssicherheit benötigen, aber keine großen Anforderungen an die Verzögerungszeit haben. Geräte, die den Kanal für den Internetzugang verwenden, ignorieren dann die Teile des Kanals, die eine andere Konfiguration verwenden. Auch das Scheduling Intervall, das in LTE immer 1 ms lang ist, wurde bei 5G NR flexibilisiert. Bleibt anzumerken, dass auch 2022 diese Flexibilität noch größtenteils ungenutzt ist, da aktuelle 5G Netzwerke hauptsächlich für einen schnellen Internetzugang sorgen. LTE Erweiterungen wie NB-IoT und CAT-M1 gibt es bisher in der Praxis bei 5G noch nicht. In der Zukunft ermöglicht die Flexibilität der 5G Luftschnittstelle jedoch, solche Erweiterungen zu spezifizieren und einzusetzen, ohne rückwärtskompatibel sein zu müssen. Dies ist eine Lehre aus LTE. Hier musste man bei der Spezifikation der Narrowband Internet of Things (NB-IOT) Erweiterung viele Kompromisse eingehen, damit LTE Endgeräte weiterhin in bestehenden Netzen funktionierten.

In Frequenzbändern unterhalb von 6 GHz (FR1) ist die maximale Kanalbandbreite von 5G NR auf 100 MHz erweitert worden. Ein 5G NR Kanal kann somit bis zu 5 Mal breiter als ein LTE Kanal sein, der auf maximal 20 MHz spezifiziert wurde.

Da viele Netzbetreiber weniger Spektrum zur Verfügung haben, ist die Kanalbandbreite in Schritten von 5 und 10 MHz definiert. Die verfügbare Kanalbandbreite hängt von diversen Faktoren ab, wie z. B. dem Frequenzband, der Anzahl der Netzbetreiber in einem Land und wie viel Spektrum in einem Band freigegeben wurde. In den meisten Teilen der Welt außer in den USA sind im Band n78 im 3,5 GHz Bereich bis zu 500 MHz Spektrum zugeteilt worden. In Ländern wie Finnland und Korea wurde genügend Spektrum freigegeben, um jedem Netzbetreiber mindestens einen 100 MHz Kanal zu ermöglichen. In anderen Ländern wurde zum Teil, wie in Tab. 2.2. gezeigt, weniger Spektrum freigegeben. In Deutschland wurde z. B. nur 300 MHz Spektrum in Band n78 an die öffentlichen Mobilfunknetzbetreiber versteigert. Es folgte ein scharfer Bieterwettbewerb zwischen vier Interessenten und entsprechend hohen Auktionserlösen. Letztlich konnten zwei Netzbetreiber jeweils 90 MHz ersteigern, also 10 MHz weniger als die maximal mögliche Kanalbandbreite. Die zwei anderen Bieter konnten 50 und 70 MHz an Spektrum in Band n78 ersteigern. In Großbritannien sind drei der vier Netzbetreiber sogar auf nur 40 bis 50 MHz in Band n78 beschränkt. Es gibt somit in Europa große Unterschiede, wie viel zusätzliche Netzkapazität Netzbetreiber in dicht besiedelten Gebieten und Hotspots wie Bahnhöfen, Flughäfen, Messegeländen und Stadien mit 5G bereitstellen können. In manchen Ländern wurde bisher auch nur ein Teil des n78 Bandes an Netzbetreiber verteilt. Somit können diese in Zukunft unter Umständen weiteres Spektrum beziehen.

Für 5G wurde im Vergleich zu LTE eine leicht modifizierte OFDM Version spezifiziert. Diese Version wurde CP-OFDM genannt und wird sowohl im Uplink als auch im Downlink verwendet⁷. Als Alternative wurde auch DFT-S-OFDM für den Uplink spezifiziert.

Tab. 2.2 Typische Beispiele für Zuteilungen im Band n78 (3,5 GHz) in 2022

Land	Netzbetreiber (Auswahl)	Spektrum in Band n78 (3,5 GHz)
Deutschland	Vodafone, Deutsche Telekom	90 MHz
Deutschland	Telefonica	70 MHz
Deutschland	1&1	50 MHz
Italien	TIM, Vodafone	80 MHz
China	China Mobile, China Unicom	100 MHz
Korea	KT, SK Telecom	100 MHz
Korea	LG U+	80 MHz
Finnland	DNA, Telia, Elisa	130 MHz
Großbritannien	3 UK	80 MHz + 40 MHz (nicht zusammenhängend!)
Großbritannien	Vodafone	50 MHz
Großbritannien	Telefonica	40 MHz

2.3.2 Das 5G Resource Grid: Symbole, Slots, Resource Blocks und Frames

Die kleinste Datenübertragungseinheit auf der Luftschnittstelle ist das Symbol. In einem Symbol können in Abhängigkeit der Modulation mehrere Datenbits übertragen werden. Tab. 2.3 gibt einen Überblick über die Modulationsarten, die in NR im Downlink verwendet werden, sowie die dazugehörige Anzahl Bits. In der Praxis zeigt sich, dass die 256-QAM Modulation nur bei sehr guten Übertragungsbedingungen verwendet wird, die nur in einem sehr kleinen Teil einer Zelle vorhanden sind, also meist auf sehr kurze Distanz. Im Uplink wird 16-QAM im größten Teil der Zelle verwendet und 64- oder 256-QAM bei sehr guten Bedingungen.

Die Übertragungszeit eines Symbols ist von der gewählten ‚Numerology‘ des Kanals abhängig. Werden, wie in einer typischen n78 Konfiguration, 30 kHz Subcarrier verwendet, beträgt die Übertragungszeit eines Symbols 33,33 µs. Werden 15 kHz Subcarrier verwendet, eine typische Konfiguration für NR FDD Kanäle unterhalb von 3 GHz, beträgt die Übertragungszeit 66,67 µs. Hier ist die Übertragungszeit somit zur Übertragungszeit eines LTE Symbols identisch. Dies ist wichtig, da dies der Grundstein für die gleichzeitige Nutzung von LTE und NR auf einem Übertragungskanal ist. Weitere Details dazu finden sich später in diesem Kapitel.

Um die Entscheidungen bei der Spezifikation der NR Luftschnittstelle besser zu verstehen, lohnt sich zunächst ein Blick auf die Struktur der LTE Luftschnittstelle. Bei LTE werden 7 Symbole in einen Slot auf der Zeitachse zusammengefasst, wobei zwei Slots einen Subframe bilden. 10 Subframes werden dann zu einem LTE Radio Frame zusammengefasst. Die Slotlänge ist mit 0,5 ms spezifiziert, die Länge eines Subframes ist somit 1 ms. Das Scheduling wird bei LTE auf Subframe Basis organisiert, d. h. das Netzwerk kann jede Millisekunde entscheiden, von und zu welchen Endgeräten Daten übertragen werden.

Auf der NR Luftschnittstelle besteht ein Slot aus 14 Symbolen. Deren Übertragungszeit bestimmt sich aus der gewählten Numerology. Ein Slot ist die kleinste Scheduling Einheit, wobei Slot Aggregation erlaubt ist. Nicht-Slot basiertes Scheduling wurde zusätzlich für Applikation definiert, die eine sehr geringe Latenz benötigen. Dieses Konzept wurde Mini-Slot Scheduling genannt, wird aber in der Praxis bisher nicht verwendet.

Tab. 2.3 Modulationsarten der NR Luftschnittstelle

Modulationsart	Bits pro Symbol (Übertragungsschritt)
QPSK (Quadrature Phase Shift Keying)	2
16-QAM (Quadrature Amplitude Modulation)	4
64-QAM	6
256-QAM	8

Während NR das Scheduling auf Slot Basis durchführt, verwendet LTE den Subframe als kleinste Scheduling Einheit. Darum hat auch die Länge des Subframe bei NR keine Bedeutung für das Scheduling mehr. Die Übertragungszeiten für Subframes und Frames sind auf 1 und 10 ms festgelegt, die Anzahl der Slots pro Subframe und per Frame sind daher von der Bandbreite des Subcarriers abhängig. Mit einer auf Band n78 üblichen Subcarrierbreite von 30 kHz passen zwei Slots in einen 1 ms Subframe. Werden 15 kHz Subcarrier in Bändern unterhalb von 3 GHz verwendet, gibt es nur ein Slot pro Subframe.

Wie in LTE werden auch auf der NR Luftschnittstelle mehrere Symbole, die auch als Resource Elemente (RE) bezeichnet werden, in einem Physical Resource Block (PRB) zusammengefasst. Ein PRB ist die kleinste Einheit, die einem Endgerät für die Datenübertragung zugeteilt werden kann. In LTE werden 12 Symbole auf der Frequenzachse und 7 Symbole auf der Zeit Achse (1 Slot) zu einem PRB zusammengefasst. NR bündelt ebenfalls 12 Symbole auf der Frequenzachse, während die Anzahl der Symbole auf der Zeitachse variabel ist und von der gewählten Numerology abhängt.

Abb. 2.4 zeigt eine typische Konfiguration der Luftschnittstelle für Band n78 mit einer Kanalbandbreite von 100 MHz. Auf der Frequenzachse werden 273 Physical Resource Blocks verwendet. Dies entspricht $273 * 12$ (Subcarrier) * 30 kHz (Breite eines Subcarriers)=98,28 MHz. Der Rest wird als Schutzabstand zum nächsten Kanal (Guard Band) an beiden Seiten des Kanals verwendet. Für einen 90 MHz Carrier werden 243 PRBs auf der Frequenzachse verwendet.

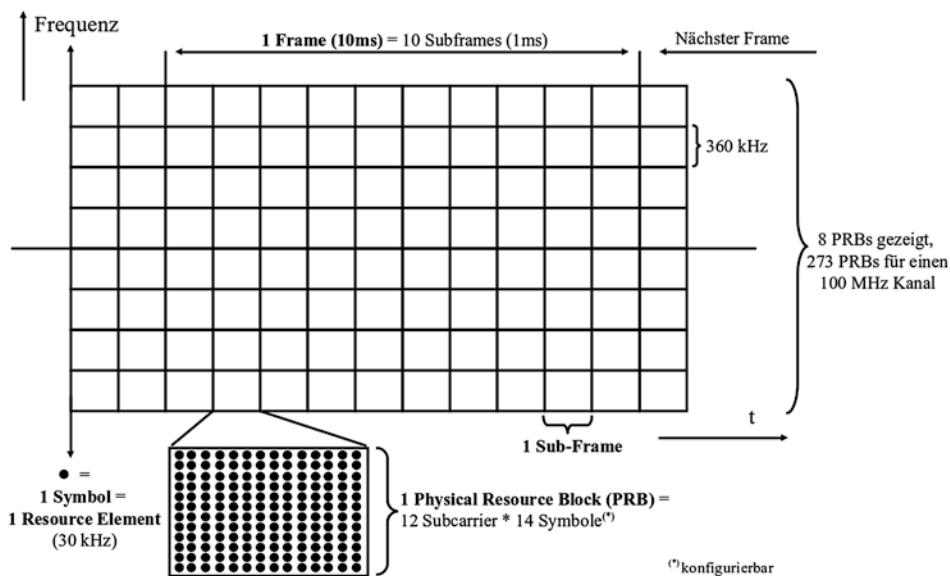


Abb. 2.4 Typische Konfiguration der NR Luftschnittstelle in Band n78

2.3.3 Synchronisation und Referenz Signale

Um das Netzwerk nach dem Einschalten zu finden und danach synchronisiert zu bleiben, sucht ein Endgerät zunächst nach den Synchronisationsinformationen im Kanal. Bei NR ist diese Information im Synchronization Signal Block (SSB) enthalten. Dieser setzt sich aus den PRBs zusammen, die die Primary Synchronization Signals (PSS), die Secondary Synchronization Signals (SSS), sowie den Physical Broadcast Channel (PBCH) enthalten. Abb. 2.5 zeigt deren Lage in der Zeit- und Frequenzachse in einer 30 kHz Subcarrier Konfiguration. Ein SSB verwendet immer 20 PRBs auf der Frequenzachse und benötigt somit 7,2 MHz in dieser Konfiguration. Auf der Zeitachse belegen der PSS und SSS ein OFDM Symbol, während der PBCH über 3 OFDM Symbole gesendet wird. Insgesamt werden 4 Symbole für den SSB über 20 PRBs verwendet. Das Intervall des SSB ist ebenfalls konfigurierbar. In Band n78 mit 30 kHz Subcarriern und aktiviertem Beamforming werden üblicherweise zwei SSBs in jedem Downlink Slot übertragen, also ein SSB alle 2,5 ms.

Während auf der LTE Luftschnittstelle die Synchronisations- und Broadcastinformationen immer in der Mitte des Kanals übertragen werden, können diese auf der NR Luftschnittstelle an jeder beliebigen Position übertragen werden. Diese Flexibilität ist nötig, damit auf einem Kanal in Zukunft mehr als nur eine „Numerology“, also mehr als eine Konfiguration verwendet werden kann. So könnte z. B. eine Numerology auf einem Großteil des Kanals für den normalen Internet Zugang verwendet werden, während auf einem separaten Teil eine andere Numerology für langsame Machine Type Communication (MTC) Anwendungen verwendet wird und ein dritter Teil des Kanals für Ultra-Low Latency konfiguriert wird.

Ein weiterer deutlicher Unterschied zur LTE Luftschnittstelle ist die Verwendung von Referenzsignalen. Bei LTE werden individuelle Symbole in einer festgelegten Anordnung und einer festgelegten Sendeleistung über die gesamte Kanalbandbreite verwendet. Dies ermöglicht es einem Endgerät den Kanal zu vermessen und hilft somit,

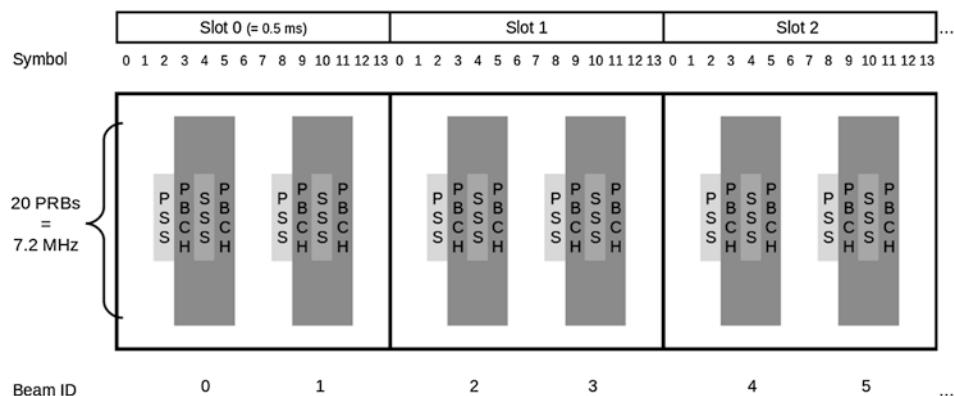


Abb. 2.5 Beispielkonfiguration von Synchronisierungs- und Broadcastinformationen

die Nutzdaten zu dekodieren. Bei NR können die Referenzsignale nicht mehr über den ganzen Kanal verteilt werden, da unterschiedliche Konfigurationen in unterschiedlichen Teilen und zu unterschiedlichen Zeiten im Kanal verwendet werden können. Darum werden auf der NR Luftschnittstelle kanal- und nutzerspezifische Referenzsignale verwendet, die nur dann auf dem Kanal eingefügt werden, wenn Daten übertragen werden.

2.3.4 Massive-MIMO für Beamforming und Multi-User Datenübertragung

Eine wichtige Funktionalität der 5G NR Luftschnittstelle im Vergleich zu früheren Technologien ist der Einsatz von ‚massive Multiple Input Multiple Output‘ (massive-MIMO) Übertragungen. Wie zuvor bereits im LTE Kapitel besprochen, nutzen MIMO Übertragungen die Eigenschaft eines Radiokanals aus, dass ein Signal aufgrund von Reflexionen an Gegenständen den Empfänger aus unterschiedlichen Richtungen und zu leicht unterschiedlichen Zeiten erreicht. Statt nur einen Datenstrom zu übertragen und die unterschiedlichen Signalanteile, die zu unterschiedlichen Zeiten beim Empfänger ankommen, zu kombinieren, kann der Empfänger dies auch nutzen, um unterschiedliche Datenströme auf diese Weise zu empfangen. Dazu werden mehrere Antennen auf der Sende- und auf der Empfangsseite benötigt. Üblich ist heute bei LTE die 2×2 Datenübertragung, da hier ein kreuz-polarisiertes Antennenpaar (horizontal und vertikal) verwendet werden kann, das in einer physischen Antenne untergebracht ist. Auf diese Weise lässt sich eine Datenübertragung mit zwei Strömen sehr einfach realisieren. Mit aufwendigeren Antennen und bei optimalen Übertragungsbedingungen kann auch 4×4 MIMO verwendet werden, das 4 Antennen beim Sender und auch beim Empfänger benötigt. Auf diese Art kann theoretisch die Datenrate gegenüber einer Single-Stream Übertragung vervierfacht werden. Wie in Abb. 2.6 gezeigt, wird dieses Prinzip in 5G NR deutlich erweitert. Mit 32 oder 64 Sendeantennen können zwei Ziele erreicht werden:

Eine Möglichkeit ist die Nutzung der zahlreichen kleinen Antennen für Beamforming, also die Konzentration der Sendeleistung in eine bestimmte Richtung, um damit die Reichweite des Signals oder die Übertragungsgeschwindigkeit für Nutzer mit guten Empfangsbedingungen zu steigern. Hier ist es wichtig zu verstehen, dass Endgeräte nicht nur wie bisher eine Zelle sehen, sondern viele „Beams“ einer Zelle die in unterschiedliche Richtungen zeigen. Manche dieser Beams kann das Endgerät empfangen, andere hingegen nicht. Beams werden über die Synchronization Signal Block Beam IDs (SSB IDs) identifiziert, die über das Antennenarray mit 32 oder 64 Elementen in unterschiedliche Richtungen eines Sektors gesendet werden. In 5G müssen Endgeräte daher nicht nur die Cell IDs suchen, sondern zusätzlich auch noch nach den Beam IDs der gefundenen Zellen.

Eine weitere Möglichkeit das massive-MIMO Array zu verwenden ist, mehrere Beams gleichzeitig an mehrere Benutzer zu senden, die sich an unterschiedlichen Orten

Passive 2x2 MIMO Panel Antennen

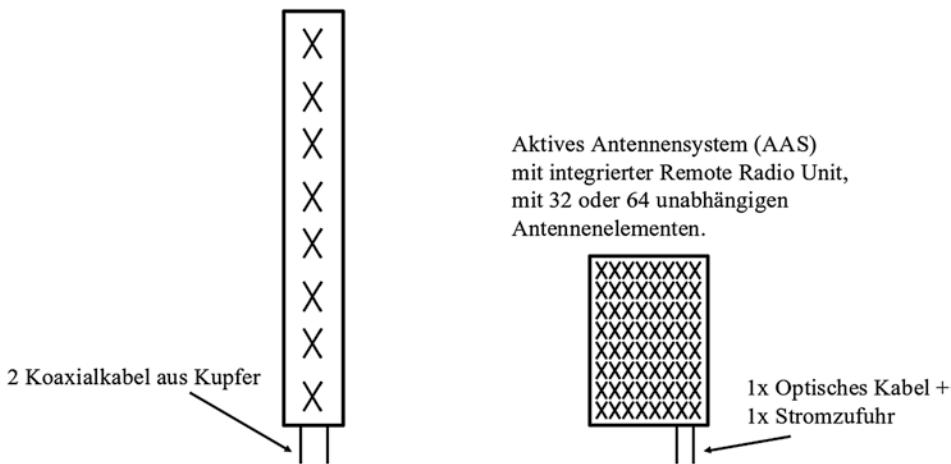


Abb. 2.6 Traditionelle 2×2 MIMO Antenne für LTE gegenüber einem 5G Active Antenna System (AAS)

in einer Zelle aufhalten. Dies wird als Multi-User MIMO bezeichnet. Die Sendeleistung müssen sich die Endgeräte jedoch teilen. Während dies nicht die Reichweite der Zelle steigert und auch nicht die maximale Datenrate eines einzelnen Endgerätes, ist es auf diese Weise aber möglich, die Kapazität einer Zelle zu steigern.

In der Praxis können beide Möglichkeiten gleichzeitig verwendet werden, wie folgendes Beispiel veranschaulichen soll: Mehrere Endgeräte sind mit einer 5G NR Zelle verbunden. Zwei Endgeräte befinden sich in einem Abstand vom Zellmittelpunkt, während vier andere Geräte sich nahe am Zentrum befinden und gleichmäßig über den 120 Grad Abdeckungsbereich des Sektors verteilt sind. Die zwei weiter entfernten Geräte werden jeweils über Beamforming erreicht, was deren Datendurchsatz steigert. Der Kanal kann somit wieder schneller für andere Geräte verwendet werden. Die vier anderen Geräte können gleichzeitig mit Daten versorgt werden, z. B. jeweils mit 2 MIMO Streams, zusammen also mit 8 MIMO Streams, von denen jeder in eine andere Richtung fokussiert wird. Die Basisstation errechnet dazu den Winkel für jede MIMO Übertragung, damit jedes Endgerät möglichst nur die Signalenergie seiner eigenen Datenströme empfängt. Die Übertragung von 8 unabhängigen Datenströmen erhöht die Zellkapazität im Vergleich zu Single-User MIMO Übertragungen mit jeweils nur einer 2×2 oder einer 4×4 Übertragung zu einer Zeit.

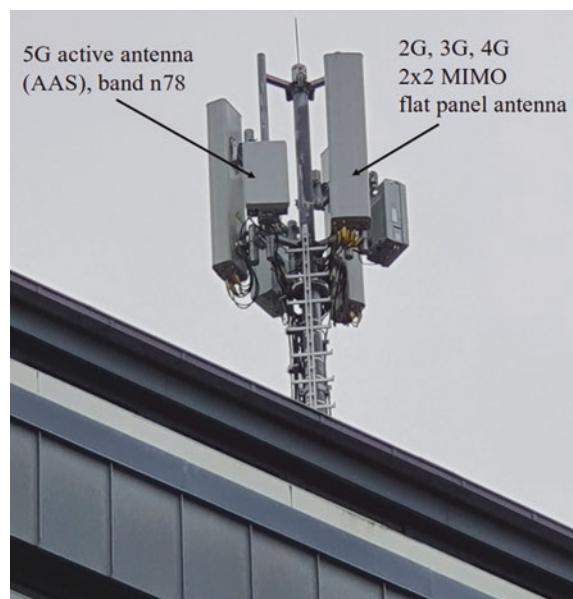
Manche Netzbetreiber verwenden auch passive 8×8 Panel Antennen, die etwas größer als traditionelle 2×2 oder 4×4 Panel LTE Antennen sind. Damit ist eine Grundform von massive MIMO Datenübertragung möglich, aber Reichweite und Anzahl der gleichzeitig angesprochenen Endgeräte sind geringer als mit Active Antenna Arrays

(AAS). Und schließlich ist es auch möglich, normale 2×2 oder 4×4 Pannel Antennen wie bei LTE ohne massive-MIMO für die 5G Luftschnittstelle zu verwenden.

Abb. 2.6 zeigt auch den ungefähren Größenunterschied einer klassischen Pannel Antenne und eines Active Antenna Arrays. Während Pannel Antennen üblicherweise dünn und etwa 2,5–3 m hoch sind, sind 5G NR Active Antenna Arrays für Band n78 deutlich kürzer, dafür aber deutlich breiter. In der Praxis werden 5G NR Array Antennen für Band n78 neben schon vorhandenen 2×2 oder 4×4 Antennen installiert und sind so einfach zu identifizieren. Dies ist in Abb. 2.7 gezeigt.

In der Praxis kann Beamforming auf zwei Weisen gesteuert werden. Wenn ein Endgerät eine Verbindung aufbaut, werden die Synchronization Signal Blocks (SSBs) verwendet, um den ersten Beam für die Verbindung auszuwählen. Jeder der 8 möglichen Beams einer FR1 Zelle hat dafür seinen eigenen SSB. Alle SSBs werden in unterschiedliche Richtungen mit der gleichen Signalstärke gesendet. Das bedeutet, dass ein Endgerät jeden SSB mit einer anderen Signalstärke empfängt. Bei Kontaktaufnahme mit dem eNB wählt es dann einen mit diesem SSB verknüpften Random Access Übertragungszeitpunkt. Sobald ein Radio Bearer aufgebaut ist, ist es dann Aufgabe des Netzwerkes, den Beam für das Endgerät nachzuregeln. Das Netzwerk kann den Beam für das Endgerät auch weiter einengen, um die Signalstärke weiter zu erhöhen. Dazu werden Channel State Information Reference Signale (CSI-RS) während der Datenübertragung gesendet. Das Endgerät wertet diese Übertragungen aus und sendet ein Feedback, wie diese empfangen wurden.

Abb. 2.7 Eine Dachinstallation eines klassischen 2×2 MIMO Multi-Band Antennensystems zusammen mit einem 5G Antennenarray



Beam Steering kann auf zwei Arten erfolgen: Eine Methode ist, das Endgerät über eine RRC-Reconfiguration Nachricht aufzufordern, Beam Reporting auf dem MAC Layer des Protokollstacks zu aktivieren. Somit kann das Netzwerk Informationen sammeln, wie das Pre-Coding in der Downlink Richtung auszusehen hat, um die Signalenergie zu bündeln. Das Feedback wird dazu auf dem MAC Layer vom Endgerät gesendet und enthält direkt die Information, welche Pre-Coding Matrix aus einem standardisierten Pre-Coding Matrix Codebook ausgewählt werden soll. Auf dem MAC Layer wird dazu der Pre-coding Matrix Indicator (PMI) gesendet. Zusammengefasst untersucht also das Endgerät das eingehende Signal und gibt dem gNB dann Rückmeldung, wie der Beam geändert werden soll.

Die zweite Option für das Netzwerk den Beam zu steuern ist, das Endgerät anzulegen, periodische Sounding Reference Signale (SRS) im Uplink zu senden. Der gNB weiß, wie die SRS Übertragungen aussehen sollten und vergleicht die eingehende Übertragung dann mit dem Optimalzustand. Über die Abweichung kann dann wiederum berechnet werden, wie der Beam anzupassen ist.

In beiden Fällen wird die RRC Signalisierung nur verwendet, die Beam Control Funktion zu aktivieren. Nach der Konfiguration wird das Feedback in Uplink Richtung und Konfigurationsänderungen für den Beam auf dem MAC Layer übertragen, und es findet keine weitere RRC Signialisierung statt. Auf diese Art kann eine Rückmeldung und eine Anpassung der Übertragung sehr schnell erfolgen. An dieser Stelle sei angemerkt, dass dieser Mechanismus große Ähnlichkeit mit der Kontrolle von LTE Carrier Aggregation hat. Wie im Kapitel zu LTE beschreiben, wird LTE Carrier Aggregation ebenfalls durch RRC Nachrichten initial konfiguriert, aber die Aktivierung und Deaktivierung der Secondary Cells findet dann ebenfalls auf dem MAC Layer statt.

Sollte das Endgerät den Beam während der Datenübertragung verlieren, startet es eine Beam Failure Recovery Prozedur. Dies geschieht über den Random Access Kanal und durch Auswahl einer neuen SSB ID.

Ergebnisse einer Beam Messung können auch in einem RRC Measurement Report gesendet werden. Dies dient jedoch nicht dem Beam Level Mobility Management, sondern wird für das Handover zwischen den Zellen verwendet (Inter-Cell Mobility). Es findet also ein Wechsel zwischen zwei Beams statt, die in unterschiedlichen Zellen ausgesendet werden. Im Unterschied zu LTE verwendet 5G NR also zwei Mobility Management Prozeduren: Beam Level Mobility wird vom MAC Layer kontrolliert, während Cell Level Mobility (Handover) über den RRC Layer gesteuert wird.

2.3.5 TDD Slot Formate

Auf einem TDD Kanal (Carrier) werden Downlink und Uplink Übertragungen auf der Zeitachse getrennt. Eine typische Konfiguration für Band n78 im 3,5 GHz Bereich nutzt 3 Slots mit je 14 Symbolen in Downlink Richtung (D), 1 Slot mit 14 Symbolen für den

Uplink (U) und einen „mixed special“ Slot (S) dazwischen. Diese Konfiguration ist in Abb. 2.8 gezeigt und wird als DDDSU Konfiguration bezeichnet.

Für das Netzwerk gibt es mehrere Möglichkeiten, wie es dem Endgerät die gewählte Slotkonfiguration signalisieren kann. Die einfachste Möglichkeit, die in der Praxis verwendet wird, ist, die UL-DL Konfiguration in einer RRC-Reconfiguration Nachricht beim Hinzufügen des NR Carrier zu einer bereits bestehenden LTE Verbindung zu senden. 3GPP TS 38.213 zeigt in Abschn. 11.1¹⁸, wie dieser im tdd-UL-DLConfigurationCommon Information Element (IE) beschrieben ist. Der folgende Ausschnitt zeigt die oben beschriebene Konfiguration:

```
tdd-UL-DL-ConfigurationCommon
{
    referenceSubcarrierSpacing 30 kHz,           [[1]]
    pattern1
    {
        dl-UL-TransmissionPeriodicity 2.5 ms,   [[2]]
        nrofDownlinkSlots 3,                     [[3]]
        nrofDownlinkSymbols 10,                  [[4]]
        nrofUplinkSlots 1,                      [[5]]
        nrofUplinkSymbols 2                     [[6]]
    }
}
```

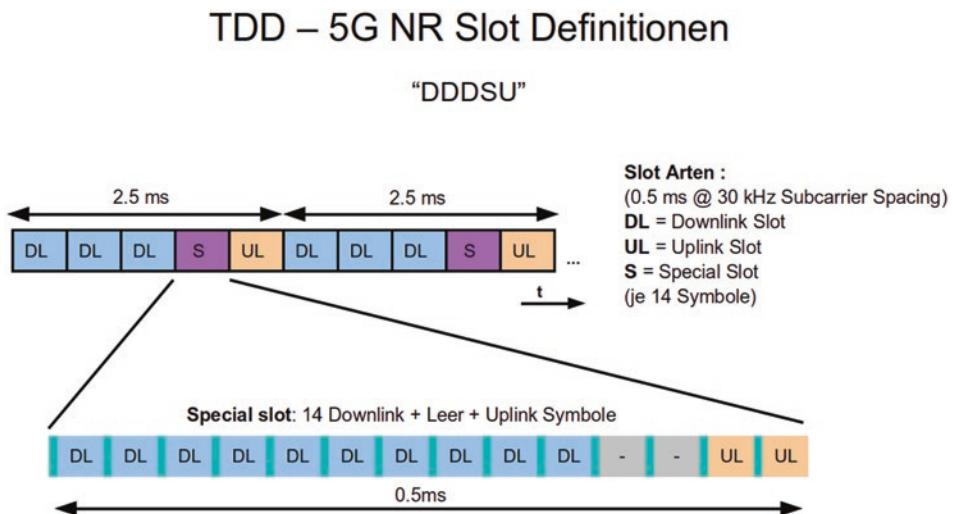


Abb. 2.8 Typische Konfiguration der NR Luftschnittstelle auf Band n78

Durch die Verwendung von 30 kHz Subcarrieren hat jeder Slot eine Länge von 0,5 ms und eine Übertragungseinheit enthält somit 5 Slots mit jeweils 14 Symbolen. Von diesen 5 Slots werden 3 für den Downlink und 1 Slot für den Uplink verwendet. Im Special Slot werden einige Symbole für den Downlink verwendet, manche bleiben frei, und die restlichen werden für den Uplink verwendet. Dies ist notwendig, um dem Radiomodul die Möglichkeit zu geben, zwischen Senden und Empfangen umzuschalten und um eine Überlappung von Senden und Empfangen aufgrund von Übertragungslaufzeiten zu vermeiden. Im oben gezeigten Beispiel werden im Special Slot 10 Symbole für den Downlink verwendet, 2 Symbole für den Uplink und 2 Symbole bleiben frei. Es werden also die meisten Symbole für die Downlinkübertragung verwendet. Somit werden also fast 4 Slots für den Downlink verwendet, während für den Uplink nur wenig mehr als 1 Slot verwendet wird. Die Uplink/Downlink Ratio ist deshalb etwa 3,8:1.

Die 3GPP Spezifikationen beschreiben etwa 60 solcher „Slot Konfigurationen“. Dies kann für Verwirrung sorgen, da die Tabelle nicht in Beziehung zur Uplink/Downlink Ratio gesetzt werden kann und die Konfigurationsnummer des Slots nicht an das Endgerät gesendet wird.

Eine weitere Option, die Uplink/Downlink Konfiguration an das Endgerät zu senden, ist über individuelle Downlink Control Information (DCI) Scheduling Nachrichten. Diese erlauben es, für einzelne Endgeräte spezielle UL/DL Konfigurationen einzustellen. In der Praxis wird diese Methode jedoch bisher nicht verwendet.

Ein weiteres Merkmal von TDD Systemen ist, dass Netzbetreiber, die ein zusammenhängendes Spektrum verwenden, ihre Radionetzwerke synchronisieren müssen, damit zwischen den Netzen keine Interferenz auftritt. Heute verwenden die meisten Netzwerke in Europa Frequency Division Duplexing (FDD). Hier ist eine Synchronisation zwischen den Netzen nicht nötig, da für Uplink und Downlink unterschiedliche Frequenzbereiche verwendet werden. Somit können Uplink Übertragungen in einem Netzwerk die Downlinkübertragungen in einem anderen Netzwerk nicht stören. In TDD Netzwerken werden Downlink und Uplink jedoch im gleichen Kanal übertragen und durch Timeslots voneinander getrennt. Sind zwei nebeneinanderliegende Netzwerke nicht synchronisiert, kann es somit zu Störungen kommen, da es dann vorkommen kann, dass ein Endgerät zu einer Zeit sendet, während ein benachbartes Endgerät, das über das andere Netzwerk kommuniziert, gerade im Empfangsmodus ist. Betreiber müssen daher ihre Radionetzwerke mit einer Toleranz von nur $\pm 1,5 \mu\text{s}$ synchronisieren⁹.

Die Synchronisation von Uplink- und Downlinkübertragungen zwischen benachbarten TDD Netzwerken bedeutet auch, dass alle Netzbetreiber die gleiche Downlink/Uplink Konfiguration verwenden müssen. Somit kann ein Netzbetreiber nicht mehr alleine entscheiden, welche Konfiguration verwendet wird.

2.3.6 Downlink Control Kanäle

Wie auch LTE strukturiert 5G NR die Datenübertragung auf der Luftschnittstelle in logische Kanäle, die dann auf Transportkanäle abgebildet werden und von dort auf physische Kanäle, die dann in Physical Resource Blocks (PRBs) im Resource Grid des Kanals eingebettet werden. Die wichtigsten Kanäle werden in Abb. 2.9 gezeigt und haben folgende Aufgaben:

Der Downlinkkanal, dem die meisten PRBs zugeordnet werden, ist der Physical Downlink Shared Channel (PDSCH). Dieser transportiert die Nutzdaten aller mit der Zelle verbundenen Endgeräte. Nutzdaten eines Endgeräts werden auf der logischen Schicht in einem Dedicated Traffic Channel (DTCH) übertragen. Zusätzlich zu den DTCHs überträgt der PDSCH auch alle Dedicated- und alle Broadcast Control Informationen. Dedicated Control Informationen sind notwendig, um einem Endgerät Downlink und Uplink Ressourcen zuzuweisen.

Teil der Broadcast Control Informationen sind die Konfigurationsinformationen der lokalen Zelle, sowie die der Nachbarzellen und werden in den 5G System Information Nachrichten zusammengefasst. Im 5G Non-Standalone Modus werden jedoch die meisten 5G System Information Nachrichten nicht benötigt, da den Endgeräten alle 5G spezifischen Parameter über die Signalisierungskanäle auf der LTE Seite mitgeteilt werden.

Weitere Broadcast Informationen für die erste Zellsuche nach Einschalten des Endgerätes werden im Master Information Block (MIB) gesendet, der Teil des Physical Broadcast Channels (PBCH) ist, der an bekannten Stellen des Resource Grids übertragen wird. Der PBCH enthält auch die Demodulation Reference Signals (DMRS), also Symbole mit einem fest definierten Inhalt. Mithilfe des DMRS kann ein Endgerät dann die Kanalqualität bestimmen. Wie in Abb. 2.5 gezeigt, enthält das Resource

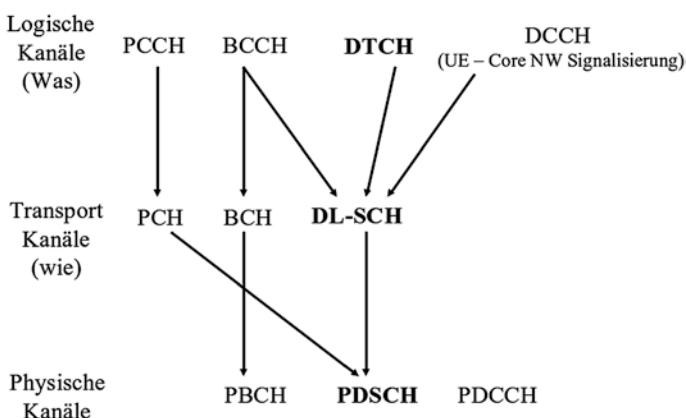


Abb. 2.9 5G NR Air-Interface Downlink Kanäle

Grid zusätzlich noch die Primary Synchronisation Signals (PSS), sowie die Secondary Synchronization Signals (SSS).

Wie jedes Mobilfunknetzwerk muss auch 5G NR alle Endgeräte im Idle Zustand ohne aktiven Radiokanal informieren, wenn neue Daten aus dem Internet für sie eingehen. Diese Prozedur wird Paging genannt, für die es einen separaten Logical- und Transportkanal auf der Luftschnittstelle gibt. Der dazugehörige Transportkanal ist auf PRBs abgebildet, die Teil des PDSCH sind. Im 5G NR Non-Standalone Mode wird der Paging Kanal jedoch nicht verwendet, da hier das Paging die LTE Ankerzelle übernimmt.

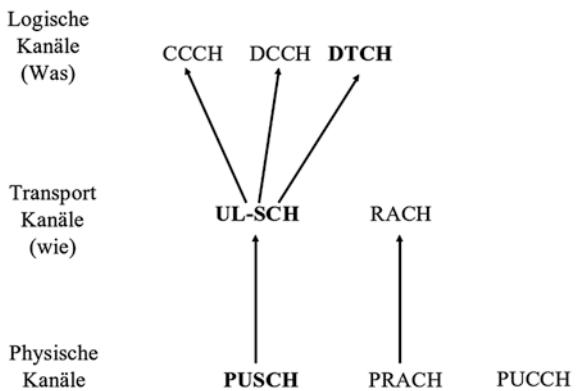
2.3.7 Uplink Kanäle

Im Uplink nimmt der Physical Uplink Shared Channel (PUSCH) den meisten Raum ein (Abb. 2.10). Wie im Downlink werden auf dem PUSCH Nutzdatenpakete und Kontrollinformationen von aktiven Geräten übertragen, wie z. B. Rückmeldungen über den korrekten Empfang von Downlinkdaten (Hybrid Automatic Repeat Request, HARQ).

Zusätzlich sind eine Anzahl von Resource Blocks im Uplink für den Random Access (RA) Zugriff reserviert. Diese werden benötigt, damit Endgeräte im RRC Idle State einen Radiokanal aufbauen können. Und schließlich gibt es den Physical Uplink Control Channel (PUCCH), der für HARQ Feedback verwendet wird, für Scheduling Request (SR) Nachrichten, sowie für Channel State Information (CSI) Feedback von Endgeräten, die bisher keine Ressourcen für diesen Zweck auf dem PUSCH zugeteilt bekommen haben.

Und schließlich werden eine Anzahl von PRBs in der Uplink Richtung für endgerätespezifische Sounding Reference Signals (SRS) reserviert, die der gNB verwendet, um die Signalqualität einzelner Endgeräte besser bewerten zu können.

Abb. 2.10 5G NR Uplink
Kanäle



2.3.8 Bandwidth Parts

Ursprünglich wurde die LTE Luftschnittstelle nur für den schnellen Internet Zugang optimiert. Dies ermöglichte es, alle wichtigen Parameter der Luftschnittstelle statisch zu definieren. Die 5G NR Luftschnittstelle jedoch wurde mit dem Ziel standardisiert, möglichst viele Anwendungen mit unterschiedlichen Anforderungen parallel zu betreiben. Um Anwendungen nur einen Teil der Bandbreite zuweisen zu können, wurde deshalb in 5G NR das Konzept der Bandwidth Parts (BWP) eingeführt.

In LTE beträgt die maximale Kanalbandbreite 20 MHz und alle Endgeräte müssen diese auch unterstützen. Die meisten Endgeräte unterstützen heute sogar eine deutlich größere Bandbreite und das Netzwerk kombiniert dann mehrere Kanäle per Carrier Aggregation (CA). Während dies für einen schnellen Internetzugang ideal ist, wird für die Dekodierung eines 20 MHz breiten Kanals relativ viel Energie benötigt und ist deswegen nicht für Geräte geeignet, die nur wenig Daten übertragen und wenig Energie verbrauchen dürfen. Aus diesem Grund erweiterte 3GPP dann viele Jahre nach der ursprünglichen Einführung von LTE die Spezifikation für solche Geräte. Details dazu finden sich im LTE Kapitel zum Thema LTE NB-IoT (Narrowband Internet of Things) und CAT-M. Da die LTE Spezifikation ursprünglich Endgeräte, die nur einen Teil des Kanals verwenden, nicht vorsah, mussten in der NB-IoT und CAT-M Erweiterung viele Kompromisse eingegangen werden, um die Rückwärtskompatibilität zu gewährleisten.

Für 5G wurde beschlossen, die Konfiguration des Kanals flexibel zu halten. Schon in der ersten 5G Spezifikation (3GPP Release 15) wird beschrieben, dass ein Endgerät nicht die volle 100 MHz Bandbreite eines Kanals unterhalb von 6 GHz unterstützen muss, bzw. die vollen 400 MHz im mmWave Spektrum. Wenn sich ein 5G Endgerät mit dem Netzwerk verbindet, kann es in einer Bitmap dem Netzwerk mitteilen, welche Bandbreiten unterstützt werden. Dabei wurde es leider versäumt, die Unterstützung von 90 MHz Bandbreite in die Bitmap aufzunehmen. Dies musste später in einem separaten Information Element dem Standard hinzugefügt werden. Notwendig wurde dies, da z. B. in Deutschland manche Netzbetreibern „nur“ 90 MHz Spektrum für Band n78 erwerben konnten.

Alle 5G NR Endgeräte wie Smartphones, Tablets und Router, die heute auf dem Markt sind, unterstützen eine Carrier Bandbreite von 100 MHz. Die 3GPP Spezifikation erlaubt bis zu vier gleichzeitige BWPs. Wird nur ein BWP verwendet, teilt das Netzwerk einem Endgerät bei Verbindungsaufnahme nur ein BWP zu, das die ganze Kanalbandbreite abdeckt. Manche Netze teilen einem Endgerät bei Verbindungsaufnahme jedoch zunächst nur ein „initial BWP“ zu, das nur einen Teil des Carriers abdeckt. Danach wird ein weiteres BWP konfiguriert, das die ganze Kanalbreite umfasst. Somit kann in Zeiten, in denen nur wenige Daten übertragen werden, das ‚kleinere‘ BWP verwendet werden und somit der Stromverbrauch des Endgerätes gesenkt werden.

Für das Netzwerk gibt es mehrere Möglichkeiten, zwischen den BWPs zu wechseln. Üblich ist, bei der Resourcenzuweisung im Physical Downlink Control Channel

(PDCCH) anzugeben, welcher BWP verwendet werden soll. BWPs können auch unterschiedliche Bandbreiten im Uplink und Downlink haben, dies wird in der Praxis bisher jedoch nicht verwendet.

Der folgende Ausschnitt einer RRC-Connection Reconfiguration Nachricht zeigt, wie ein BWP während des Hinzufügens einer 5G NR Zelle zu einer LTE Verbindung signalisiert wird:

```
SCS-specificCarrier
  offsetToCarrier: 0
  subcarrierSpacing: 30 kHz
  carrierBandwidth: 217
  [...]
initialDownlinkBWP
  genericParameters
  locationAndBandwidth: 16499
  subcarrierSpacing: 30 kHz
  [...]
```

Während das Subcarrier Spacing Information Element einfach zu interpretieren ist, ist dies beim Location and Bandwidth Information Element deutlich schwieriger. Wie der Name andeutet, wird hier der erste Resource Block (RB) des BWP relativ zum unteren Ende des Carriers angegeben, sowie die Anzahl der RBs auf der Frequenzachse. Eine Formel und eine Tabelle ist nötig, um aus diesem Wert die entsprechenden zwei Parameter zu dekodieren¹⁰.

Im oben gezeigten Beispiel ist die Carrier Bandbreite 217 RBs. Bei einem Subcarrier Abstand von 30 kHz entspricht das einer Kanalbandbreite von $217 \text{ RBs} \times 12 \text{ Symbole} \times 0,03 \text{ MHz} = 78,12 \text{ MHz}$. Mit einem Schutzabstand (Guard Band) am Anfang und Ende des Carriers ergeben sich dann 80 MHz. Wird die oben referenzierte Formel verwendet und wird weiterhin angenommen, dass das BWP bei 0 beginnt und alle 217 RBs abdeckt, resultiert die Formel in einem „Location And Bandwidth“ Wert von 16.499. Für einen 90 MHz breiten Kanal mit 245 RBs ist dieser Wert 8799.

2.3.9 Der Downlink Control Channel und das Scheduling

Nachdem einem Endgerät ein BWP zugeteilt wurde, muss es den Kanal nach Informationen abhören, welche RBs Daten enthalten und welche RBs in Uplink Richtung verwendet werden können. In LTE werden dazu ein oder mehrere der 14 Symbole eines Subframes auf der Zeitachse und alle Symbole auf der Frequenzachse einmal pro Millisekunde für den Physical Dedicated Control Channel (PDCCH) verwendet.

In 5G NR wird hierfür ein erweiterter Mechanismus verwendet. Dies ist notwendig, da unterschiedliche Teile des Carriers für unterschiedliche Anwendungen verwendet werden können, die nicht unbedingt den gleichen Subcarrierabstand verwenden. Beispielsweise könnte in Zukunft zwar ein Großteil des Kanals für schnellen Internetzugang verwendet werden, während ein kleiner Teil jedoch für Machine-Type Communication (MTC) verwendet wird. Diese Anwendungsart benötigt nur wenig Bandbreite, fordert jedoch schnelle Reaktionszeiten. Um dies zu gewährleisten, könnte ein Netzbetreiber in diesem Teil des Carriers einen anderen Subcarrier Abstand verwenden. Daraus folgt, dass es nicht mehr möglich ist, einen einzigen PDCCH zu verwenden und für diesen die gesamte Kanalbandbreite zu verwenden.

In 5G wird der PDCCH in sogenannten Control Regions ausgestrahlt, die nicht die ganze Bandbreite des Kanals verwenden müssen. Einem Endgerät können dann eine oder mehrere Control Regions zugewiesen werden, die innerhalb der für das Endgerät konfigurierten BWPs liegen müssen. Dies wird als Control Region Set (CORESET) bezeichnet. Mit anderen Worten stellt das CORESET eine Anzahl von Bereichen innerhalb der BWPs eines Gerätes dar, in denen der PDCCH zu finden ist.

In der Praxis ist bisher der schnelle Internetzugang die einzige Applikation in öffentlichen Netzwerken, und das Netzwerk teilt die komplette Kanalbandbreite über ein BWP den Endgeräten zu. Somit wird der PDCCH wie auch bei LTE über die gesamte Kanalbandbreite ausgesendet. Durch die beschriebene Flexibilität ist es jedoch in Zukunft möglich, dies ohne Modifikation der Spezifikationen zu ändern.

Das folgende Beispiel zeigt, wie ein CORESET in einer RRC Connection Reconfiguration Nachricht konfiguriert wird:

```
ControlResourceSet
  controlResourcesetId: 1
  frequencyDomainResources: ffffffff00 [1 bit represents 6 RBs]
  duration: 1 [symbol]
```

- 240 Resource Blocks werden auf der Frequenzachse verwendet.
- 1 OFDM Symbol wird auf der Zeitachse verwendet.
- Implizite Verwendung von QPSK für die Modulation, ein Symbol repräsentiert 2 Bits.

Da jeder Resource Block auf der Frequenzachse 12 Symbole enthält, können im CORESET dann $240 \times 12 \times 2$ Bits übertragen werden. Bei einer Länge des Slots von 0,5 ms ist die Datenrate für die Zuteilung von Downlink und Uplink Resourcen für alle Geräte somit $(4/5 \text{ (DDDSU)} \times 5760 \text{ Bits})/0,0005 \text{ s} = 9,2 \text{ Mbit/s}$.

Die Daten in einem CORESET sind wie folgt strukturiert: Ein Resource Block (12 Symbole auf der Frequenzachse) wird als Resource Element Group (REG) bezeichnet. 6 REGs werden dann zu einem Control Channel Element (CCE) zusammengefasst.

Ein CCE ist die kleinste Einheit für eine Scheduling Nachricht. Da drei Demodulation Reference Symbole alle 12 Symbole eingefügt werden, können in einer Scheduling Nachricht somit 6×9 (Symbole) $\times 2$ Bits (QPSK Modulation) verwendet werden. Dies entspricht einer Nachrichtenlänge von 108 Bits. Um die Redundanz zu erhöhen, können mehrere CCEs für eine Scheduling Nachricht zusammengefasst werden. Typischerweise werden 4, 8 oder 16 CCEs gebündelt. Dies kann je nach Signalqualität entsprechend angepasst werden.

Ein oder mehrere CCEs, die die Scheduling Nachrichten enthalten, werden als Downlink Control Information (DCI) Element bezeichnet. Ein DCI kann dabei unterschiedliche Formate annehmen. Downlink Scheduling Informationen sind im DCI Format 1_0 enthalten, Uplink Scheduling Informationen nutzen Format 0_1.

Eine Downlink DCI Nachricht enthält unter anderem folgende Parameter:

- Resourcezuteilung auf der Frequenzachse, also wie viele Symbole auf der Frequenzachse einem Nutzer für den Datenempfang zugeteilt werden. Details dazu folgen weiter unten.
- Resourcezuteilung auf der Zeitachse. Details werden ebenfalls weiter unten beschrieben.
- Modulation und Kodierungsverfahren für die Datenübertragung.
- Downlink Assignment Index (welcher BWP verwendet werden soll).
- Ob neue Daten übertragen werden oder die Zuteilung einer anderen Redundanzversion bereits zuvor übertragenen Daten enthält (Hybrid Automatic Repeat Request, HARQ).
- Die HARQ Queue Nummer, zu der die Daten gehören. Bis zu 8 ACK/NACK HARQ Queues können auf der 5G Luftschnittstelle verwendet werden.
- Transmit Power Control (TPC) für das Endgerät.
- Eine Information, wo Bestätigungen (ACKs) für die empfangenen Daten gesendet werden sollen (PUCCH Resource Indicator).
- Ein HARQ Feedback Indikator, der das Endgerät anweist, wann es die Bestätigung der eingegangenen Daten senden soll.

Ein Uplink DCI enthält unter anderem die folgenden Parameter:

- Resourcezuteilung auf der Frequenzachse, also wie viele Symbole über die Frequenz einem Nutzer für den Datenempfang zugeteilt werden. Details dazu folgen weiter unten.
- Resourcezuteilung auf der Zeitachse. Details werden ebenfalls weiter unten beschrieben.
- Modulations- und Kodierungsverfahren.
- HARQ Informationen (siehe oben).

Im Uplink werden Ressourcen auf der Frequenzachse ähnlich wie in LTE zugeteilt. Eine Option ist die Verwendung einer Bitmap. Diese beschreibt, welche Resource Blocks (RBs) im Frequenzbereich einem Gerät zugeteilt werden. Die zweite Option, ist die Startposition und die Anzahl der Resource Blocks auf der Frequenzachse zu beschreiben, die einem Endgerät zugeteilt werden.

In LTE gibt es kein Scheduling im Zeitbereich, da es genau zwei RBs in einem Subframe gibt. Wenn ein Endgerät eine Zuweisung auf der Frequenzachse bekommt, werden auf der Zeitachse auch beide RBs zugeordnet. In 5G NR wird dies nun flexibler gestaltet. Eine Start- und eine Längeninformation beschreiben hier, wie Downlink- und Uplinkzuweisungen in einem Slot konfiguriert sind und wie viele Symbole verwendet werden. Zuweisungen dürfen jedoch Slotgrenzen nicht überschreiten, es können also maximal 14 aufeinander folgende Symbole zugeteilt werden. Mit einem im 3,5 GHz Band typischen Subcarrierabstand von 30 kHz kann eine Zuweisung für bis zu 0,5 ms erfolgen. Es ist jedoch möglich, einem Endgerät Ressourcen nicht nur in einem Slot, sondern auch in den nachfolgenden Slots mit einer Konfiguration zuweisen.

Ein weiterer Unterschied zu LTE ist das Resource Assignment Timing. In FDD LTE gelten Uplinkzuweisungen für den Slot, der 4 Subframes später übertragen wird. Mit leeren Puffern und der benötigten Rechenzeit (Processing Overhead) beträgt die Round Trip-Zeit auf dem LTE Interface somit zwischen 8 und 10 ms. Auf der TDD NR Luftschnittstelle hingegen gelten Uplink Zuweisungen für den darauf folgenden Uplink Slot. Somit verringert sich die Verzögerung auf der Luftschnittstelle. Zusätzlich dazu kann ein Scheduling alle 0,5 statt 1 ms wie bei LTE erfolgen.

Diese Vorteile können jedoch nur genutzt werden, wenn Daten auf einem 5G NR Träger mit 30 kHz Subcarrier übertragen werden. In Netzwerken, die 5G NR EN-DC mit Split Downlink Bearer verwenden, werden jedoch Daten über die LTE- und die NR Luftschnittstelle gleichzeitig übertragen.

Ein weiterer Unterschied der 5G NR TDD Luftschnittstelle zu FDD LTE ist der geänderte HARQ Mechanismus. Statt synchron wird dieser nun asynchron verwendet. Es gibt somit keine fixe Zeit mehr zwischen Downlinkübertragung und der Bestätigung (Acknowledgment) im Uplink. Stattdessen enthält die DCI Nachricht Informationen, wann und wo in Downlink Richtung übertragen wird und wo die Empfangsbestätigung für einen Datenblock dann gesendet werden soll.

2.3.10 Downlink Geschwindigkeit in Theorie und Praxis

Durch die maximale Kanalbandbreite von 100 MHz in Band n78 kann mit 5G NR ein großes Stück zusätzliches Spektrum für den Mobilfunk erschlossen werden. Dieses Unterkapitel gibt nun einen Überblick wie die theoretisch maximale Geschwindigkeit erreicht werden kann. In der Praxis ist jedoch nicht die maximale Geschwindigkeit eines einzelnen Nutzers wichtig, sondern die signifikante Kapazitätssteigerung, die mit 5G NR in Band n78 für alle Teilnehmer ermöglicht wird.

Die theoretisch maximale Datenrate eines Kanals in Band n78 kann wie folgt berechnet werden:

- 5G NR n78 Kanalbandbreite: 100 MHz.
- Subcarrierabstand: 30 kHz. Somit werden 28 Symbole pro Millisekunde gesendet.
- Anzahl an PRBs auf der Frequenzachse: 273. Dies ist die maximale Anzahl für einen Kanal mit 100 MHz Bandbreite.
- Modulation: 256-QAM, es werden also 8 Bits in jedem Symbol kodiert.
- Anzahl der gleichzeitigen Übertragungswege zum Endgerät: 4×4 MIMO.
- Benötigte Kapazität für Control Channel und Signalisierung: 15 %.
- Uplink/Downlink Pattern: DDDSU mit einer Special (S) Slot Konfiguration wie oben beschrieben. Das Verhältnis des Downlinks zum gesamten Kanal ist somit 3,8:5.

Mit diesen Werten kann die maximale Datenrate in einem 100 MHz breiten Kanal wie folgt berechnet werden:

$$\text{Maximale Datenrate} = 273 \text{ (PRBs)} \times 12 \text{ (Subcarrier)} \times 28 \text{ Symbole} \times 8 \text{ (256 QAM)} \times 1000 \text{ (Millisekunden)} \times 4 \text{ (MIMO)} \times (3,8/5) \times 0,85 \text{ (15 \% Overhead)} = \mathbf{1,896 \text{ Gbit/s.}}$$

In der Praxis konnte vom Autor auf einem solchen Kanal unter sehr guten Übertragungsbedingungen bisher eine Geschwindigkeit von etwa 1,3 Gbit/s erreicht werden.

Die hier berechnete maximal mögliche Datenrate ist jedoch nur der 5G Teil eines LTE/5G Split Bearers. Um die kombinierte Datenrate zu erhalten, muss noch der Durchsatz des LTE Teils der Verbindung addiert werden. In der Praxis können heute Netzwerke und Endgeräte drei bis vier LTE Kanäle in einer EN-DC Verbindung mit einer Bandbreite von 50 bis 60 MHz bündeln. Da der Datendurchsatz eines LTE Kanals pro Megahertz ungefähr identisch zu 5G NR ist, kommen theoretisch nochmals 900 Mbit/s hinzu. In der Praxis unterstützten Netze und auch Endgeräte 4×4 MIMO, jedoch nicht in auf allen Frequenzbändern. Somit kommen durch den LTE Anteil des Split Bearers 400 bis 500 Mbit/s hinzu und die gesamte Datenrate liegt bei etwa bei 1,7 Gbit/s. Da Netzwerke jedoch nur selten gering ausgelastet sind, ist es zwar möglich, aber üblicherweise schwierig, eine solche Datenrate in der Praxis zu reproduzieren.

Tab. 2.4 zeigt eine LTE/5G Split Bearer Konfiguration eines typischen europäischen Netzbetreibers, dem 100 MHz Bandbreite auf Band n78 zur Verfügung stehen.

Die gerade dargestellten Berechnungen und die in der Praxis ermittelten Werte basieren auf einem einzelnen Endgerät in einer Zelle mit wenig anderem Datenverkehr. Nicht berücksichtigt sind somit weitere positive Effekte die z. B. die Multi-User MIMO Übertragung über aktive Antennenarrays.

Tab. 2.4 Maximale Datenrate eines einzelnen Endgeräts mit einem LTE/5G Split Downlink Bearer

Bandnummer	Frequenzband	Bandbreite	Ungefähr max. Datenrate eines Nutzers in der Praxis
n78	3500 MHz	100 MHz	1000 Mbit/s
20	800 MHz	10 MHz	60 Mbit/s
3	1800 MHz	20 MHz	200 Mbit/s
1	2100 MHz	10 MHz	100 Mbit/s
7	2600 MHz	20 MHz	200 Mbit/s
Total		160 MHz	1560 Mbit/s

2.3.11 Downlink Datendurchsatz

In der Uplinkrichtung ist die Datenübertragungsrate eines Endgeräts wesentlich geringer als im Downlink. Dies hat vor allem folgende Gründe:

- Die maximale Sendeleistung eines Endgeräts beträgt nur 0,2 W (23 dBm).
- MIMO wird im Uplink wegen der geringen Sendeleistung nicht verwendet.
- Im Uplink wird vorwiegend das 16-QAM Modulationsverfahren verwendet. 64-QAM und in seltenen Fällen 256-QAM sind unter idealen Bedingungen möglich.
- Ein einzelner Sender im Endgerät kann nur ein Signal für ein Band zu einer Zeit erzeugen. Dies limitiert die Verwendung von Carrier Aggregation in der Uplink Richtung. In der Praxis unterstützen deshalb auch teure Endgeräte nur 2-CA im Uplink in einem limitierten Subset von Bändern, während in der Downlink Richtung bis zu 5-CA mit vielen Bandkombinationen möglich ist.
- 5G Endgeräte benötigen zwei Sender, einen für die LTE Seite und einen für den 5G Teil des Modems.
- Durch die Verwendung von TDD in Band n78 und einem DL/UL Verhältnis von etwa 4:1 kann eine maximale Datenrate in einem 100 MHz 5G Carrier in Uplink Richtung mit 16-QAM Modulation ohne MIMO von etwa 90 Mbit/s erreicht werden.

Während die erreichbare Datenrate im Uplink somit eine Größenordnung niedriger als im Downlink ist, sollte an dieser Stelle auch erwähnt werden, dass auch das Verhältnis von tatsächlich übertragenen Daten zwischen Downlink und Uplink auch in stark genutzten Zellen üblicherweise bei 10:1 liegt. Typischerweise ist bei hoher Last deshalb der Downlink pro Nutzer sehr langsam, während der Uplink noch weiterhin mit fast voller Geschwindigkeit zur Verfügung steht.

2.3.12 Das TDD Air Interface in den mmWave Bändern (FR2)

In manchen Regionen der Welt, wie z. B. in den USA, begannen Netzbetreiber schon früh mit der Verwendung von Spektrum im 24 GHz Bereich für ihre 5G Netzwerke. Da hier die Ausbreitung eines Radiosignals anderen Effekten als in niedrigeren Frequenzbereichen unterworfen ist, entschloss sich 3GPP, zwei Physical Layer Spezifikation zu entwerfen. Frequenzbänder unterhalb von 6 GHz werden als Frequency Range 1 (FR1) bezeichnet und können dort im FDD oder TDD Modus betrieben werden. Frequenzbänder ab 24 GHz, die mit 5G zum ersten Mal für den Mobilfunk verwendet werden, werden dem Frequency Range 2 (FR2) zugeordnet. In den Medien wird FR2 auch als Millimeterwellenspektrum (mmWave Spectrum) bezeichnet, da die Wellenlänge eines Signals oberhalb von 30 GHz kleiner als 10 mm ist. Tab. 2.5 gibt eine Übersicht der Bandnummern, die in 3GPP für die Nutzung in unterschiedlichen Regionen der Welt spezifiziert wurden.

In FR1 beträgt die Bandbreite eines gesamten Frequenzbandes typischerweise weniger als 100 MHz. Ausnahme ist Band n78 im 3,5 GHz Bereich. Hier steht in vielen Ländern bis zu 500 MHz Bandbreite zur Verfügung, von der 300 bis 400 MHz an Mobilfunknetzbetreiber zugeteilt wurden. In FR2 umfasst ein Frequenzband, wie in Tab. 2.5 gezeigt, mehrere GHz an Spektrum.

Einzelne FR2 Kanäle sind ebenfalls wesentlich breiter als in FR1 Bereich. LTE Kanäle haben heute typischerweise eine Bandbreite von 10–20 MHz im Downlink und ebenso im Uplink. 5G NR Kanalbandbreiten, besonders in Band n78 bei 3,5 GHz in Europa und Asien sind typischerweise 80–100 MHz. Die kleinste definierte Kanalbandbreite im FR2 Bereich ist dagegen 50 MHz, maximal dürfen 400 MHz verwendet werden. Das ist um den Faktor 4 mehr als im FR1 Bereich.

Durch die deutlich breiteren Kanäle und die geringere Reichweite des Signals sind für FR2 Subcarrier Abstände von 120 und 240 kHz spezifiziert. Mit einem Subcarrier Abstand von 120 kHz enthält ein Subframe mit einer Dauer von 1 Millisekunde 8 Slots mit jeweils 14 Symbolen. Zum Vergleich: In FR1 Bereich in Band n78 finden in einem Subframe 2 Slots mit jeweils 14 Symbolen Platz.

FR2 wird bisher hauptsächlich in den USA verwendet, da hier ursprünglich kein zusätzliches Spektrum in FR1 zur Verfügung stand. In der Praxis erreichbare Datenraten

Tab. 2.5 FR2 Bänder

Band Nummer	Band Name: Region ¹¹	Bandumfang
n257	28 GHz (Südkorea, Japan)	26,50–29,50 GHz
n258	26 GHz (Europa, China)	24,25–27,50 GHz
n259	39 GHz	39,5–43,5 GHz
n260	39 GHz Untermenge von Band n259 (USA) ¹²	37,00–40,00 GHz
n261	28 GHz Untermenge von Band n257 (USA)	27,50–28,35 GHz

sind im Bereich von 1 bis 1,5 Gbit/s, also in etwa ähnlich der Datenrate, die in einem 100 MHz breiten FR1 Kanal im Band n78 erreicht werden kann.

Ein großer Nachteil des FR2 Spektrums ist jedoch die sehr geringe Reichweite des Signals. Da Frequenzen im FR2 Bereich schon durch Wände und Fenster stark gedämpft werden, eignen sich diese nicht für die Versorgung innerhalb eines Gebäudes von externen Basisstationen. Externe Mobilfunkstandorte können weiterhin nur einen sehr kleinen Bereich außerhalb von Gebäuden mit einem Signal versorgen. Daraus folgt, dass die Anzahl von FR2 Standorten im Vergleich zu FR1 Standorten wesentlich höher sein muss, um eine größere Fläche möglichst nahtlos abzudecken. Dies ist vor allem aus Kostengründen problematisch und stellt Betreiber auch vor große Herausforderungen bezüglich der Backhaul Planung, da FR2 Stationen per Glasfaser an das Netzwerk angeschlossen sein sollten. Somit eignen sich FR2 Bänder aktuell hauptsächlich für die Bereitstellung einer sehr hohen Kapazität in relativ kleinen öffentlichen Bereichen wie Fußballstadien, sowie Konzert- und Ausstellungshallen.

Endgeräte, die FR2 Bänder unterstützen, brauchen speziell für diese Frequenzbänder entwickelte Antennen, die entlang der Ränder eines Endgerätes verbaut werden¹³, sowie spezielle Sende- und Empfangsbausteine. Aktuell sind diese in Geräten, die in Europa und den meisten anderen Teilen der Welt verkauft werden, nicht enthalten.

2.4 Die 5G FDD Luftschnittstelle

3GPP Release 15 enthält auch eine Beschreibung der 5G FDD Luftschnittstelle. Diese wird hauptsächlich in Frequenzbändern unterhalb von 3 GHz verwendet, um zusätzlich zu vorhandenen LTE Kanälen weitere Kapazität bereitzustellen, oder LTE Kanäle nach und nach zu ersetzen. Hier werden die von LTE bekannten Bandnummern verwendet und ein „n“ vorangestellt. Band 3 für LTE im 1800 MHz Bereich wird bei 5G NR als Band n3 bezeichnet.

Ein Grund, 5G nicht nur im 3,5 GHz Band, sondern auch in niedrigeren Frequenzbereichen zu verwenden ist die geringe Reichweite von Band n78. Für eine landesweite Abdeckung und Empfang auf dem Land ist es notwendig, die 5G Luftschnittstelle auch auf weit niedrigeren Bändern zu nutzen. Tab. 2.6 gibt eine Übersicht über Bänder unterhalb von 3,5 GHz, die Netzbetreiber heute oder in naher Zukunft für 5G verwenden, bzw. verwenden werden.

Wird 5G NR in einem Frequenzband ausgerollt, das schon für andere Radiotechnologien verwendet wird, können Netzbetreiber üblicherweise die schon vorhandenen Antennen weiterverwenden. Dies reduziert die Kosten signifikant und beschleunigt den Rollout, da es dann oft möglich ist, nur die digitalen Basebandmodule am Fuß der Station zu erweitern oder zu ersetzen. Wird zusätzliches Spektrum verwendet, müssen Netzbetreiber zusätzlich zu den Hardware- und Softwaremodifikationen auch vom nationalen Regulierer eine Genehmigung für zusätzliche Sendeleistung einholen. Werden

Tab. 2.6 5G Frequenzbänder unterhalb von 3,5 GHz (aktuelle Beispiele)

Bandnummer	Frequenzbereich	Typische Kanalbandbreite	Nutzungsszenario	Region
n28	700 MHz	10 MHz	<ul style="list-style-type: none"> • Versorgung auf dem Land • Versorgung in Gebäuden 	Europa
n3	1800 MHz	20 MHz	<ul style="list-style-type: none"> • Erweiterte Abdeckung über die Reichweite von Band n78 hinaus • Teilweise für ländliche Abdeckung 	Europa, Asien
n1	2100 MHz	10–20 MHz	<ul style="list-style-type: none"> • Erweiterte Abdeckung über die Reichweite von Band n78 hinaus • Teilweise für ländliche Abdeckung • Ersetzte UMTS nach dessen Abschaltung 	Europa, Asien
n7	2600 MHz	20 MHz	<ul style="list-style-type: none"> • Übergang von LTE nach 5G • Alternative zu n78, falls dieses Band nicht verfügbar ist 	Europa, Asien
n71	600 MHz	10–20 MHz	<ul style="list-style-type: none"> • Versorgung auf dem Land 	USA, T-Mobile
n5	850 MHz	5–10 MHz	<ul style="list-style-type: none"> • Versorgung auf dem Land • Versorgung in Gebäuden 	USA, AT&T

bestehende Antennen verwendet, ist dann auch 5G üblicherweise auf eine Datenübertragung mit 2×2 MIMO begrenzt.

Eine große Limitation der Verwendung von 5G NR in niedrigeren Frequenzbändern ist die geringe Bandbreite im Vergleich zu den bis zu 100 MHz breiten Kanälen im 3,5 GHz Band. Besonders im Frequenzbereich zwischen 600 und 900 MHz sind Netzbetreiber typischerweise auf 10 MHz breite Kanäle limitiert. Im mittleren Frequenzbereich zwischen 1800 und 2600 MHz sind Kanalbandbreiten von 15 bis 20 MHz typisch. Somit ist also vor allem in den niedrigen Frequenzbereichen die Geschwindigkeit von 5G NR deutlich langsamer als in Band n78.

An dieser Stelle sei noch erwähnt, dass 5G NR einen kleinen Vorteil gegenüber LTE im sub-3 GHz Spektrum hat. Wie im LTE Kapitel gezeigt wurde, bringt LTE in einem 20 MHz breiten Kanal 100 LTE Resource Blocks unter. Durch bessere Filter an den Kanalgrenzen können auf der gleichen Kanalbandbreite auf der 5G NR Luftschnittstelle 106 Resource Blocks untergebracht werden, also 6 % mehr.

Von diesem Unterschied abgesehen ist auf sub-3 GHz Kanälen die Konfiguration der 5G NR Luftschnittstelle jedoch der von LTE sehr ähnlich. Auch hier wird ein 15 kHz Subcarrier Abstand verwendet, statt 30 kHz wie in Band n78.

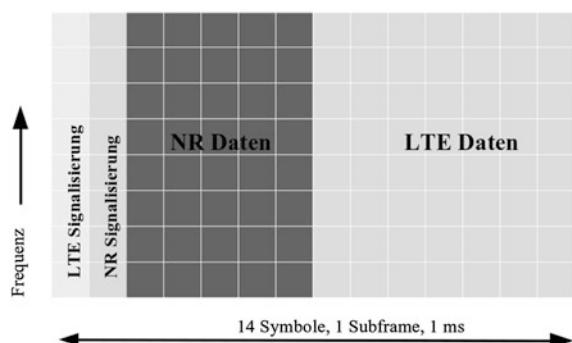
2.4.1 Refarming und Dynamic Spectrum Sharing

Hat ein Netzbetreiber bisher nicht verwendetes Spektrum für 5G FDD NR im Frequenzbereich unterhalb von 3 GHz, kann dies in der Praxis sehr schnell nutzbar gemacht werden. Oftmals möchten Netzbetreiber jedoch FDD NR in einem Teil des Spektrums verwenden, das noch für GSM, UMTS oder LTE verwendet wird. Der Prozess, diese älteren Technologien durch eine Neue zu ersetzen, wird als „Spectrum Refarming“ bezeichnet.

Im Falle von GSM und UMTS ist die Nutzung oft rückläufig, da der meiste Sprach- und Datenverkehr heute über LTE fließt. Die meisten Nutzer haben heute LTE-fähige Endgeräte, und Voice over LTE (VoLTE) ist in den meisten LTE Netzen verfügbar. In diesem Fall ist es recht einfach, das für GSM oder UMTS verwendete Spektrum zu reduzieren, oder eine dieser Technologien abzuschalten. Spektrum, das auf diese Weise freigeräumt werden kann, ist jedoch sehr limitiert. Deshalb möchten viele Netzbetreiber heute auch 5G in Spektrum verwenden, das bereits von LTE belegt ist. Dies ist jedoch nicht ohne weiteres möglich, da der komplette Wechsel von LTE nach 5G in sub-3 GHz Kanälen die Bandbreite für Teilnehmer mit LTE Endgeräten reduzieren würde.

Ein Lösung für dieses Problem ist das Dynamic Spectrum Sharing (DSS). Die Idee von DSS ist, LTE und NR gleichzeitig auf einem Kanal zu verwenden. Dies ist möglich, da die LTE und die NR Luftschnittstelle sehr ähnlich sind. Wie in Abb. 2.11 gezeigt,

Abb. 2.11 Dynamic Spectrum Sharing (DSS) zwischen LTE und NR



übertragen mit DSS sowohl LTE als auch NR ihre Signalisierungs- und Kontrollinformationen über den Kanal. Auf diesem Weg kann dann die Bandbreite dynamisch zwischen dem LTE Physical Downlink Shared Channel (PDSCH) und dem NR PDSCH geteilt werden. Wie die Verteilung geschieht, hängt von der Implementierung und Konfiguration ab. Wenn z. B. nur LTE Endgeräte in einer Zelle sind, kann die komplette Kapazität für den LTE PDSCH verwendet werden. Wenn LTE und 5G Endgeräte in der Zelle sind, kann ein Ansatz sein, die vorhandene Bandbreite gleichmäßig den zwei Technologien zuzuordnen. Es ist aber auch möglich, NR Endgeräte gegenüber LTE Endgeräten zu bevorzugen.

Ist 5G NR Standalone in einem Netzwerk noch nicht aktiviert, oder ein älteres 5G Endgerät unterstützt nur 5G NR Non-Standalone, wird der DSS Kanal wie folgt verwendet:

Im Ausgangszustand muss zunächst sichergestellt werden, dass ein Gerät nicht den LTE Teil des DSS Kanals im Idle-Zustand als aktive Zelle auswählt. Dies kann erreicht werden, in dem LTE Kanäle mit DSS eine niedrige Idle Mode „Camping“ Priorität vom Netzbetreiber zugeteilt bekommen. Somit wird also erreicht, dass nur LTE Kanäle ohne DSS vom Endgerät ohne aktive Radioverbindung ausgewählt werden. Wenn dann eine Radioverbindung aufgebaut wird, kann der 5G Teil eines DSS Kanals zur bereits bestehenden LTE Verbindung hinzugenommen werden. Da der DSS Kanal als 5G FDD Kanal zur Verbindung hinzugenommen wird, sucht das Endgerät dort keine LTE Synchronisationsinformationen und sieht somit nur den 5G Teil des Kanals. Zusätzlich können dann noch weitere LTE Kanäle mit LTE Carrier Aggregation eingebunden werden.

Für Geräte, die nur LTE unterstützen, läuft die Verbindungsaufnahme mit dem Netzwerk wie folgt: Da für diese Geräte die gleichen Kanalprioritäten (Cell Reselection Criteria) gelten wie für 5G ENDC Geräte, verwenden auch diese den DSS Kanal nicht im Idle Zustand. Bauen sie dann eine Verbindung mit dem Netzwerk auf, wird LTE Carrier Aggregation aktiviert und der LTE Teil des DSS Kanals wird als LTE CA Component Carrier in die Verbindung aufgenommen. Falls das Netzwerk die Kanalkapazität gleichmäßig zwischen LTE- und 5G-FDD fähigen Geräten aufteilt, würden beide Geräte etwa die gleiche Übertragungsgeschwindigkeit erreichen.

Ein Nachteil auf einem Kanal LTE und auch NR Systeminformationen auszustrahlen ist natürlich die zusätzliche Nutzung von RBs für die Signialisierung in beiden Systemen. Dies benötigt zusätzlich etwa 15 % der vorhandenen Ressourcen. Das bedeutet, dass die Kapazität für Nutzdaten auf diesem Kanal reduziert ist.

Ein wichtiges technische Detail, das in Abb. 2.11 nicht gezeigt wird, sind die LTE Channel State Indication – Reference Signals (CSI-RS), die gleichmäßig über den ganzen Kanal verteilt sind. LTE Endgeräte benötigen die CSI-RS Symbole für die Synchronisation mit dem Netzwerk und um die Signalqualität des Kanals messen zu können. Dies ist jedoch ein Problem für die 5G NR Seite, da NR seine eigenen Synchronisations-, Broadcast- und Kontrollkanäle an fest definierten Stellen im

Resource Grid senden will. Diese kollidieren somit mit den LTE Referenzsignalen. In der Praxis werden deshalb folgende Methoden verwendet, um diese Kollisionen zu vermeiden:

In frühen Versionen der LTE Spezifikation wurden Multimedia Broadcast Single Frequency Network (MBSFN) Subframes definiert, die ursprünglich für die Übertragung von TV Kanälen gedacht waren. In der Praxis wurden MBSFN Subframes dafür aber nie verwendet. Eine Eigenheit von MBSFN Subframes ist es jedoch, die Übertragungen dort unabhängig von LTE zu gestalten. Somit werden in diesen Subframes nur die ersten zwei Symbole für LTE Downlink Information verwendet, wie z. B. der Physical Hybrid ARQ Indicator Channel (PHICH) und die LTE Kontrollkanäle. All weiteren Symbole werden für LTE nicht genutzt und es werden keine CSI-RS Symbole im Resource Grid eingefügt. LTE Endgeräte ignorieren solche Subframes. Da keine CSI-RS Symbole in diesen Subframes vorhanden sind, können diese nun bei DSS verwendet, den NR Teil zu übertragen.

LTE Endgeräte werden über die Verwendung von MBSFN Subframes im System Information Block 2 informiert. Das folgende Beispiel zeigt ein Ausschnitt aus einer solchen Nachricht:

```
MBSFN-SubframeConfig
    radioframeAllocationPeriod: 4
    radioframeAllocationOffset: 0
    subframeAllocation: Four Frames
    mbsfnPattern: 1100 0000 0000 1000 0000 0000
```

In diesem Beispiel wird ein MBSFN Subframe Muster konfiguriert, das sich alle 4 LTE Frames (40 ms) wiederholt. Der „mbsfnPattern“ Parameter enthält dazu eine Bitmap, die die Lage der MBSFN Subframes beschreibt. Die Bitmap hat eine Länge von 4 (Frames) \times 6 Subframes = 24 bit, da nur 6 von 10 Subframes in einem 10 ms Frame für MBSFN verwendet werden können. Das oben gezeigte Muster informiert das Endgerät also über 2 MBSFN Subframes im ersten Frame, wie in Abb. 2.12 gezeigt, und einen weiteren MBSFN Subframe im dritten Frame. Das bedeutet, dass in 40 Subframes, die in einer Periode von 40 ms übertragen werden, 3 Subframes leer bleiben und für die NR Seite des DSS Kanals verwendet werden können.

Die leeren Subframes werden nun verwendet, um verschiedene NR Kanäle wie den PSS, den SSS und den Broadcast Kanal mit dem Master Information Block zu übertragen. Zusätzlich wird das erste Symbol im MBSFN Bereich für NR Kontrollkanäle verwendet, die verbleibenden Symbole werden für den NR Physical Downlink Shared Channel verwendet. Weiterhin werden zwei Symbole auf der Zeitachse für NR Demodulation Reference Signals (DMRS) verwendet, die den gleichen Zweck wie CSI-RS Symbole bei LTE erfüllen.

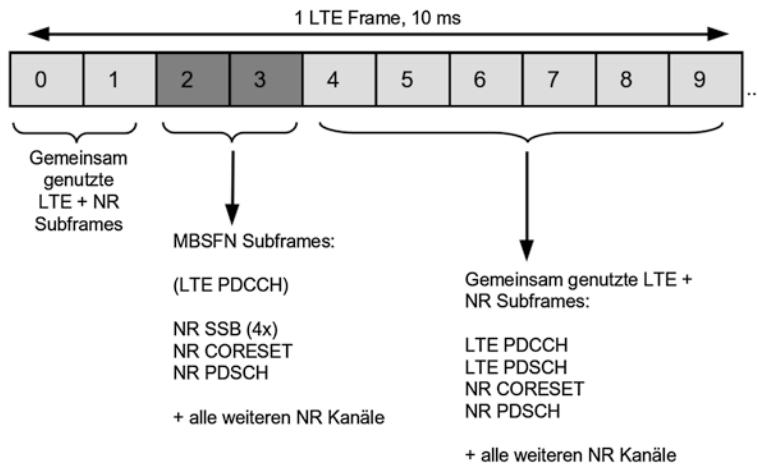


Abb. 2.12 Ein LTE Frame mit MBSFN Subframes

Würden nur MBSFN Subframes für NR verwendet, wäre die resultierende Datenrate natürlich sehr gering. Es ist jedoch möglich, normale Subframes für LTE und NR in Zeiten zu verwenden, in denen keine SSBs gesendet werden müssen. Da genügend SSBs in drei MBSFN Subframes alle 40 ms gesendet werden können, ist es somit möglich, NR und LTE flexibel zu mischen. Um eine Kollision zwischen LTE CSI-RS und NR DMRS Symbolen zu vermeiden, müssen Netzwerk und Endgeräte zusätzlich in der Lage sein, die DMRS Symbole in ein anderes Symbol auf der Zeitachse in jedem Subframe schieben zu können.

Das folgende Beispiel zeigt unterschiedliche Möglichkeiten, wie Refarming und DSS in der Praxis eingesetzt werden können und wie mit diesen Methoden ein Mobilfunknetzwerk über die Zeit weiterentwickelt werden kann:

Zunächst ist einen ländlicher Standort eines Netzbetreibers wie folgt ausgestattet:

- 10 MHz GSM in Band 8 (900 MHz)
- 10 MHz LTE in Band 20 (800 MHz)
- 20 MHz LTE in Band 3 (1800 MHz)
- 15 MHz UMTS in Band 1 (2100 MHz)

Basierend auf dieser Konfiguration entscheidet sich dann ein Netzbetreiber, zwei der drei UMTS Träger für 5G zu verwenden. Dynamic Spectrum Sharing würde somit nicht verwendet werden und der 5G FDD Kanal in Band 1 würde zu einer LTE Verbindung in der gleichen Art hinzugenommen werden, wie ein 5G TDD Kanal in Band n78. 5G-fähige Endgeräte könnten dann eine gesamte LTE Kanalbandbreite von 30 MHz in Band 3 und 20 verwenden und zusätzlich 10 MHz von Band n1 mit 5G. Diese würde die zur Verfügung stehende Bandbreite für 5G-fähige Geräte an diesem Standort von 30 MHz

auf 40 MHz erhöhen. 5G Endgeräte am Zellrand profitieren hiervon nicht direkt, da nur Band 20 die nötige Reichweite besitzt. Auch LTE Endgeräte würden von dieser Art des Refarmings nicht direkt profitieren, da sie die zusätzliche Kapazität in Band 1 nicht nutzen könnten.

Eine andere Möglichkeit für einen Netzbetreiber in dieser Situation wäre es, die vorhandenen 10 MHz Spektrum in Band 1 sowohl für LTE und auch für 5G NR Endgeräte durch Aktivierung von DSS für diesen Kanal zu verwenden. Beim Aufbau eines Kanals ergäbe sich dadurch kein Unterschied für 5G Geräte. LTE Endgeräte können jedoch die zusätzlichen 10 MHz Spektrum in Band 1 nun ebenfalls mit LTE Carrier Aggregation verwenden. Somit können dann mehr Kunden direkt von der zusätzlichen Kapazität profitieren.

Während das Hinzufügen eines 5G FDD Kanals zu einem LTE Anker große Ähnlichkeit mit der Aufnahme eines 5G TDD Kanals in Band n78 hat, gibt es jedoch einen großen Unterschied aus Kapazitätssicht. Wird ein 100 MHz n78 5G TDD Kanal zu einer LTE Verbindung hinzugefügt, dominiert die Bandbreite des 5G Kanals. In diesem Beispiel jedoch dominiert die LTE Seite, da hier durch Carrier Aggregation mehr Bandbreite zur Verfügung steht.

Da die 5G Flächenversorgung auf dem Land über Band n1 durch die begrenzte Reichweite nicht gewährleistet werden kann, ist es wichtig, 5G NR zusätzlich auch in einem niedrigen Frequenzbereich zu aktivieren. In Europa könnte dies zum Beispiel Band n28 im 700 MHz Bereich sein.

Wie auch für die NR TDD Luftschnittstelle, benötigt ein Endgerät auch für NR FDD zwei Sendeeinheiten. Ein Sender wird für den LTE Ankerkanal in Uplink Richtung benötigt, der Zweite für das Senden von Uplinkdaten über den NR Kanal. Erste NR FDD-fähige Endgeräte waren in der Kombination von LTE und NR Uplink sehr begrenzt und unterstützten keine Mid/Mid oder Low/Low Kombinationen. Das bedeutete, dass diese Endgeräte 5G auf Band n1 (2100 MHz) nicht mit einem LTE Anker auf Band 3 (1800 MHz) kombinieren konnten und deswegen auf Band 8 (900 MHz) oder Band 20 (800 MHz) als Anker ausweichen mussten. Dies war jedoch von großem Nachteil, da in diesen Bändern nur sehr wenig Kapazität zur Verfügung stand. Heute unterstützen jedoch alle neuen Endgeräte sehr viel mehr LTE Anker/NR Kombinationen und Mid/Mid und Low/Low Kombinationen werden jetzt unterstützt.

In einer weiteren Evolutionsstufe führen Netzbetreiber dann ein 5G Kernnetzwerk ein, und Endgeräte, die den 5G Standalone Modus unterstützen, können dann auch 5G NR Kanäle ohne einen LTE Anker verwenden. Nachteil in den ersten Jahren ist jedoch, dass die maximale Datenrate gegenüber einer LTE+5G NR Verbindung im Non-Standalone Mode geringer ist, da ja keine LTE Kanäle mit hinzugenommen werden können. Zwar sieht der Standard auch dies vor, in der Praxis wurde dies jedoch nicht umgesetzt. Hier hilft es also nur, weitere Kanäle auf 5G zu migrieren oder per DSS dort LTE und 5G gleichzeitig zu verwenden.

2.5 EN-DC Bearer und Scheduling

5G NR Non-Standalone Netzwerke nutzen LTE als Anker für die Signalisierung und die Übertragung von Nutzdaten, während der 5G NR Teil nur für die Übertragung von Nutzdaten verwendet wird. Der Aufbau einer Verbindung zwischen Endgerät und Netzwerk ist somit dem LTE Verbindungsaufbau sehr ähnlich und wurde nur leicht erweitert. Wie im Kapitel über LTE schon ausführlich beschrieben wurde, kann sich ein Endgerät in unterschiedlichen Zuständen befinden. Ist es komplett vom Netzwerk getrennt, wird dies oft auch als „Flight-Mode“ bezeichnet. Ist es am Netzwerk angemeldet und besitzt eine Radioverbindung, befindet es sich im RRC-Connected State (RRC=Radio Resource Control). Ist es angemeldet, es besteht jedoch keine Radioverbindung, befindet es sich im RRC-IDLE Zustand. Im IDLE Zustand behält es somit zwar seine IP Adresse, die Sende- und Empfangseinheit ist jedoch die meiste Zeit komplett abgeschaltet. Nur zu vordefinierten Zeiten wird kurz der Empfänger aktiviert, um auf dem Paging Kanal zu überprüfen, ob im Netzwerk neue Datenpakete oder Signalisierung warten. Empfängt ein Endgerät eine Paging Nachricht, verbindet es sich wieder mit dem Netzwerk und wartende Datenpakete werden dann an das Endgerät weitergereicht. Ist ein Endgerät im RRC-IDLE Zustand und es werden neue IP Pakete auf dem Endgerät erzeugt, wechselt es ebenfalls in den RRC-Connected Zustand und überträgt diese dann an das Netzwerk.

Der Wechsel von RRC-IDLE nach RRC-Connected wird auch als Aufbau eines Radio Access Bearers (RAB) bezeichnet. Oftmals wird auch nur die Abkürzung ‚Bearer‘ verwendet. Ein Radio Access Bearer ist eine logische Verbindung zwischen Endgerät und einem eNB. Im Betriebssystem des Endgeräts wird dieser durch eine IP Adresse und ein Netzwerkinterface repräsentiert. Typischerweise hat ein Endgerät mehrere Bearer gleichzeitig aufgebaut, einen für den Internetzugang und einen weiteren für den Voice over LTE Sprachdienst. Im Betriebssystem sind also zwei Netzwerkinterfaces sichtbar, jedes mit seiner eigenen IP Adresse. Normale Applikationen erhalten jedoch nur Zugriff auf das Interface, das die Verbindung zum Internet herstellt. IP Adresse und Netzwerkinterfaces bleiben auch im RRC-IDLE Zustand erhalten, Radio Access Bearer jedoch nicht. Diese werden üblicherweise schon nach ein paar Sekunden Inaktivität abgebaut.

Nachdem ein Radio Bearer aufgebaut ist, weist der eNB Scheduler Uplink und Downlink Ressourcen an jedes aktive Gerät zu. Im Downlink berücksichtigt das Scheduling die Priorität eines Gerätes, die Prioritäten der aufgebauten Bearer, den Umfang der Daten im Sendepuffer und andere proprietäre Parameter. In Uplinkrichtung melden alle Endgeräte dem Netzwerk, ob Daten für die Übertragung anstehen. Der eNB teilt dann den Endgeräten wiederum entsprechend ihrer Priorität Teile des Uplink Resource Grids zu bestimmten Zeiten zu.

Für LTE/NR EN-DC gibt es unterschiedliche Möglichkeiten, Daten zu transferieren:

- Nutzdaten, die über einen bestimmten Bearer übertragen werden, werden über die NR Seite der LTE/NR Verbindung übertragen.

- Nutzdaten eines bestimmten Bearers werden über die LTE Seite einer LTE/NR Verbindung übertragen. Dies ist z. B. für den VoLTE Bearer der Fall. Aus Sicht des IP Layers ist dies nicht unbedingt notwendig, da der IMS Dienst keine Informationen über das Radio Access Netzwerk und die Bearerkonfiguration hat. Die meisten Netzbetreiber verwenden jedoch für den VoLTE Bearer nur die LTE Seite einer LTE/NR Verbindung, um eine minimale und auch konstante Verzögerung zu gewährleisten. Zudem ist die Datenrate für einen Sprachanruf sehr gering, die höhere Datenrate der NR Seite bietet somit also keinen Vorteil. Außerdem wird für die meisten LTE/NR Verbindungen der LTE Teil der Verbindung auf einer niedrigeren Frequenz verwendet. Diese hat eine größere Reichweite und es ist somit auch nicht nötig, die Radioverbindung am Zellrand umzukonfigurieren.
- Nutzdaten eines Bearers werden über die LTE Seite und auch über die NR Seite einer Verbindung gleichzeitig übertragen. Dies wird als „Split Bearer“ bezeichnet und wird typischerweise für den Internetzugang verwendet.

2.5.1 Split Bearer und Flusskontrolle

Üblicherweise wird heute ein Split Bearer in einer EN-DC Verbindung im Downlink und auch im Uplink verwendet. Ein großer Unterschied zu LTE Carrier Aggregation ist jedoch, dass der LTE- und der NR Scheduler komplett unabhängig voneinander sind. Das bedeutet, dass eNB und gNB unabhängig voneinander entscheiden können, wann und wie viele Daten sie zum Endgerät senden wollen. In der EN-DC Option 3x ist der 5G gNB der Master eines Downlink Split Bearers und alle Pakete, die aus dem Internet ankommen, laufen zunächst bei ihm ein. Der gNB teilt dann die Datenpakete auf (Split), überträgt einen Teil davon über seine eigene 5G NR Luftschnittstelle und leitet einen anderen Teil zum eNB weiter, der diesen dann über die LTE Luftschnittstelle überträgt. Auf der LTE Seite aktiviert der eNB typischerweise Carrier Aggregation (CA), um die Daten über mehrere Frequenzbänder zu verteilen. Ein typischer gut ausgebauter Standort verwendet beispielsweise folgende Bänder und Technologien für einen Split Bearer:

- 5G NR Teil
 - Band n78 (3,5 GHz), 100 MHz
- 4G LTE Teil (mit Carrier Aggregation)
 - Band 3 (1,8 GHz), 20 MHz
 - Band 3 (1,8 GHz), 10 MHz
 - Band 7 (2,6 GHz), 20 MHz
 - Band 20 (800 MHz), 10 MHz

In diesem Beispiel werden 160 MHz Bandbreite für den Downlink Split Bearer verwendet. Durch die sich verändernden Radiobedingungen und Datenraten von allen Nutzern des eNB und gNB, ändert sich die Datenrate des Split Bearer ständig. Aus

diesem Grund informiert der LTE eNB den 5G NR gNB regelmäßig über seinen Downlink Buffer Status über das X2 Interface¹⁴.

Sequenznummern in den eingehenden Datenpaketen erlauben es dem Endgerät, die ursprüngliche Reihenfolge der Pakete wieder herzustellen, und die Datenpakete dann an die höheren Schichten des Protokollstapels weiterzureichen.

Auch im Uplink kann ein Split Bearer verwendet werden. Die dort erreichbare Datenrate ist jedoch deutlich geringer als im Downlink. Dies hat folgende Gründe:

Da ein Endgerät auf eine Sendeleistung von 0,2 W begrenzt ist, ist es schon bei LTE üblich, die Uplink Übertragung auf zwei Kanäle zu limitieren. Dies ist auch deshalb nötig, da ein Sender im Endgerät nur auf einem Kanal senden kann und deshalb für die Mehrkanalübertragung weitere Sendeeinheiten notwendig sind. Dies hat sich auch mit NR nicht geändert. Deshalb verwenden Geräte heute typischerweise nur einen Kanal auf der LTE Seite und einen Kanal auf der 5G Seite, und verzichten auf LTE Carrier Aggregation in Kombination mit 5G. Einzige Ausnahme ist, wenn die LTE Kanäle im gleichen Band und direkt nebeneinander liegen.

Ein anderer, oft übersehener Grund ist, dass auf dem High-Speed NR Kanal TDD verwendet wird und es somit ein Ungleichgewicht zwischen Uplink und Downlink gibt. In Deutschland sind TDD n78 Kanäle in einem 3,8:1 Downlink/Uplink Split konfiguriert. Außerdem wird MIMO nur in Downlink Richtung verwendet, was die Datenrate in der Uplink Richtung im Verhältnis zum Downlink weiter verringert. Somit ist im Uplink eines 100 MHz NR Kanals eine maximale Datenrate von etwa 50 Mbit/s möglich. Zusammen mit weiteren 50 Mbit/s eines 20 MHz LTE Kanals kann somit unter idealen Bedingungen eine Datenrate im Uplink von 100 Mbit/s erreicht werden. Dies ist etwas höher als die Datenrate, die mit LTE Carrier Aggregation in der Praxis heute erreicht werden kann.

2.5.2 Zwei Sender für EN-DC

Ein Nachteil eines LTE/NR EN-DC Bearers im Vergleich zu LTE Carrier Aggregation ist, dass zwei Sender im Endgerät benötigt werden. Diese müssen die maximale Sendeleistung von 0,2 W unter sich aufteilen. Dies ist auch dann nötig, wenn Daten nur auf der LTE Seite der EN-DC Verbindung im Uplink übertragen werden. Wie in Abb. 2.13 gezeigt, wird für einen EN-DC Split Bearer ein HARQ Feedback sowohl auf der LTE-, als auch auf der NR Seite benötigt, da die zwei Scheduler unabhängig voneinander arbeiten.

Während sich das Endgerät in der Nähe eines Sendestandortes befindet und nicht die komplette Leistung im Uplink genutzt wird (Power Headroom), ist die Nutzung von zwei Sendeeinheiten unproblematisch. Am Zellrand wird jedoch die limitierte Sendeleistung des Endgerätes zum Problem. Um dieses Problem zu reduzieren, kann jedoch das Netzwerk beim Erreichen der maximalen Sendeleistung des Endgeräts die HARQ Uplink Übertragungen so verteilen, dass zu einer Zeit nur jeweils ein Sender verwendet

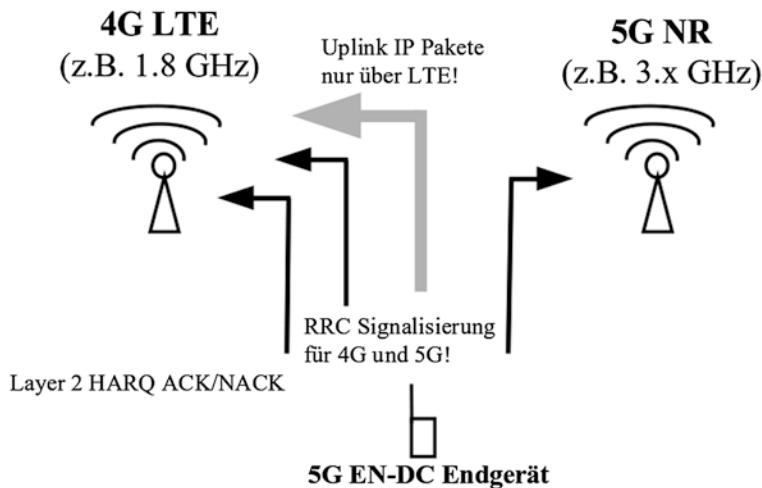


Abb. 2.13 Verwendung von zwei Sendern im Endgerät für EN-DC Downlink Split Betrieb und einem LTE-only Uplink Bearer

wird und dieser somit die ganze Sendeleistung verwenden kann. Dies reduziert natürlich die maximale Datenrate im Downlink, da das Endgerät nicht immer Daten senden kann, da auf das HARQ Feedback gewartet werden muss. Außerdem erfordert die HARQ Koordinierung eine sehr enge Zusammenarbeit des LTE- und des NR-Schedulers.

Während das HARQ Feedback auf der LTE- und auch auf der NR-Seite einer EN-DC Verbindung gesendet werden muss, wird die RRC-Signalisierung nur auf der LTE Seite übertragen. Wie die eigentlichen Nutzdaten (IP Pakete) im Uplink übertragen werden, kann jedoch sehr flexibel gehandhabt werden. Ist der Empfang gut, kann das Netzwerk ein Endgerät so konfigurieren, dass IP Pakete auf einem Split-Uplink Bearer übertragen werden. Wird der Empfang schlechter, kann das Netzwerk die Verbindung umkonfigurieren und alle IP Pakete werden dann nur auf der LTE Seite zum Netzwerk übertragen. Auf diese Weise steht dann mehr Sendeleistung für die LTE Sendeeinheit zur Verfügung. Außerdem verwendet typischerweise der LTE Teil einer EN-DC Verbindung ein niedrigeres Band als der 5G NR Teil mit Band n78. Hier ist die Reichweite somit deutlich besser. Weiterhin ist es möglich, bei guten Übertragungsbedingungen nur den NR Uplink zu verwenden, und den LTE Uplink für andere Endgeräte zu verwenden. Verschlechtern sich die Bedingungen, kann dann ebenfalls schnell umkonfiguriert werden.

2.6 Grundsätzliche Prozeduren und Mobility Management im Non-Standalone Mode

Nachdem nun das Konzept des 5G Resource Grids und der Bearer bekannt sind, beschreibt das folgende Unterkapitel, wie ein 5G Non-Standalone Split Bearer aufgebaut wird. Wurde ein Bearer aufgrund von Inaktivität oder Wechsel in den Flugmodus abgebaut, wird dieser aus folgenden Gründen dann wieder erneut aufgebaut:

- Das Modem empfängt von höheren Schichten des Protokollstapels neue Daten, die zum Netzwerk zu senden sind.
- Das Endgerät antwortet auf eine eingehende Paging Nachricht, die vom Netzwerk ausgesendet wird, wenn neue Daten aus dem Internet für das Endgerät eintreffen.
- Der Nutzer deaktiviert den Flugmodus und das Gerät verbindet sich erneut mit dem Netzwerk. Hier wird nicht nur der Bearer wieder neu aufgebaut, sondern es findet auch eine erneute Authentifizierung statt und die Ende-zu-Ende Verbindung mit dem Internet wird wiederhergestellt. Diese umfangreichere Prozedur wird nachfolgend beschrieben.

2.6.1 Aufbau eines LTE-Only Bearers als 5G Anker aus dem Flugmodus

Im Non-Standalone Modus ist die Grundlage für einen 5G Split Bearer zunächst der Aufbau einer LTE Verbindung zum eNB. Der LTE Anker für den Split Bearer wird dabei fast genauso wie eine normale LTE Verbindung aufgebaut. Dieser Vorgang ist im Kapitel über LTE bereits im Detail beschrieben und wird nachfolgend nochmals zusammengefasst.

Nach verlassen des Flugmodus sucht das Endgerät zunächst den Broadcast Kanal der umliegenden LTE Zellen. Um sich mit einem Netzwerk zu verbinden, wählt es dann die beste Zelle aus und führt eine Random Access Prozedur auf dem Random Access Channel (RACH) durch. Dies wird in Abb. 2.14 gezeigt. Der Kanal erhielt seinen Namen, weil Resource Zuweisungen in LTE und 5G im Uplink und Downlink vom Netzwerk gesteuert werden. Für die erste Verbindungsaufnahme bzw. später aus dem RRC-IDle Zustand ist dies jedoch nicht möglich, da der eNB das Endgerät bisher noch nicht kennt und ihm somit auch keine Ressourcen zuweisen kann. Das Endgerät muss deshalb selber zu einer zufälligen (random) Zeit zu senden beginnen und tut dies auf dem RACH, für den im Resource Grid festgelegte RBs zur Verfügung stehen. Die genaue Position des RACH wird in den Broadcast Nachrichten der Zelle bekanntgegeben.

Nachdem ein Gerät in Uplinkrichtung die ersten Ressourcen zugeteilt bekommen hat, sendet es eine kurze Radio Resource Control (RRC) Connection Setup Request Nachricht, die als Grund für die Verbindungsaufnahme „Mobile Originated Signaling“

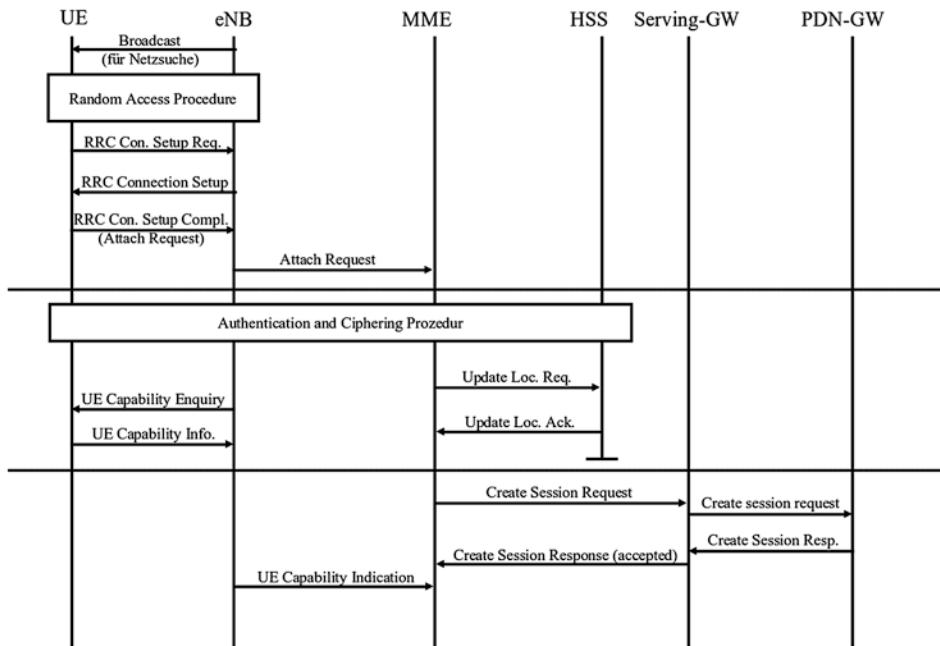


Abb. 2.14 Aufbau eines LTE Ankers für einen 5G EN-DC Bearer – Teil 1

enthält. Das Netzwerk baut dann einen logischen Signaling Radio Bearer (SRB-1) auf und informiert das Endgerät über die dessen Parameter in einer RRC Connection Setup Nachricht. Danach teilt der eNB weitere Ressourcen im Uplink zu und das Endgerät kann dann eine RRC Connection Setup Complete Nachricht über den neu aufgebauten SRB-1 Signaling Bearer übertragen.

Der wichtigste Teil dieser Access Stratum (AS) Nachricht ist die eingebettete Non-Access Stratum (NAS) „Attach Request“ Nachricht, die der eNB transparent an die Mobility Management Entity (MME) im Kernnetz weiterleitet. Wenn das Endgerät Dual-Connectivity mit NR unterstützt, setzt es ein Bit im UE Network Capability Information Element in der Attach Request NAS Nachricht und informiert das Kernnetzwerk außerdem über seine 5G Authentifizierungs- und Verschlüsselungsparameter am Ende der Nachricht. Dies kann dann von der MME für die PDN- und S-GW Auswahl verwendet werden¹⁵.

Nach der Authentifizierung und Aktivierung der Verschlüsselung bereiten die Kernnetzkomponenten den Nutzdatentunnel zwischen Endgerät und dem Internet vor. In der Zwischenzeit sendet der eNB eine UE Capability Request Nachricht an das Endgerät, um eine Liste der unterstützten Radiofunktionen des Endgerätes zu bekommen. Wenn das Endgerät EN-DC unterstützt, sendet es eine kurze Liste mit den wichtigsten Informationen über seine 5G Fähigkeiten am Ende der LTE Capabilities Liste:

```

EN-DC-r15: supported (0)
supportedBandListEN-DC-r15: 1 item
  Item 0
    SupportedBandNR-r15
      bandNR-r15: 78
      [...]

```

Diese Liste enthält die unterstützten 5G EN-DC Bandnummern und andere Informationen, wie z. B. die Anzahl der unterstützten MIMO Streams im Downlink und im Uplink. Falls der LTE eNB mit einem oder mehreren gNBs verbunden ist, frägt der LTE eNB dann das Endgerät nach einer erweiterten 5G Funktionsliste. Auf diese Anfrage sendet das Endgerät dann viele weitere Parameter, wie z. B. eine Liste der geeigneten LTE Anker Bänder für jedes unterstützte EN-DC Band und welche LTE Kanäle zusätzlich zum Dual Connectivity Modus aggregiert werden können. Diese Informationen werden auch an die MME im Kernnetz weitergeleitet.

Unterstützt ein Endgerät zusätzlich zu LTE Carrier Aggregation auch NR Carrier Aggregation für eine EN-DC Kombination, steigt die Anzahl der Bandkombinationen nochmals beträchtlich. Beispielsweise können die meisten LTE Bänder als Anker für NR Band n78 verwendet werden. Dies ist jedoch nicht für niedrigere NR Bänder wie n20 (800 MHz) oder n78 (700 MHz) der Fall. Wie schon zuvor beschrieben, werden für LTE und NR eigene Sendeeinheiten im Endgerät benötigt, die auf der Frequenzachse nicht dicht beieinander gleichzeitig betrieben werden können. Für NR auf niedrigen Frequenzen unterstützen manche Endgeräte nur LTE Anker in höheren Frequenzbändern wie Band 3 (1800 MHz), Band 1 (2100 MHz) oder Band 7 (2600 MHz). Somit hat die LTE Anker Zelle eine wesentlich geringere Reichweite als die NR Zelle auf einer niedrigeren Frequenz. Vor allem in ländlichen Szenarien kann es deshalb nötig werden, den EN-DC Mode zu deaktivieren und in den LTE-only Modus umzuschalten, wenn es einen LTE Kanal auf einem niedrigeren Frequenzband gibt.

Nachdem das Kernnetz eine Session für das Endgerät erzeugt hat, d. h. eine IP Adresse mit einem Quality of Service (QoS) Flow, sendet die MME eine Initial Context Setup Request Nachricht an den eNB. Diese enthält, wie in Abb. 2.15 gezeigt wird, eine Attach Accept NAS Nachricht für das Endgerät und alle Informationen, die der eNB benötigt, um einen LTE Radio Bearer für den IP Nutzdatenstrom aufzubauen.

Nach der Konfiguration des Radio Bearers auf der eNB Seite sendet dieser eine RRC Connection Reconfiguration Nachricht an das Endgerät. Diese enthält unter anderem folgende wichtige Parameter:

- Alle Parameter, um die Radio Bearer aufzubauen.
- Die Attach Accept Nachricht der MME.

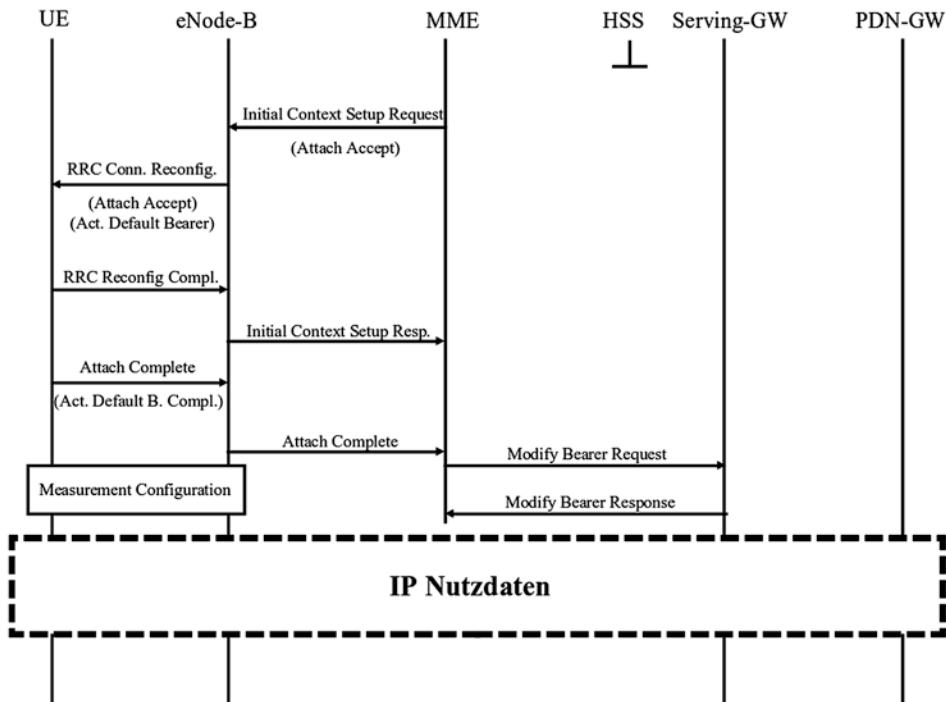


Abb. 2.15 Aufbau eines LTE Ankers für einen 5G EN-DC Bearer – Teil 2

Die Attach Accept Nachricht enthält eine eingebettete Activate Default EPS Bearer Context Request Nachricht, die wiederum die IP Adresse für das Endgerät, die IPv6 Konfigurationsinformationen, die DNS Server Adressen, sowie andere Bearer spezifische Informationen enthält.

Um unterschiedliche Datenströme eines Endgeräts auf der Luftschnittstelle auseinanderzuhalten, wie z. B. die IP Pakete von und zum Internet und die IP Pakete des VoLTE IMS Systems, hat jeder Bearer eine eigene EPS Bearer ID und die dazugehörigen Konfigurationswerte. Ein Unterschied zwischen diesen ist zum Beispiel die Konfiguration von Radio Link Control (RLC) Bestätigungen (Acknowledgements). Während IP Pakete von und zum Internet im RLC Acknowledged Mode übertragen werden, und somit im Falle des Verlusts auf niedrigeren Protokollschichten wiederholt werden, werden IP VoLTE Sprachdatenpakete nicht wiederholt.

Nachdem der Radio Bearer aufgebaut ist, sendet das Endgerät dann eine RRC Connection Reconfiguration Complete Nachricht an den eNB. Zum Core Netzwerk sendet das Endgerät dann eine Attach Complete Nachricht, sowie eine Activate Default EPS Bearer Context Complete Nachricht. Ab diesem Zeitpunkt können dann das Endgerät, das Radio Netzwerk und das Kernnetzwerk Nutzdaten weiterleiten. Eine weitere

Operation, die der eNB während dieses Prozesses ausführt, ist die Konfiguration von Radiomessungen mit einer RRC Connection Reconfiguration Nachricht. Mit Hilfe der Messkonfiguration kann das Endgerät dann dem eNB Measurement Reports schicken, um z. B. bei schlechterem Signal einen Handover zu einer anderen Zelle einzuleiten.

In einem weiteren Schritt, der in den Abbildungen nicht gezeigt wird, versucht der eNB außerdem, LTE Carrier Aggregation zu aktivieren. Dies geschieht durch die Konfiguration von Messungen der Nachbarkanäle und Bänder. Wenn ein Endgerät dort Signale empfängt, sendet es entsprechend Measurement Reports. Die Messungen werden dann beendet und LTE Carrier Aggregation wird durch hinzufügen von Secondary Cells zur Verbindung aktiviert. Details hierzu sind im LTE Kapitel beschrieben.

2.6.2 Hinzufügen einer NR Zelle im Non-Standalone Modus

Nach oder auch schon während des LTE Verbindungsauflaufbaus werden zusätzlich auch noch Messungen für 5G NR Zellen vom eNB konfiguriert. Dies ist der erste Schritt eines Dual Connectivity Setup, der in 3GPP TS 37.340¹⁶ beschrieben wird und in Abb. 2.16 dargestellt ist.

Alle Radio Resource Control (RRC) spezifischen Nachrichten verwenden das generische RRC Connection Reconfiguration Kommando und es ist üblich, Bearer Konfigurationsinformationen und Meßkonfigurationsdetails in einer einzigen Nachricht

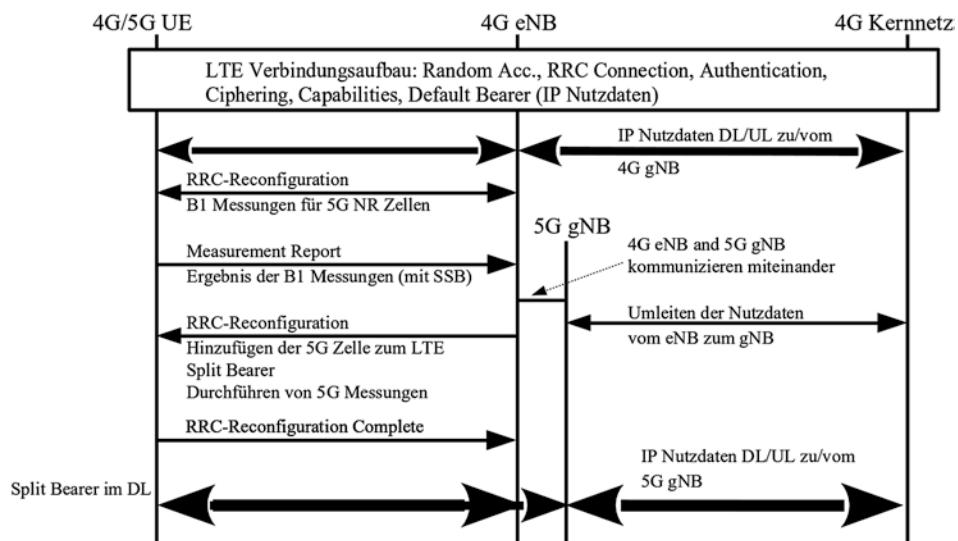


Abb. 2.16 Aufbau eines LTE/NR Split Bearer

zu finden. Da diese unabhängig voneinander sind, kann das Endgerät separat auf die einzelnen Bestandteile reagieren und auch antworten. Der folgende Ausschnitt zeigt ein Measurement Object, das Teil eines Measurement Setups ist, um nach einer 5G NR Zelle zu suchen:

```
measObjectNR
  carrierFreq: 643654
  rs-ConfigSSB
  measTimingConfig
  periodicityAndOffset: 20 subframes
  ssb-Duration: 5 subframes
  subcarrierSpacingSSB: 30 kHz
[...]
bandNR: 78
```

Während EUTRAN Measurement Objekte verwendet werden um LTE Kanäle zu beschreiben, werden NR Measurement Objekte verwendet, um die grundsätzlichen Parameter eines 5G Kanals zu beschreiben. Das gerade gezeigte Beispiel enthält die NR Carrier (Kanal) Frequenz, bei der das Endgerät nach einem Träger suchen soll. Der Wert kann über eine Formel in eine Frequenz von 3654,81 MHz umgerechnet werden¹⁷. Das Endgerät wird außerdem aufgefordert, nach SSBs zu suchen, die in 5 Subframes über eine Periode von 20 ms verteilt sind. Der verwendete Abstand der Subcarrier in diesem Kanal beträgt 30 kHz.

Zusätzlich zum Measurement Object wird außerdem ein Measurement Report Configuration Object benötigt. Dieses beschreibt, wie die Zellen zu messen und zu melden sind:

```
reportConfigInterRAT
  triggerType: event
  event
    eventId: eventB1-NR
    eventB1-NR
      b1-ThresholdNR: nr-RSRP
      nr-RSRP: -115 dBm
      hysteresis: 1 dB
      timeToTrigger: 40 ms
    maxReportCells: 8
    reportInterval: 5.12 seconds
    reportAmount: 1
    reportQuantityCellNR
      ss-rsrp: true
      ss-rsrq: false
```

```

ss-sinr: false
maxReportRS-Index
reportQuantityRS-IndexNR
ss-rsrp: true
ss-rsrq: false
ss-sinr: false

```

In diesem Beispiel wird das Endgerät aufgefordert, eine Inter-Radio Access Technology (inter-RAT) Messung durchzuführen, da 5G NR aus Sicht von LTE eine andere Technologie ist. Der Measurement Type ist auf Event B1-NR gesetzt, und es werden somit Zellen gemeldet, die eine konfigurierte Signalstärke überschreiten. In diesem Beispiel sind die Reference Signal Received Power (RSRP) Messungen auf ein Überschreiten von -115 dBm während 40 ms konfiguriert. Außerdem sollen maximal 8 Zellen gemeldet werden und das Intervall für die Reports wird auf 5,12 s festgelegt, wobei nur ein Report gesendet werden soll. Da Beamforming eine elementare Funktion von 5G ist, enthält die Messkonfiguration auch Informationen, wie Beams gemessen und gemeldet werden sollen. In diesem Beispiel soll ebenfalls der RSRP der SSBs gemessen werden. Die Reference Signal Received Quality (RSRQ) und die Signal to Interference and Noise Ratio (SINR) sollen dagegen nicht gemessen werden.

Der dritte Teil einer Messkonfiguration in der RRC Reconfiguration Nachricht ist die Measurement ID. Diese kombiniert Measurement Objekte mit Report Konfigurationen. Auf diese Weise kann die gleiche Report Konfiguration für mehrere Measurement Objects verwendet werden. Falls nur wenige Zellen gefunden und gemeldet werden müssen, benötigt eine Messung und Dekodierung der SSBs typischerweise nur wenige hundert Millisekunden. Im folgenden Beispiel wurde die Primary LTE Zelle (PCell) mit einer Signalstärke von -89 dBm gemessen, sowie mit einer Signalqualität von -7 dB. Die Liste der Nachbarzellen enthält dann das Resultat der B1-NR Messungen. Im folgenden Beispiel wurde nur eine Zelle mit der Physical Cell ID (PCI) 42 gefunden, und zwar mit einer Signalstärke des SSB 4 von -98 dBm.

```

measurementReport
measurementResults
measurementId: 3
measResultPCell
    rsrp-Result: -89dBm
    rsrq-Result: -7dB
measurementResultNeighborCellListNR: 1 item
    Item 0
        MeasurementResultCellNR
            pci: 42
        measurementResultCell
            rsrp-Result: -98dBm

```

```

measurementResultRS-IndexList: 1 item
  Item 0
    MeasurementResultSSB-Index
      ssb-Index: 4
      measurementResultSSB-Index
        rsrp-Result: -98dBm

```

Nachdem der eNB den Measurement Report erhalten hat, wird die Suche nach NR Zellen deaktiviert und der eNB nimmt Verbindung mit dem für die Zelle mit PCI 42 zuständigen gNB auf. Hat der gNB genug Kapazität für einen weiteren Nutzer, konfiguriert er daraufhin einen Split Bearer. Außerdem ändert der gNB dann den S1 Bearer, damit alle IP Nutzdatenpakete vom Serving-Gateway nicht länger zum eNB übertragen werden, sondern direkt zum gNB. Danach sendet der eNB eine RRC Connection Reconfiguration Nachricht an das Endgerät, um den neuen EN-DC Bearer zu aktivieren. Diese Nachricht enthält einige hundert Konfigurationsparameter für die NR Luftschnittstelle. Somit ist diese Nachricht um ein vielfaches größer als bei LTE. Die folgende Liste gibt einen Überblick über die wichtigsten Parameter dieser Nachricht:

- Die ID des Bearers, zu dem der 5G Teil hinzugefügt wird. Typischerweise wird EN-DC nur für den Internet Bearer aktiviert, während alle IP Pakete des VoLTE Bearers nur auf der LTE Luftschnittstelle übertragen werden.
- Discontinuous Reception (DRX) Parameter.
- Die maximale Sendeleistung, die das Endgerät im Uplink verwenden darf.
- Die Absolute Radio Frequency Channel Number (ARFCN) des SSB Blocks. Wie zuvor beschrieben, kann die Frequenz mit diesem Parameter berechnet werden.
- Das Subcarrier Spacing des NR Carriers (z. B. 30 kHz für Band n78).
- Die Kanal (Carrier) Bandbreite (z. B. 60, 80, 90 oder 100 MHz für Band n78).
- Die für das Endgerät konfigurierten Bandwidth Parts (BWPs).
- Parameter, die den Physical Downlink Control Channel (PDCCH) beschreiben, wie z. B. Message Aggregation Levels, etc.
- Die Konfiguration des Physical Downlink Shared Channels (PDSCH).
- Wie der Physical Uplink Control Channel (PUCCH) für HARQ Feedback konfiguriert ist.
- Wie der Physical Uplink Shared Channel (PUSCH) für die Nutzdatenübertragung konfiguriert ist.
- Uplink/Downlink Slot Konfiguration (z. B. DDDSU) und die Konfiguration des Special (S) Slots.
- Die ID des Endgerätes, die der gNB auf der 5G Luftschnittstelle verwenden wird.
- Konfiguration des Random Access Channels (RACH).
- Timer für Events wie z. B. Radio Link Failures (RLFs).
- Welche Modulation und Kodierung im Downlink verwendet werden können.
- Reference Signal (CS-RSI) Konfiguration.

- Konfiguration der optionalen Sounding Reference Signale (SRS). Werden diese konfiguriert, muss das Endgerät diese im Uplink senden.
- Codebook Konfiguration für das Kanalfeedback.
- Messkonfiguration für den 5G Teil der Verbindung, z. B. für die Meldung an das Netzwerk, wenn der Signalpegel der Zelle unter einen konfigurierten Wert fällt (A2-NR Event), oder falls der Signalpegel einer Nachbarzelle einen konfigurierten Wert überschreitet (A3-NR Event).
- Konfiguration der PDCP Schicht.
- Data Split Threshold für den Uplink. Mit diesem Parameter kann bestimmt werden, ob der LTE- oder der 5G-NR Uplink für Daten verwendet werden soll und ab welchem Pufferfüllstand beide Seiten der EN-DC Verbindung für Uplinkdaten verwendet werden sollen. Ein konfigurierter Pufferfüllstand von unendlich (infinity) wird verwendet, um für den Uplink nur die LTE Seite oder nur die NR Seite zu verwenden.
- Aktivierung und Verwendung der Verschlüsselung. Bei Aktivierung eines Split Bearers wird der Secondary Key verwendet (NR Seite). Wird der Split Bearer später wieder abgebaut, wird wieder zum Primary Key (LTE Seite) gewechselt.

Nachdem das Endgerät diese Reconfiguration Nachricht empfangen hat, bestätigt es dem eNB den Empfang und führt dann eine Random Access Prozedur auf 5G NR durch, um seine Verfügbarkeit auch dem gNB zu signalisieren. Aber diesem Zeitpunkt ist dann der gNB für den Split der Nutzdaten im Downlink zuständig.

Im Beispiel in Abb. 2.17 wurde für den Uplink ein LTE-only Bearer konfiguriert. Somit sendet das Endgerät seine Nutzdatenpakete nur dem eNB, der diese dann an den gNB weiterleitet. Von dort aus werden die Nutzdaten vom gNB an das Serving-Gateway im Kernnetz weitergeleitet und von dort aus dann zum Internet. Heute ist es üblich, auch im Uplink einen Split Bearer zu konfigurieren. Nur in Fällen, in denen wenig Daten im Uplink Puffer des Endgeräts bereitstehen, oder bei begrenzter Sendeleistung, werden diese dann nur über den LTE Pfad übertragen.

In den 3GPP Spezifikationen werden die LTE Zellen eines Dual-Connectivity Bearers als Primary Cell Group (PCG) bezeichnet. Ein PCG enthält den Primary Component Carrier (PCC), der den Anker der EN-DC Verbindung darstellt, und typischerweise einen oder auch mehrere Secondary Component Carrier (SCC), die Teil der Carrier Aggregation Funktion sind. Der 5G NR Teil der Dual Connectivity Verbindung wird als Secondary Cell Group (SCG) bezeichnet. Dieser umfasst den Primary NR Component Carrier und, optional, ein oder mehrere Secondary Component Carrier, falls auch auf der NR Seite Carrier Aggregation verwendet wird. In der Praxis wird heute jedoch NR Carrier Aggregation noch nicht von allen Netzwerken unterstützt und auch ältere Endgeräte können keine NR Kanäle bündeln.

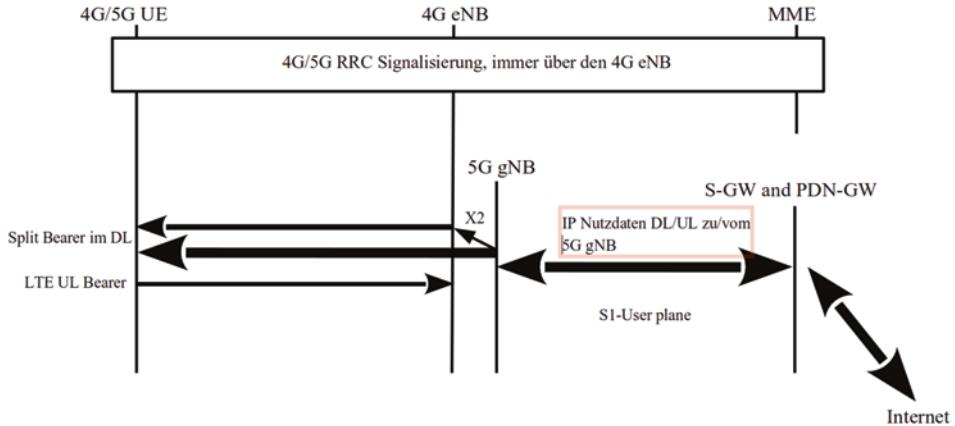


Abb. 2.17 Split Bearer Konfiguration im Downlink und LTE-only Bearer im Uplink

2.6.3 5G Anzeige im Display

Eine interessante Frage für EN-DC Netzbetreiber und Endgerätehersteller ist, wann ein 5G Logo im Display angezeigt werden soll. Bei GSM und LTE wird ein entsprechendes Logo angezeigt, wenn das Endgerät in einem solchen Netzwerk eingebucht ist. Würde der gleiche Ansatz auch für EN-DC verwendet werden, könnte das 5G Logo nur angezeigt werden, während ein NR Teil zu einer LTE Anker Verbindung aktiv ist. Somit würde der Netzwerkindikator im Display ständig zwischen 4G und 5G wechseln. Auch im RRC Idle State würde nur das LTE Logo gezeigt werden.

Um das 5G Logo auch im RRC-Idle Zustand anzeigen zu können, sowie in Fällen, in denen für einige Zeit nur ein LTE Anker verwendet wird, spezifizierte 3GPP das Upper Layer Indication Bit in der LTE System Information Broadcast Nachricht 2 in 3GPP TS 36.331¹⁸. Trotz Diskussionen, diesem Bit einen Namen zu geben, der dessen Zweck besser beschreibt, blieb dieser bis heute unverändert bestehen¹⁹.

```

SIB2
[...]
freqInfo
    ul-Bandwidth: n100 (5)
    additionalSpectrumEmission: 1
timeAlignmentTimerCommon: sf1920
plmn-InfoList-r15: 1 item
    Item 0
        PLMN-Info-r15
            upperLayerIndication-r15: true

```

In der Praxis gibt es mit diesem Ansatz zwei Herausforderungen: Zum einen ist das Bit nicht direkt mit der 5G Aktivierung der Zelle verbunden. Ein Netzbetreiber könnte also das Bit setzen, auch wenn keine 5G Zelle in der Nähe vorhanden ist. Auch kann das Netzwerk nicht gewährleisten, dass ein Endgerät das 5G Band unterstützt, das in einem Gebiet verwendet wird. Frühe 5G Endgeräte unterstützten beispielsweise nur 5G in Band n78. In ländlichen Gebieten verwenden Netzbetreiber jedoch 5G auch in Band n28, n3 und n1. Wird das Bit hier auf ‚true‘ gesetzt, was durchaus aus Sicht eines Netzbetreibers Sinn macht, werden nicht alle Endgeräte dort auch 5G verwenden können.

Wie und wann das Bit verwendet wird, um das 5G Logo anzuzeigen, ist Netzerkbetreiber und Endgerätehersteller spezifisch. Manche Endgerätehersteller verwenden z. B. leicht unterschiedliche 5G Logos, um unterschiedliche Zustände anzuzeigen. Ein Ansatz ist beispielsweise, ein 5G Logo anzuzeigen, wenn das Bit in einer LTE Zelle gesetzt ist, dieses jedoch leicht optisch zu modifizieren, wenn ein 5G Kanal auch tatsächlich zu einer LTE Verbindung hinzugenommen wird.

Netzbetreiber gewähren üblicherweise nicht allen Kunden den Zugang zum 5G Netzwerk und benötigen deswegen eine Möglichkeit, dem eNB und dem Endgerät mitzuteilen, ob 5G zu einer LTE Verbindung hinzugenommen werden darf. Dazu wurde in der Teilnehmerdatenbank, dem Home Subscriber Server (HSS, siehe Kapitel zu LTE), ein 5G Aktivierungsparameter pro Teilnehmer hinzugefügt. Während des Attach Vorgangs informiert das HSS die Mobility Management Entity (MME), ob das Dual Connectivity (LTE-NR) Information Element im Datenbankeintrag des Kunden gesetzt ist. Ist dies nicht der Fall, informiert die MME dann während des Einbuchvorgangs das Endgerät durch Setzen des Dual Connectivity-New Radio Restriction (DCNR) Bit in der Attach Accept Nachricht. Ist das Bit gesetzt, ignoriert das Endgerät dann das Bit in SIB-2. Die MME informiert ebenfalls den eNB, damit dieser für das entsprechende Endgerät keinen 5G Bearer zu einer LTE Verbindung hinzuschaltet, auch wenn dies vom Endgerät unterstützt wird.

2.6.4 Handover Szenarien

Nachdem ein Dual Connectivity Bearer aufgebaut wurde und das Endgerät Messinstruktionen für den LTE- und den NR-Teil der Verbindung erhalten hat, prüft es ständig den Kanal und meldet periodisch oder bei Erreichen der konfigurierten Schwellwerte die Signalstärke und Qualität der aktuellen Zelle und eventuell gefundenen Nachbarzellen. Dies kann sehr flexibel konfiguriert werden und Netzbetreiber und Hersteller haben unterschiedliche Ansätze, wie dieses Mobility Management in der Praxis durchgeführt wird. Wie im Kapitel über LTE schon ausführlich gezeigt, werden hierzu die Events A1 bis A5, sowie B1 und B2 verwendet. Diese existieren nicht nur für LTE, sondern auch für 5G NR. Da die LTE und NR Teile einer EN-DC Verbindung unabhängig voneinander sind, werden diese auch für jede Seite unabhängig vom Endgerät gemeldet. Sowohl LTE,

als auch NR Messergebnisse werden jedoch zum eNB gemeldet, der dann die NR Messergebnisse an den gNB für die Analyse weiterleitet.

Wird die NR Netzabdeckung verlassen, während LTE weiterhin vorhanden ist, meldet das Endgerät dem gNB, dass die aktuelle (Serving) NR Zelle zu schwach geworden ist (NR Event A2). Der gNB beendet dann den NR Teil des EN-DC Bearers und gibt die Kontrolle zurück an den eNB. Somit wird dann aus der Verbindung über die Luftschnittstelle wieder ein ganz normaler LTE Bearer.

Bewegt sich das Endgerät zwischen zwei LTE/NR Zellen, gibt es mehrere Möglichkeiten, einen Handover durchzuführen. Im einfachsten Fall wartet das Netzwerk, bis eine LTE Nachbarzelle stärker als die aktuelle LTE Zelle wird. Das Endgerät sendet in diesem Fall ein LTE Event A3. Das Netzwerk konfiguriert daraufhin einen Handover, aus der EN-DC Verbindung wird nach dem Handover in die Nachbarzelle jedoch zunächst eine LTE-only Verbindung. Der neue eNB konfiguriert dann Inter-RAT Event B1, um nach einer neuen 5G NR Zelle zu suchen. Diese wird dann nach Rückmeldung des Endgerätes wieder hinzugefügt. Dies bedeutet in der Praxis, dass 5G für einen kurzen Moment nach dem Handover nicht verwendet wird.

Da der LTE- und der NR-Teil einer Verbindung voneinander unabhängig sind, ist es auch möglich, den Handover des LTE-Teils und den Handover des NR-Teils einer Verbindung unabhängig voneinander durchzuführen. Dies wird in 3GPP TS 37.340 beschrieben. Dazu ist es notwendig, dass ein LTE eNB mit einem NR gNB an einem anderen Standort kommunizieren kann. Dies geschieht über das logische X2 Interface, das unterschiedliche Zellstandorte miteinander verbindet. Dieses Interface gibt es schon bei LTE und wurde für die Verwendung für 5G NR und EN-DC entsprechend erweitert. Abb. 2.18 zeigt ein solches Szenario.

Zunächst ist das Endgerät in Schritt 1 mit einem eNB und einem gNB am gleichen Standardort verbunden. Bewegt sich der Nutzer dann in ein Gebiet, in dem die benachbarte NR Zelle stärker als die aktuelle (Serving) NR Zelle wird, sendet das Endgerät ein NR Measurement Event A3 zum gNB. Der folgende Ausschnitt zeigt, wie ein solches Event laut 3GPP TS 38.331 gemeldet wird. In diesem Beispiel liegt der Empfangsspegel

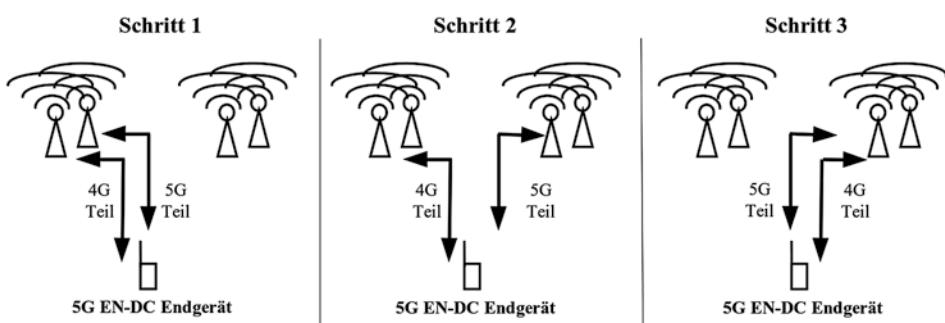


Abb. 2.18 Durchführung einer EN-DC Handover Prozedur

der Serving Cell (Reference Signal Received Power, RSRP) bei -101 dBm , während die Nachbarzelle deutlich besser mit -90 dBm empfangen werden kann. Zusätzlich zum Empfangspegel meldet das Endgerät auch die ID des SSB Beams, der am besten Empfangen werden kann.

```
[...]
measResultServingMOList
    MeasurementResultServMO
        servCellId: 16
        measResultServingCell
            physCellId: 64
            measResult
                cellResults
                    resultsSSB-Cell
                        SS-RSRP: -101dBm
[...]
measResultNeighNRCells
    measResultListNR
        MeasResultsNR
            physCellId: 65
            measResults
                cellResults
                    resultsSSB-Cell
                        SS-RSRP: -90dBm
            rsIndexResults
                resultsSSB-Indexes: 1 item
                    Item 0
                    ResultsPerSSB-Index
                        ssb-Index: 4
[...]
```

Nachdem der gNB das Messergebnis empfangen hat, sucht er nach dem gNB, der für die gemeldete PCI verantwortlich ist und baut zu diesem eine Verbindung auf. Dazu wird das X2 Interface verwendet, auch wenn der neue gNB lediglich ein anderer Sektor am gleichen Standort ist. Befindet sich der neue gNB an einem anderen Standort, fließen die Daten dann über das X2 Interface zunächst zum ersten Aggregation Router und von dort, statt weiter zum Kernnetzwerk, zurück zum anderen Standort. Auch wenn das X2 Interface logisch eine direkte Verbindung zwischen den eNBs und gNBs herstellt, ist dies physisch somit in diesem Szenario nicht der Fall. Kann der neue gNB die zusätzliche Verbindung bedienen, wird Kapazität auf der NR Luftschnittstelle für das neue Endgerät reserviert und eine Antwort an den aktuellen gNB gesendet. Dieser erstellt dann eine RRC Connection Reconfiguration Nachricht und schickt diese an den LTE gNB, der

diese dann wiederum an das Endgerät weiterleitet. Das Endgerät führt daraufhin einen Handover zum neuen gNB durch. In einem für das Endgerät transparenten Schritt wird während dieser Prozedur die Verbindung für die Nutzdaten mit dem Serving-Gateway im Kernnetz geändert, da nun der neue gNB dafür zuständig ist. Der neue gNB leitet dann einen Teil der Nutzdaten über das X2 Interface zum eNB am anderen Standort weiter. Zu diesem Zeitpunkt sendet und empfängt das Endgerät dann seine Nutzdaten von zwei unterschiedlichen Standorten. Dies ist in Schritt 2 in Abb. 2.18 gezeigt.

Da sich die Signalstärken des 4G und 5G Teils einer Verbindung sehr ähnlich verändern, wird nach einer gewissen Zeit auch der LTE-Teil der Verbindung zum einem neuen Standort wechseln. Dies geschieht, sobald das Endgerät eine im Vergleich zur aktuellen Zelle stärkere LTE Zelle meldet (LTE Event A3). Daraufhin wird dann eine Handover Prozedur für den LTE Teil der Verbindung gestartet. Zusätzlich zum Handover auf der Luftschnittstelle wird auch hier die X2 Verbindung modifiziert, damit das Routing der Nutzdaten zwischen dem 5G gNB und der LTE-Seite entsprechend angepasst wird. Nach dem Handover ist dann die neue LTE Zelle der Anker der Verbindung. Dies ist in Teil 3 in Abb. 2.18 gezeigt.

Natürlich ist es in der Praxis auch möglich, dass zunächst der LTE-Teil einer Dual-Connectivity Verbindung an einen anderen Standort weitergegeben wird. Ist der Abdeckungsbereich des LTE Teils und des NR Teils nicht identisch, z. B. weil unterschiedliche Antennen verwendet werden, die nicht in die gleiche Richtung zeigen, ist es sogar möglich, und oft in der Praxis auch der Fall, dass ein Endgerät für längere Zeit mit unterschiedlichen Standorten kommuniziert.

2.6.5 EN-DC Signaling Radio Bearers

Während Dedicated Radio Bearer (DRB) für Nutzdatenpakete verwendet werden, gibt es für EN-DC Signalisierungsinformationen mehrere Signaling Radio Bearer (SRB). LTE SRB-0 nutzt den LTE Common Control Channel im Uplink und Downlink während des Aufbaus einer neuen Verbindung bis zum Zeitpunkt, ab dem LTE SRB-1 verwendet werden kann. Zusätzlich dazu wird nach der Authentifizierung und Aktivierung der Verschlüsselung LTE SRB-2 aufgebaut, über den RRC Nachrichten mit eingebetteten NAS Nachrichten von und zum Kernnetzwerk übertragen werden können. Für die NR Seite einer EN-DC Verbindung gibt es unterschiedliche Möglichkeiten, Signalisierungsnachrichten zu übertragen²⁰:

Option 1 (wird heute typischerweise verwendet): Der eNB nutzt LTE SRB-1/2 für seine RRC Nachrichten. Wenn der gNB RRC Nachrichten mit dem Endgerät austauschen will (z. B. für Messergebnisse), werden die RRC Nachrichten über das X2 Interface zwischen eNB und gNB ausgetauscht. Vom eNB werden die Signalisierungsnachrichten dann über SRB-1/2 von und zum Endgerät weitergeleitet.

Option 2: Der eNB nutzt LTE SRB-1/2 für seine RRC Nachrichten, und der gNB baut seinen eigenen Signaling Bearer, SRB-3, zum Endgerät auf. Die Unterstützung von SRB-3 im Endgerät ist jedoch optional. Während eine solche Konfiguration ideal ist, um die LTE- und NR-Signalisierung zu trennen, erfordern manche Operationen jedoch eine Koordinierung von eNB und gNB. Dazu gehört z. B. der Aufbau weiterer Component Carrier. In solchen Fällen kann dann SRB-3 nicht verwendet werden. Stattdessen wird Option 1 verwendet.

Option 3: Der eNB verwendet einen Split-Bearer für die Signalisierung. SRB-1/2 Nachrichten und gNB RRC Nachrichten, die in SRB-1/2 Nachrichten eingebettet sind, werden entweder über die eNB Luftschnittstelle, über die gNB Luftschnittstelle oder über beide Pfade gleichzeitig übertragen. Dafür ist kein SRB-3 notwendig, es gibt dem Endgerät jedoch zusätzliche Flexibilität, schnell auf veränderte Radiobedingungen zu reagieren.

2.6.6 5G Non-Standalone und VoLTE

VoLTE ist das von der GSMA spezifizierte 3GPP IMS-basierte Sprachprofil für LTE. Wie im VoLTE Kapitel beschrieben, definiert das Profil, welche Optionen der IMS Spezifikation Netzwerke und Endgeräte implementieren müssen, um interoperabel zu sein. Im 5G Non-Standalone Modus wird der VoLTE Sprachdienst ohne Änderungen übernommen. Da VoLTE ein IP-basierter Dienst ist, wäre es ohne Weiteres möglich, einen 4G/5G Split Bearer für die IMS Signalisierung und auch für die Sprachdatenpakete zu verwenden. In der Praxis haben sich jedoch die meisten Netzbetreiber entschieden, den VoLTE Bearer nur über das LTE Netzwerk zu leiten. Der Grund dafür ist, dass sich der LTE-Anker für den EN-DC Bearer der Internet Verbindung meist auf einem niedrigeren Frequenzband als 5G befindet, was sich positiv auf die Signalqualität vor allem am Zellrand auswirkt. Somit ist die Sprachqualität auch während schneller Zellwechsel besser. Außerdem wird ein VoLTE Gespräch nicht von der Umkonfiguration eines NR Radio Bearer beeinflusst. Manche Netzbetreiber beenden sogar aktiv einen EN-DC Internet Bearer beim Aufbau des Dedicated Bearers für ein VoLTE Gespräch. Andere Netzbetreiber wiederum behalten den EN-DC Split Bearer für die Internet Verbindung zunächst bei und bauen diesen erst während eines Handovers ab. Der Vorteil dieser Ansätze ist, dass die Sendeleistung des Endgeräts ohne den EN-DC Bearer komplett von der LTE-Seite des Modems verwendet werden kann.

2.7 Netzwerkplanung und Rollout Aspektse

Nachdem im bisherigen Teil dieses Kapitels die 5G NR TDD und FDD Luftschnittstelle in High-, Mid- und Low-Bändern des Frequency Range 1 eingeführt wurde, betrachtet dieses Unterkapitel nun einige Aspekte der 5G Netzwerkplanung und des Netzausbau.

2.7.1 Die Reichweite von Band n78

Das wichtigste Band für 5G NR in Europa und Asien zur Steigerung der Netzkapazität ist Band n78 im 3,5 GHz Bereich. Da höhere Frequenzbereiche eine geringere Reichweite im Vergleich zu niedrigeren Frequenzbereichen haben, ist der Abdeckungsbereich eines n78 Signals kleiner, als der Bereich, der vom gleichen Standort über niedrigere Frequenzen erreicht werden kann. Vor allem im städtischen Umfeld, in dem dieses Band hauptsächlich verwendet wird, ist die Standordtichte jedoch sehr hoch. Der typische Radius eines Standorts im städtischen Bereich in Europa ist etwa 200 m. Mit dieser Distanz werden LTE Handover schon bei einer typischen Signalstärke von -95 dBm durchgeführt. In ländlichen Gebieten, in denen Standorte mehrere Kilometer auseinanderliegen, wird typischerweise ein Handover bei -110 dBm durchgeführt. Dies ist ein Unterschied von 15 dB, die Signalstärke ist beim Handover hier also um den Faktor 30 geringer. Das bedeutet, dass sich im städtischen Umfeld sogar 5G NR Zellen auf Band n78 deutlich überlappen, jedoch, verglichen mit einem LTE Kanal in Band 3 (1800 MHz), mit einem etwas geringeren Signalpegel am Zellrand. Aus diesem Grund gibt es typischerweise auch keine Bereiche außerhalb von Gebäuden in Städten, die nicht mit Band n78 abgedeckt werden können, ohne dabei die Anzahl der Standorte zu erhöhen.

In der Presse ist oftmals zu lesen, dass die 5G Versorgung in Band n78 sehr schnell endet, wenn sich ein Nutzer vom Zellmittelpunkt entfernt. Dies wird dann üblicherweise den schlechteren Ausbreitungseigenschaften von Band n78 angelastet. In vielen Fällen verliert ein Endgerät die 5G Abdeckung jedoch nicht weil kein Signal mehr vorhanden wäre, sondern weil das Netzwerk einen Handover des LTE Anker zu einer benachbarten Station durchführt, die nur über LTE Komponenten verfügt, die keine inter-Site Anchoring Verbindung mit einer benachbarten 5G Zelle aufbauen können.

Im städtischen Bereich ist die n78 Abdeckung eher bei der Versorgung innerhalb von Gebäuden limitiert, da hier typischerweise aufgrund der großen Wand- und Fensterdämpfung niedrigere Frequenzen einen Vorteil haben. Auf der LTE Seite sind dafür z. B. Band 20 (800 MHz) und Band 8 (900 MHz) geeignet.

2.7.2 Backhaul Betrachtungen

Aktuell sind LTE Standorte typischerweise über IP Router und optischen 1 Gbit/s SFP (Small Form Factor Pluggable) Transceivern²¹ oder per schnellem Mikrowellenfunk an das Kernnetz angebunden. Diese Verbindung wird auch als „Backhaul“ bezeichnet. Eine Geschwindigkeit von 1 Gbit/s ist üblicherweise für einen Standort mit drei GSM und LTE Sektoren ausreichend. Die kombinierte theoretisch mögliche Spitzendatenrate aller drei Sektoren übersteigt zwar oft diesen Wert, vor allem dann, wenn mehrere LTE Kanäle an einem Standort für Carrier Aggregation verwendet werden. Es ist jedoch unwahrscheinlich, dass alle drei Sektoren gleichzeitig voll ausgelastet sind, und nur mit

Endgeräten kommunizieren, die Daten mit der höchsten möglichen Geschwindigkeit auf der Luftschnittstelle empfangen können. Für 5G n78, mit dem das genutzte Spektrum eines Standorts verdreifacht werden kann, ist eine 1 Gbit/s Backhaul Verbindung nicht mehr ausreichend. Ist der Mobilfunkstandort mit einer Glasfaser angebunden, kann hier durch Tausch der Hardware und der Fiber Transceiver mit SFP+ Modulen auf eine Geschwindigkeit von 10 Gbit/s aufgerüstet werden. Die Glasfaser selber muss nicht ersetzt werden. Wird Mikrowellenfunk für den Backhaul verwendet, sind heute auch Systeme auf dem Markt, die eine ähnliche Übertragungsgeschwindigkeit über mehrere Kilometer beherrschen²².

2.8 Die 5G NR Standalone (SA) Architektur und grundsätzliche Prozeduren

Während die meisten Netzbetreiber ursprünglich ihren 5G Rollout mit EN-DC Dual Connectivity gestartet hatten, und somit das vorhandene 4G Kernnetzwerk (EPC, Evolved Packet Core) weiterverwenden konnten, benötigt der nächste 5G Evolutions-schritt ein 5G Kernnetz (5G Core, 5GC). Dieses ist in folgenden 3GPP Dokumenten spezifiziert:

- 3GPP TS 23.501: Beschreibung der 5G System Architektur²³
- 3GPP TS 23.502: Beschreibung der Non-Access Stratum (NAS) Prozeduren, z. B. für das Mobility und Session Management²⁴
- 3GPP TS 24.501: Parameterdefinition für NAS Nachrichten²⁵

Da die Hauptaufgaben des 5GC große Gemeinsamkeiten mit denen des LTE Kernnetz aufweisen, soll hier nun zunächst ein Überblick über die 5GC Funktionen, Parameter und Prozeduren gegeben werden, sowie ein Vergleich zu 4G angestellt werden.

Ein wesentlicher Unterschied zwischen früheren Kernnetzwerken ist, dass 3GPP bei der Spezifikation des 5G Standards die neuesten Entwicklungen der IT Welt einbezogen hat, und den 5G Core in einer „Cloud Native“ Fassung spezifiziert hat. Das 5G Kernnetz kann somit in einer sogenannten Service Based Architektur umgesetzt werden kann. Dies wird im letzten Teil dieses Kapitels genauer beschrieben.

2.8.1 Funktionen des 5G Kernnetz

Abb. 2.19 zeigt die wichtigsten Komponenten des 5GC, die für ein funktionierendes Netzwerk notwendig sind. Grundsätzlich hat das 5GC eine große Ähnlichkeit mit dem LTE Kernnetz. Um es von früheren Architekturen zu unterscheiden, haben alle Komponenten neue Namen und Abkürzungen erhalten. Wie im LTE Kernnetz wird in

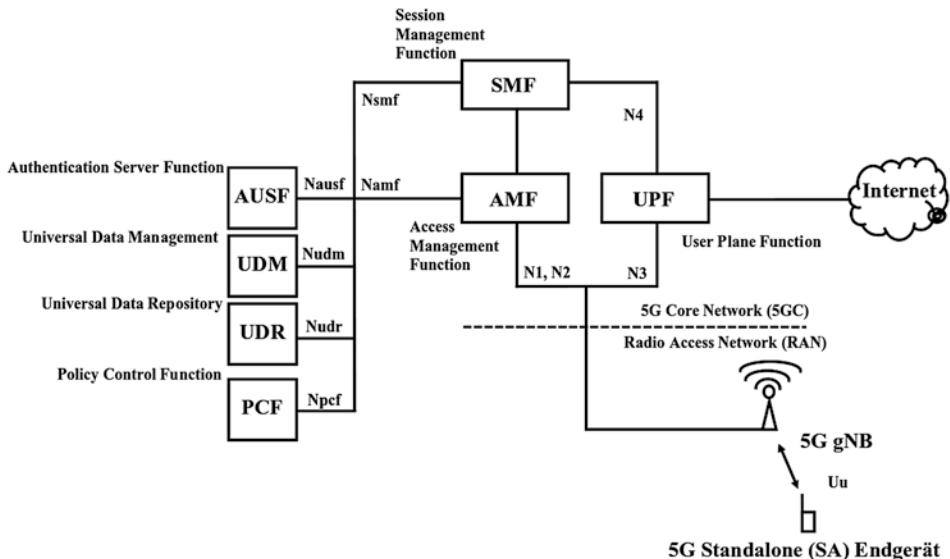


Abb. 2.19 Die wichtigsten Komponenten des 5G Kernnetz (5GC)

der Architektur zwischen Signaling Plane und User Plane unterschieden. Wie auch im Radionetzwerk werden über den logischen User Plane Teil die Nutzdaten übertragen. Die User Plane stellt also die Netzwerkverbindung für ein Endgerät her. In den meisten Fällen ist dies eine Verbindung zwischen Internet und Endgerät. Auf der anderen Seite gibt es die Signaling Plane, die von allen Netzelementen und auch dem Endgerät für Management Aufgaben verwendet wird. Management Aufgaben werden in Connection Management, Mobility Management und Session Management unterteilt. Sowohl die User Plane als auch die Signaling Plane verwenden das IP Protokoll zum Datentransport, und an vielen Stellen des Netzwerks werden diese über das selbe physische Netzinterface übertragen. Dies ist besonders auf den Backhaul Verbindungen der Fall, die die Mobilfunkstandorte mit dem Kernnetzwerk verbinden.

Wie auch in LTE sind im 5GC unterschiedliche Funktionen für die zwei Planes verantwortlich. Im 5GC wurden die notwendigen Funktionen im Vergleich zu LTE in kleinere Teile aufgeteilt. Dies ermöglicht es, diese als sogenannte Microservices in Containern in einer Cloud Umgebung zu implementieren. Bei LTE hingegen wurde noch angenommen, dass ein logisches Netzwerkelement einem physischen Server Rack entspricht. Aus diesem Grund wird im 5GC auch nicht mehr von „Network Entities“ gesprochen, sondern von „Network Functions“, die unabhängig von der Hardware sind.

Im weiteren Verlauf des Kapitels wird davon ausgegangen, dass das im LTE Kapitel beschriebene LTE Mobility- und Session Management bekannt ist. Im 5G Kern wurde die 4G Mobility Management Entity (MME) in zwei 5G Funktionen aufgeteilt: Die Access Management Function (AMF) und die Session Management Function (SMF).

Logisch kommuniziert das Endgerät mit beiden Funktionen, alle Signalisierungsnachrichten werden jedoch an die AMF gesendet. Die AMF bearbeitet dann die Connection- und Mobility Management Aufgaben und leitet alle Session bezogenen Anfragen an eine SMF weiter. Ob AMF und SMF auf dem gleichen physischen Server ausgeführt werden oder nicht, wurde im Standard bewusst nicht näher spezifiziert.

Das 4G HSS wurde im 5G Kernnetz in drei Funktionen aufgeteilt: Das Front-End zur AMF bildet die Authentication Server Function (AUSF), während die Teilnehmerinformationen von der Universal Data Management (UDM) Funktion zugänglich gemacht werden. Die eigentlichen Kundendaten (Subscription Information) sind in der Unified Data Repository (UDR) Funktion abgelegt.

Auf der User Plane, also den Komponenten, die Nutzdaten zwischen Endgerät und einem externen Netzwerk (z. B. dem Internet) weiterleiten, gibt es einen wesentlichen Unterschied zum LTE Kernnetz. Während im LTE Kernnetz die Userplane aus Serving-Gateway (S-GW) und dem PDN-Gateway (P-GW) besteht, gibt es im 5G Kernnetz nur eine Komponente, die User Plane Function (UPF). Auch im LTE Kernnetz sind S-GW und P-GW heute oft eine Komponente, da es sich gezeigt hat, dass die physische Trennung und Aufgabenverteilung auf zwei Komponenten in der Praxis keine Vorteile bringt.

Eine weitere Funktion, die heute in 5G Kernnetzen eingesetzt wird, ist die Policy Control Function (PCF). Diese ist für das Quality of Service Management der Nutzdatenbäuerer zuständig. Dazu ordnet diese Funktion jedem Bearer diverse Parameter zu, die die maximale Datenrate und Priorität des IP Datenstroms bei Überlast festlegen. Die Funktion entspricht somit der PCRF im 4G Kernnetzwerk.

Nicht in der Abbildung gezeigt ist der Security Edge Protection Proxy (SEPP). Diese Funktion wird zwischen 5G Kernnetzen unterschiedlicher Betreiber in unterschiedlichen Ländern geschaltet. Die Aufgabe dieser Funktion ist die Überwachung und Modifikation der Signialisierung zwischen Netzwerken für das internationale Roaming, und um Komponenten in anderen 5G Kernnetzwerken nur limitierten Zugang in das eigene Kernnetz zu gewähren.

2.8.2 Netzwerkschnittstellen

Damit Netzbetreiber Kernnetzfunktionen von unterschiedlichen Herstellern beziehen können, ist es notwendig, die Interaktion zwischen diesen zu standardisieren. Wie in Abb. 2.19 gezeigt, gibt es folgende Schnittstellen zwischen dem 5G RAN und dem 5GC:

- N1: Endgerät, AMF und SMF
- N2: gNB und AMF
- N3: gNB und UPF
- N4: SMF und UPF

Diese Schnittstellen sind auf „traditionelle“ Weise wie für frühere Netzstandards spezifiziert. Das bedeutet, dass Netzwerkkomponenten und Funktionen logische Verbindungen zueinander aufbauen und Signalisierungsnachrichten dann über diese logischen Verbindungen austauschen.

Zwischen Control Plane Funktionen im 5G Kernnetz werden gänzlich andere Schnittstellen verwendet:

- Namf: Kommunikation mit der Access Management Function (AMF)
- Nsmf: Kommunikation mit der Session Management Function (SMF)
- Nauf: Kommunikation mit der Authentication Function (AUSF)
- Nudm: Kommunikation mit dem Universal Data Management (UDM) Service
- Nudr: Kommunikation mit dem Unified Data Repository (UDR)
- Npcf: Kommunikation mit der Policy Control Function (PCF)

Der große Unterschied zu den „traditionellen“ Schnittstellen besteht darin, dass im 5G Kernnetz nicht mehr wie früher die Kommunikation zwischen zwei Netzwerkelementen standardisiert ist. Stattdessen bieten Netzwerkfunktionen nun eine Schnittstelle an, über die andere Funktionen im Netzwerk einen Service erreichen können. Statt einer statischen Kommunikationsbeziehung und einer zustandsbasierten Interaktion, bei der jede Komponente einen Kontext während des Datenaustausches vorhält, sind diese Schnittstellen zustandslos konzipiert. Für Anfragen an eine Funktion wird das HTTP Protokoll verwendet und Daten werden im JSON (JavaScript Object Notation) Format übertragen. Wie auch bei Webseiten steht jede Transaktion für sich und ist final, d. h. es wird kein Zustand vorgehalten, um die Kommunikation weiterzuführen. Für weitere Anfragen wird somit jedes Mal ein neuer HTTP Request gesendet, der völlig unabhängig von allen früheren Interaktionen ist.

2.8.3 Teilnehmer und Geräte IDs

In GSM und LTE Netzen wird ein Teilnehmer über die International Mobile Subscriber Identity (IMSI) identifiziert. Diese ist auf der SIM Karte und im Netzwerk im HSS hinterlegt. Im 5G Kernnetz wird ein Teilnehmer über den Subscription Permanent Identifier (SUPI) identifiziert²⁶. Auf der SIM Karte ist dieser identisch mit der IMSI. Im 5G Kern sind auch die Kommunikation mit Endgeräten ohne SIM Karte vorgesehen. Für diese Fälle wird als SUPI der Network Access Identifier (NAI) verwendet, der in RFC 4282²⁷ und 3GPP TS 23.003²⁸ beschrieben ist.

In allen bisherigen Netztechnologien sendet das Endgerät bei der ersten Verbindungsauftnahme mit dem Netzwerk seine IMSI zum Kernnetzwerk. Nach der Authentifizierung und Aktivierung der Verschlüsselung wird dann eine anonymisierte und temporäre Kennung an das Endgerät gesendet, die dann bei zukünftigen Verbindungsauftnahmen verwendet wird. Dies kann jedoch von Angreifern ausgenutzt werden, um das Endgerät

aufzufordern, seine IMSI bekanntzugeben, z. B. wenn eine gefälschte Basisstation dem Endgerät mitteilt, dass die zuvor temporär zugewiesene ID unbekannt ist. Das 5G Kernnetz verhindert dies, da das Endgerät hier eine verschlüsselte Version der SUPI senden kann. Diese wird als Subscription Concealed Identifier (SUCI) bezeichnet. Details hierzu werden weiter unten beschrieben.

Zusätzlich zur SUPI, die auf der SIM Karte gespeichert ist und den Teilnehmer identifiziert, haben auch Endgeräte ihren eigenen Identifier. In früheren 3GPP Netzen ist dies der International Mobile Equipment Identifier (IMEI). Im 5G Kernnetz wird diese ID als Permanent Equipment Identity (PEI) bezeichnet. Die PEI kann unterschiedliche Formate haben. Für Geräte, die die 3GPP NR Luftschnittstelle verwenden, sind PEI und IMEI identisch.

2.8.4 Prozeduren im 5G Kernnetz

Wie auch bei früheren Netztechnologien sind die Hauptaufgaben des Kernnetzes die Teilnehmerverwaltung, Verbindungssteuerung, Mobilitätsmanagement, sowie das Weiterleiten der IP Pakete (Nutzdaten) von und zum Teilnehmer über das Radio Netzwerk (RAN) auf der einen Seite, und einem externen Netzwerk wie dem Internet auf der anderen Seite. 3GPP TS 23.502 beschreibt folgende Aufgaben:

- Connection Management (CM)
- Registration Management (RM)
- Mobility Management (MM)
- Session Management (SM)

In den weiteren Abschnitten werden nun einige essenzielle Management Prozeduren beschrieben, wie diese zusammenspielen, und wie sich diese von GSM und LTE unterscheiden.

2.8.5 Connection Management

Möchte ein Endgerät eine Signalisierungsverbindung mit dem Kernnetz aufbauen, z. B. nachdem es eingeschaltet wurde, oder nachdem es für einige Zeit im ‚Idle‘ Zustand ohne Radioverbindung war, muss eine neue Verbindung zur Access Management Function (AMF) aufgebaut werden. Dazu sind zwei Signalisierungsverbindungen notwendig.

In 3GPP 5G NR Netzwerken wird die Verbindung zwischen Endgerät und gNB als Radio Resource Control (RRC) Verbindung bezeichnet. Diese ist für die Kommunikation über die Luftschnittstelle zwischen Endgerät und gNB notwendig. Ein Endgerät kann sich im RRC-Connected, RRC-IDLE oder RRC-Inactive Zustand befinden. Um Energie zu sparen und trotzdem schnell auf Nachrichten reagieren zu können, gibt es im

RRC-Connected Zustand noch die Discontinuous Transmission und Reception (DTX, DRX) Unterzustände.

Des Weiteren benötigt ein Endgerät auch eine Verbindung zur AMF im Kernnetzwerk für Aufgaben wie die Registrierung und den initialen Aufbau von Datenverbindungen zum Internet, sowie zum IMS Sprachnetzwerk. Außerdem informiert ein Endgerät die AMF über den Wechsel in eine neue Tracking Area (TA), damit es später im RRC-Idle State wieder vom Netzwerk gefunden werden kann. Dies wird als Connection Management (CM) bezeichnet.

Nachdem ein Endgerät logisch mit dem Netzwerk verbunden ist und ihm eine IP Adresse für die Kommunikation mit dem Internet zugewiesen wurde, können die Radioverbindung und die Verbindung zum Kernnetz zwischen unterschiedlichen Zuständen wechseln. Wie in LTE kann sich ein Endgerät im CM-Connected+RRC-Connected Zustand befinden, sowie im CM-IDLE+RRM-IDLE Zustand. In der Architektur des 5G Kernnetzes gibt es zusätzlich noch den CM-Connected+RRC Inactive Zustand. Das bedeutet, dass der 5G Kern einen Ende-zu-Ende Tunnel für die Nutzdaten aufrechterhält, während die Verbindung auf der Luftschnittstelle temporär abgebaut wird. Diese Kombination ist eine Lehre die aus LTE gezogen wurde, da hier Endgeräte heute sehr oft zwischen den RRC-Connected und RRC-IDLE Zuständen wechseln. Viele Applikationen halten heute jedoch eine Verbindung zu einem Server offen, und senden oder empfangen im Abstand von wenigen Minuten Datenpakete, um die TCP Verbindung aufrecht zu erhalten. Somit wird auch ständig die Verbindung ins Kernnetz auf- und abgebaut. Wenn ein gNB zum 5GC verbunden ist, kann dieser Aufwand reduziert werden, falls die AMF und der gNB die nötigen Funktionen implementiert haben, die Nutzdatenverbindung zwischen RAN und 5GC auch im RRC-IDLE State eines Endgerätes beizubehalten.

2.8.6 Registration Management Prozeduren

Wenn ein Gerät eingeschaltet wird, oder der Flugmodus deaktiviert wird, verbindet es sich zunächst mit der AMF, um eine gegenseitige Authentifizierungsprozedur durchzuführen. Danach folgt üblicherweise die Aktivierung der Verschlüsselung und Integritätsprüfung zwischen Endgerät und AMF, um den weiteren Austausch von Signalisierungsnachrichten zu schützen. Nachdem ein Endgerät registriert ist, möchte es dann typischerweise eine Internetverbindung aufbauen. Dies wird weiter unten noch genauer beschrieben. Bei LTE wurden diese zwei Prozeduren noch zusammen ausgeführt, d. h. eine Registrierung beinhaltete auch den Verbindungsaufbau zu einem externen Netzwerk. In 5G wurden diese zwei Aspekte jedoch wieder getrennt. Ein Grund dafür könnte sein, dass im 5G Kernnetz die Access Management Function (AMF) und die Session Management Function (SMF) klar voneinander getrennte Funktionen sind.

Abb. 2.20 zeigt die wichtigsten Schritte der Registrierungsprozedur. Diese startet durch Aufbau einer RRC Verbindung zum gNB und einer Connection Management (CM) Verbindung mit der AMF. Über diese wird dann die Registrierungsnachricht an

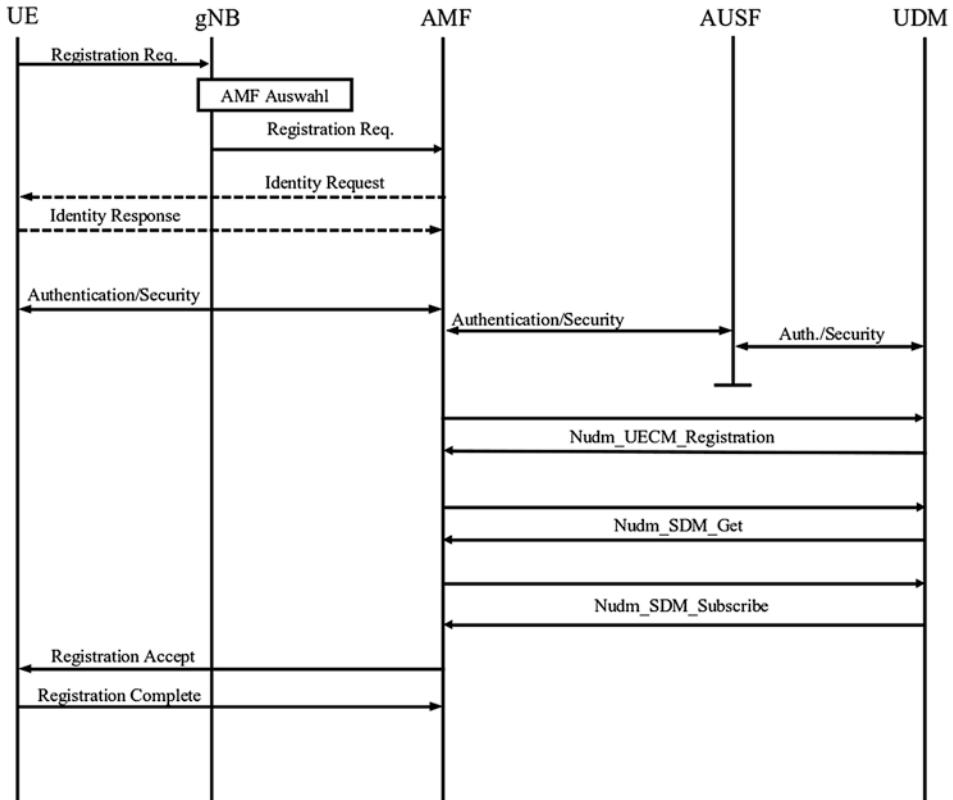


Abb. 2.20 Registrierungsprozedur zwischen Endgerät und AMF

die AMF gesendet. Falls das Endgerät schon zuvor mit dem Netzwerk verbunden war, nutzt es die 5G S-TMSI (Serving-Temporary Mobile Subscriber Identity), die zuvor vom Netzwerk zugewiesen wurde. Diese ID ist eine verkürzte Form der 5G GUTI (Globally Unique Temporary Identity) und enthält die Information, mit welcher AMF das Endgerät in der Vergangenheit verbunden war. Der gNB nutzt diese Information, um die Registration Request Nachricht an diese AMF zu senden. Dies ermöglicht es, einige Schritte der Registrierungsprozedur zu überspringen. Ist die ID unbekannt, wird die Nachricht an die Default AMF weitergeleitet.

Ist der AMF die temporäre ID des Endgeräts unbekannt, sendet sie bei Verbindungsauftnahme ein Identity Request an das Endgerät, um dessen Subscriber Concealed Identity (SUCI) zu erfahren. Die SUCI ist die verschlüsselte Variante der SUPU (Subscriber Permanent Identifier) und wurde mit dem Public Key des Netzwerkes verschlüsselt. Dies wird im Detail später beschrieben. Das Verschlüsseln ist an dieser Stelle notwendig, da der Datenaustausch mit der AMF an diesem Punkt noch nicht geschützt ist.

Im nächsten Schritt folgt eine Authentifizierungsprozedur, die ebenfalls nachfolgend noch genauer beschrieben wird. Nach erfolgreicher Authentifizierung wird dann die Verschlüsselung und die Integritätsprüfung jedes Datenpakets, das zwischen den Endgerät und der AMF ausgetauscht wird, aktiviert.

Danach registriert sich die AMF beim UDM mit einer Nudm_UECM_Registration Nachricht als für das Endgerät zuständig. Der Name der Nachricht enthält den angesprochenen Service (UDM) und die durchzuführende Aktion, also ein UE Connection Management Registration. In einer darauffolgenden Aktion fragt die AMF dann nach den Mobility Subscription Informationen des Endgeräts mit einem Nudm_SDM_Get und registriert sich schließlich für Benachrichtigungen bei Änderungen an den Daten des Nutzers in der Datenbank mit einem Nudm_SDM_Subscribe Request. Im letzten Schritt der Prozedur schließt die AMF dann die Registrierungsprozedur mit dem Endgerät ab.

Während diesen essentiellen Schritten können auch noch eine Anzahl an weiteren optionalen Operationen ausgeführt werden. Die AMF kann beispielsweise nach dem Permanent Equipment Identifier (PEI) des Endgeräts fragen. In manchen Ländern wird die PEI mit einer Liste von Geräten verglichen, die als gestohlen gemeldet wurden und die Registrierung dann verweigert. Auch ist es möglich, nur Geräte im Netzwerk zuzulassen, die zuvor registriert worden sind. In Europa ist dies jedoch nicht üblich. Die meisten Netzbetreiber verwenden die PEI Abfrage jedoch, um eine Liste an Endgerätemodellen zu pflegen, die im Netzwerk aktuell verwendet werden. Diese wird dann für statistische Zwecke und vor allem bei der Fehleranalyse verwendet.

2.8.7 Session Management

Nachdem ein Endgerät im Netzwerk registriert ist, fordert es dann eine oder mehrere Datenverbindungen (Data Sessions) an. Eine Session wird normalerweise für die Internetverbindung verwendet, eine Zweite für eine Verbindung zum IMS Telefoniesystem. Auf Applikationsebene bedeutet der Aufbau einer Datenverbindung den Bezug einer IP Adresse, eines IPv6 Prefix, oder beides.

Aus Endgerätesicht bedeutet der Aufbau einer Session, einen Tunnel aufzubauen, in dem dann IP Pakete zwischen dem Endgerät, dem Zugangsnetz und dem Kernnetz übertragen werden können. Tunnel sind notwendig, da ein Nutzer seinen Standort zu jeder Zeit beliebig ändern kann, und somit die Verbindung von einem gNB zum nächsten gNB weitergegeben wird. Würde die IP Adresse des Nutzers für das Routing der Datenpakete verwendet werden, müssten beim Wechsel des gNB die Routingtabellen aller Router im RAN und Kernnetz angepasst werden. Wie in Abb. 2.21 gezeigt, kann dies vermieden werden, wenn die IP Nutzdatenpakete durch einen IP Tunnel geleitet werden, und für die Endpunkte des Tunnels die IP Adressen des gNBs und der UPF verwendet werden. Während die IP Adressen der Nutzdatenpakete innerhalb des Tunnels unverändert bleiben, ändert sich die IP Adresse des Tunnelendpunktes, sobald ein neuer gNB für

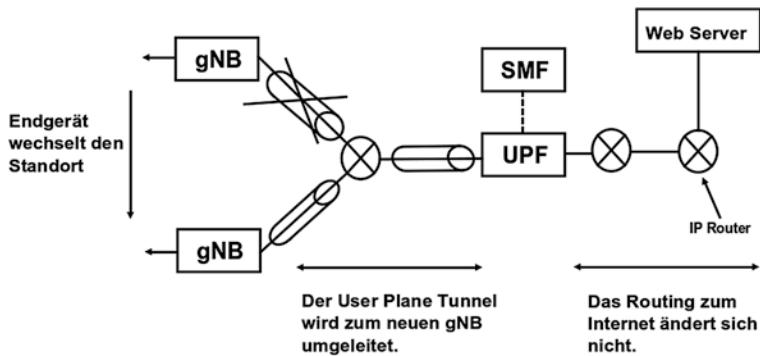


Abb. 2.21 GTP Tunnel im 5G Kernnetz

die Verbindung zuständig wird. Das bedeutet, dass statt dem Update vieler IP Routing Tabellen nur die IP Adresse des Endpunktes des Tunnels geändert werden muss.

Während sich die IP Adresse des gNB, die außerhalb des Tunnels verwendet wird, ständig ändert, bleibt die selbe UPF im Pfad des Tunnels. Dies ist nötig, da die UPF der Anker zum Internet ist. Alle IP Pakete, die aus dem Internet an das Endgerät gesendet werden, werden somit immer an diese UPF gesendet und von dort durch den User Plane Tunnel zum aktuellen gNB.

Jeder User Plane Tunnel hat seine eigene ID im Pakethader. Diese ermöglicht es den gNBs und der UPF, die darin übertragenen Nutzdatenpakete einem Endgerät zuzuordnen. Das Protokoll für den Aufbau des Tunnels und die Übertragung von IP Paketen innerhalb des Tunnels nennt sich GPRS Tunneling Protocol (GTP). Ursprünglich in den 1990er Jahren für die paketorientierte Erweiterung des 2G GSM Standards spezifiziert, wurde es seither mit nur wenigen Änderungen auch in allen darauffolgenden Netzgenerationen verwendet.

Der Session (Tunnel) Aufbau wird von der Session Management Function (SMF) gesteuert. Das Endgerät kommuniziert mit dieser jedoch nicht direkt. Stattdessen müssen alle Nachrichten durch die AMF geleitet werden, da diese den Security Context terminiert, also der Endpunkt für die Verschlüsselung der Verbindung zum Endgerät ist.

Der Aufbau einer Session wird durch das Endgerät mit einer PDU Session Establishment Request Nachricht angestoßen, die von der AMF an eine SMF weitergegeben wird. Der wichtigste Parameter in dieser Nachricht ist der Data Network Name (DNN). Dieser ist im Endgerät konfiguriert, und bestimmt, mit welchem externen Netzwerk eine Verbindung aufgebaut werden soll. In allen bisherigen Netzgenerationen wurde dieser Parameter als Access Point Name (APN) bezeichnet. Android und iOS Endgeräte haben heute eine Liste der DNNs bzw. APNs aller Netzwerkbetreiber und konfigurieren sich somit automatisch. Es ist jedoch als Benutzer auch möglich, den DNN manuell zu konfigurieren.

Nachdem die AMF die Nachricht erhalten hat, wählt sie eine Session Management Function (SMF) aus, die für diesen Nutzer und für das angefragte externe Netzwerk verwendet werden kann, und leitet die Nachricht dann an diese weiter. Die SMF stellt dann eine Anfrage an die UDM, um die Teilnehmerdaten des Benutzers aus der Datenbank zu erhalten. Außerdem registriert sich die SMF auch für Push Benachrichtigungen bei Änderung der Daten des Teilnehmers in der Datenbank. Daraufhin bestätigt sie der AMF mit einer Session Create Response Nachricht, dass sie den Teilnehmer verwaltet wird, und kontaktiert die User Plane Function (UPF) über die N4 Schnittstelle, um den Aufbau einer User Data Session zu beantragen. Wenn die UPF noch genug Ressourcen für diese Verbindung zur Verfügung hat, also z. B. genügend Rechenleistung und genügend freie Bandbreite zum externen Netzwerk sowie dem 5G Radionetzwerk, erzeugt sie einen Nutzdatentunnel, teilt diesem eine eindeutige ID zu und antwortet der SMF entsprechend. Die SMF sendet dann eine Namf_N1N2Message Transfer Nachricht an den gNB über die logische N2 Schnittstelle, um diesen über den Aufbau des Nutzdatentunnels mit der UPF zu informieren. Die AMF übersetzt diese Nachricht in eine N2 PDU Session Request Nachricht und eine PDU Session Establishment Accept Nachricht gegenüber dem Endgerät. Der gNB leitet dann die Endgeräte spezifische PDU Session Establishment Accept Nachricht über die logische N1 Schnittstelle an das Endgerät weiter und baut einen neuen Radio Bearer auf der Luftschnittstelle auf. Das Endgerät wird über den neuen Radio Bearer über eine RRC Reconfiguration Nachricht informiert.

Der gNB bestätigt dann den Bearer Aufbau über das N2 Interface zur SMF, die diese Nachricht dann an die UPF in einer Session Modification Request-Response Nachricht über das N4 Interface weiterleitet. Ab diesem Zeitpunkt können dann IPv4 Pakete mit dem externen Netzwerk ausgetauscht werden. Falls eine IPv6 oder IPv4v6 Session beantragt wurde, erzeugt die SMF außerdem eine IPv6 Address Configuration Nachricht und sendet diese über die UPF zum Endgerät. Die Nachricht enthält eine IPv6 Router Advertisement (RA) Nachricht, die das IPv6 Prefix für das Endgerät enthält. Mit dem IPv6 Prefix erzeugt das Endgerät dann, wie im LTE Kapitel beschrieben, eine IPv6 Adresse.

An dieser Stelle sei angemerkt, dass Abb. 2.22 nur den Nachrichtenaustausch zwischen den grundsätzlichen Funktionen zeigt, die für Aufbau einer Session notwendig sind. Nicht gezeigt sind:

- Die Interaktion mit der Policy Control Function (PCF) für die Reservierung von Bandbreite auf der User Plane, der Aktivierung von Paket Filtern und der Limitierung der Datenrate.
- Die Kommunikation zwischen der PCF und der SMF mit der Charging Function (CHF) für Abrechnungszwecke. Dies beinhaltet auch Funktionalitäten wie z. B. die Geschwindigkeitsreduktion, wenn Nutzer ihr monatliches Datenlimit erreicht haben. Weitere Details zum Thema 5G Charging sind in 3GPP TS 29.513²⁹ zu finden.

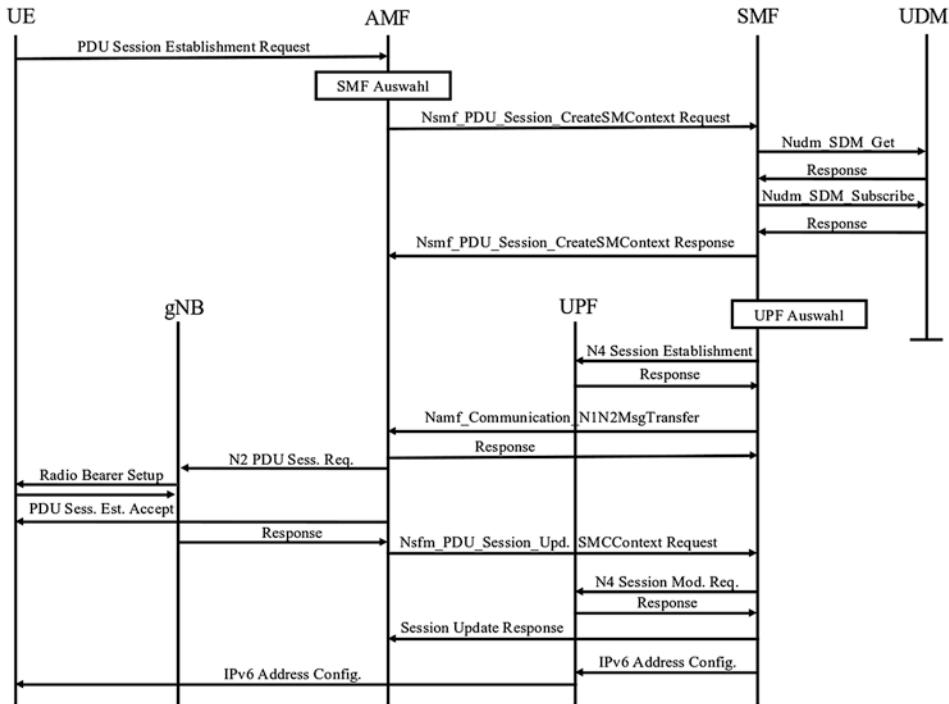


Abb. 2.22 Aufbau einer Session

- Der Austausch von Nutzdatenpaketen zwischen Endgerät, dem gNB, der UPF und dem externen Netzwerk. Dieser Datenfluss ist für die AMF und SMF transparent, da die Nutzdaten nur über die Luftschnittstelle, über das N2 Interface zwischen gNB und UPF und dem Interface zum externen Netzwerk übertragen werden.
- Die Struktur des darunterliegenden IP Netzwerkes, also IP Router im Signalisierungs- und Datenpfad.

Im Normalfall verbindet der Session Establishment Prozess das Endgerät mit dem Internet, entweder nachdem es eingeschaltet wurde, oder nachdem der Flugmodus deaktiviert wurde. Am Ende der Prozedur verfügt das Endgerät über eine IPv4 Adresse, über DNS Server Adressen für die Umwandlung von Domainnamen in IP Adressen, sowie über ein IPv6 Prefix, zusammen mit den IPv6 Adressen der DNS Server. Der Internetzugang ist jedoch nicht die einzige Verbindung, die mit einem Session Establishment aufgebaut werden kann. Für den IMS Telefoniedienst wird in gleicher Weise verfahren, das Endgerät muss jedoch noch zusätzlich über die IP Adresse der P-CSCF informiert werden. Details hierzu finden sich im Kapitel über VoLTE. Die P-CSCF Adresse(n) werden dem Endgerät in der PDU Session Establishment Response Nachricht zusammen mit den

DNS Server Adressen und den bereits oben beschriebenen Konfigurationsparametern übermittelt. Zusätzlich erhalten die IP Pakete mit Sprachdaten eine eigene Quality of Service Klasse. Dies stellt sicher, dass diese auf allen Schnittstellen im Netzwerk, inklusive der Luftschnittstelle, bevorzugt behandelt werden. In 2G und 3G Netzwerken wurde dazu der Primary PDP Kontext mit einem Secondary PDP Kontext erweitert. In LTE wird der IMS Default Bearer durch Aufbau eines Dedicated Bearers modifiziert. Im 5G Kernnetz werden zu diesem Zweck sogenannte Packet Filter Sets (PFS) verwendet. Trotz der unterschiedlichen Namen haben alle Prozeduren das Ziel, dem Endgerät und allen Routern im Übertragungspfad eine Liste von Paketfilterregeln zu senden. Diese geben vor, dass Pakete von und zu bestimmten IP Adressen und von und zu bestimmten TCP/UDP Ports zu bevorzugen sind. Außerdem können über diese Regeln auch eine maximale Datenrate für den Sprachdatenstrom festgelegt werden. Wichtig ist an dieser Stelle, dass das Endgerät ein Packet Filter Set per N1 Signalisierung und nicht über den Nutzdatenstrom bekommt.

2.8.8 Mobility Management

Ist ein Endgerät mit einem gNB verbunden und somit im RRC-Connected State, werden während des Verbindungsaufbaus auch Kanalmessungen konfiguriert. Wenn das Endgerät dann eine bessere Nachbarzelle findet, sendet es einen Measurement Report. Der (Source) gNB kontaktiert dann den benachbarten (Destination) gNB, der entweder am gleichen Standort zu finden ist, falls das Endgerät einen anderen Sektor besser bewertet, oder auch an einem Nachbarstandort. In beiden Fällen wird die Xn Schnittstelle für die Signalisierungsverbindung verwendet. Um einen Handover so schnell wie möglich auszuführen, wird der Endpunkt des Nutzdatentunnels zum Kernnetzwerk zunächst beim Source gNB belassen. Alle Daten von der UPF werden nach dem Handover dann für eine kurze Zeit über die Xn Schnittstelle zum Destination gNB weitergeleitet. In umgekehrter Richtung leitet der Destination gNB zunächst alle ankommenden Datenpakete des Endgeräts an den Source gNB weiter, der diese dann zur UPF sendet. Aus Sicht des Kernnetz ändert sich somit zunächst nichts. Nachdem der Handover auf der Luftschnittstelle erfolgreich durchgeführt wurde, startet der Destination gNB dann den letzten Teil der Handoverprozedur und modifiziert auch den Endpunkt des Nutzdatentunnels zur UPF.

Wie in Abb. 2.23 gezeigt, startet dieser letzte Teil der Handoverprozedur, sobald der Destination gNB über die N2 Schnittstelle eine Path Switch Request Nachricht an die AMF sendet. In dieser Nachricht wird der Nutzdatentunnel des Teilnehmers identifiziert und die AMF bekommt die neue IP und Port Adresse, an der dieser Nutzdatentunnel nun enden soll. Die AMF bearbeitet dann die Anfrage und leitet diese in einer Nsmf_PDUSession_UpdateSMContext Nachricht an die Session Management Funktion weiter. Von dort wird die Anfrage in einer N4SessionModification Request Nachricht an die UPF weitergeleitet. Diese ändert dann den Endpunkt des RAN Tunnels zum Destination gNB und antwortet der SMF. Daraufhin wird ein Ende Marker über das N3 Interface

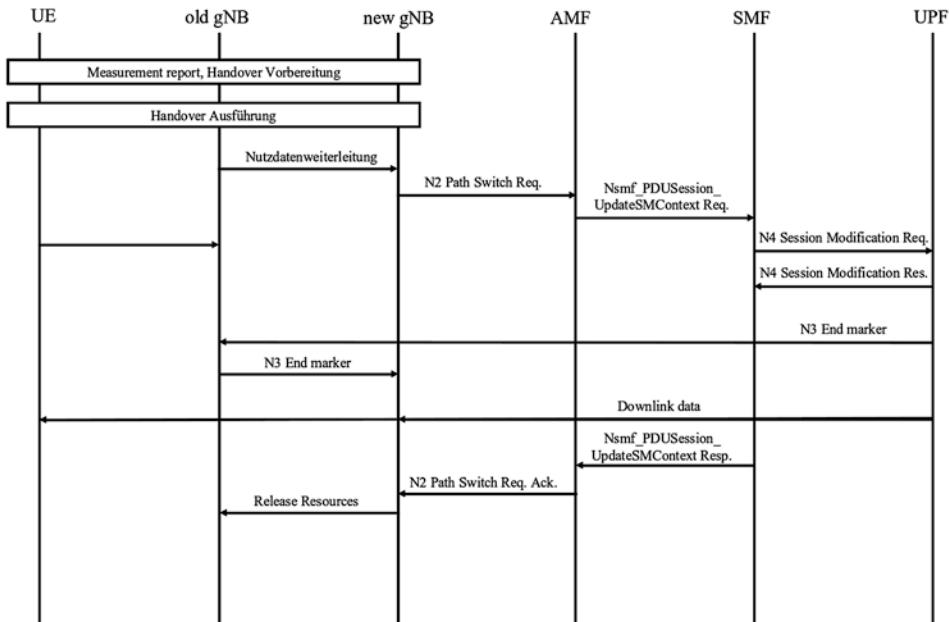


Abb. 2.23 Ein 5G Handover über die Xn Schnittstelle zwischen zwei gNBs

an den Source gNB gesendet. Dieser weiß somit, dass keine weiteren Nutzdaten vom Kernnetz für diese Verbindung mehr eintreffen werden. Dies ist wichtig, da bis zu diesem Zeitpunkt noch neue Daten ankommen konnten, die dann weitergeleitet werden mussten.

Auf der Signaling Plane endet die Prozedur durch eine Antwort der SMF an die AMF, die dieser dann an den Destination gNB weiterleitet. Der Destination gNB löscht daraufhin die temporäre Nutzdatenverbindung auf der Xn Schnittstelle durch das Senden einer Release Resources Nachricht.

Während die zuvor beschriebenen Registration-, Session- und Mobility Management Operationen einen guten Überblick über die wichtigsten Operationen in einem Mobilfunkkernnetz geben, gibt es viele weitere Prozeduren, die hier nicht beschrieben sind. Dazu zählen z. B. Connection-, Registration- und Session Release Prozeduren, Security Prozeduren, weitere Arten von Handoverprozeduren, Connection Inactivity Prozeduren, Resume Prozeduren, uvm. Details dazu sind in 3GPP TS 23.502 zu finden.

2.8.9 Neue Sicherheitsfunktionen

Durch immer leistungsfähigere Netzwerkkomponenten und der daraus resultierenden Weiterentwicklung werden auch immer wieder neue Schwachstellen in Netzwerken

gefunden, die Sicherheit und Privatsphäre gefährden. Es ist somit ständig notwendig, Netzwerk und Endgeräte entsprechend weiterzuentwickeln. Wie auch in früheren Netzgenerationen wurden von 3GPP auch in 5G neue Sicherheitsverfahren, wie in TS 33.501³⁰ beschrieben, eingeführt. Die wichtigsten Änderungen gegenüber dem 4G Kernnetz, vor allem aus Endgerätesicht, werden nun nachfolgend beschrieben.

Ein großes Defizit von früheren Netzgenerationen ist, dass das Endgerät seine International Mobile Subscriber Identity (IMSI), die auf der SIM Karte gespeichert ist, an das Netzwerk senden muss, falls die zuvor vergebene temporäre ID nicht mehr gültig ist. Dies ist zum Beispiel der Fall, wenn die SIM Karte in ein anderes Gerät eingelegt wird, wenn sich das Gerät zum ersten Mal in einem Roaming Netzwerk einbucht, oder wenn das Endgerät versucht, sich mit einer ‚gefälschten‘ Basisstation eines Angreifers zu verbinden. Solche Basisstationen werden auch als IMSI Catcher bezeichnet. Dieser fordert alle Endgeräte die sich einbuchen wollen auf, ihre IMSI zu senden. Auf diese Weise ist es dem IMSI Catcher möglich, die Teilnehmerkennungen von allen Geräten in dessen Umkreis zu erfahren. Ist die IMSI einer Zielperson bekannt, ist es somit möglich herauszufinden, ob diese sich in der Nähe befindet.

Eine weitere Schwachstelle des Authentifizierungsprozesses war bisher, dass Angreifer, die Zugriff auf die Signalisierung des weltweiten Roaming Netzwerks haben, an die temporären Schlüssel für das Ciphering mit dem Endgerät gelangen können. Somit konnten Angreifer dann eine Zielperson abhören, oder sich selber für einen Teilnehmer ausgeben. Für das 5G Kernnetz hat 3GPP deswegen Verfahren entwickelt, mit denen es nicht mehr notwendig ist, die IMSI des Teilnehmers, die jetzt als SUPI (Subscriber Permanent Identity) bezeichnet wird, im Klartext zu übertragen. Das Prinzip dieses Ansatzes ist wie folgt:

Um die SUPI zu verschleiern, bevor sie über die Luftschnittstelle, die Backhaul Verbindung, oder auch durch ein Roaming Netzwerk gesendet wird, erzeugt das Endgerät eine verschlüsselte Version, die als SUCI (Subscriber Concealed Identity) bezeichnet wird. Dazu generiert das Endgerät zunächst ein temporäres (Ephemeral) Public/Private Schlüsselpaar zur einmaligen Verwendung. Daraus wird dann ein weiter Schlüssel zusammen mit dem Public Key des Heimnetzes erzeugt, der ebenfalls auf der SIM Karte gespeichert ist. Mit diesem Schlüssel wird dann die IMSI in die SUCI verschlüsselt. Der Mobile Country Code (MCC) und der Mobile Network Code (MNC), die Teil der IMSI sind, werden nicht verschlüsselt. Diese werden benötigt, um in Roaming Netzwerken das Heimatnetzwerk des Teilnehmers zu finden. Nachdem die SUCI generiert wurde, wird sie zusammen mit dem Ephemeral Public Key an das Netzwerk geschickt. Auf der Netzwerkseite wird dann die SUPI aus der empfangenen SUCI zusammen mit dem Ephemeral Public Schlüssel des Teilnehmers und dem privaten Schlüssel des Netzbetreibers rekonstruiert. Da für jede Übertragung der SUCI ein neues Public/Private Schlüsselpaar vom Endgerät generiert wird, ist sichergestellt, dass ein Angreifer den Teilnehmer nicht über die SUCI nachverfolgen kann.

Wichtig ist an dieser Stelle noch zu erwähnen, dass die Verschleierung der SUPI den Public Key des Heimnetzbetreibers benötigt, der auf der SIM Karte gespeichert sein

muss. Dazu werden entweder neue SIM Karten benötigt, oder bestehende SIM Karten müssen mit einem Over-the-Air (OTA) Update entsprechend erweitert werden. Dieser Prozess wird schon seit langem verwendet, um z. B. während eines Auslandsaufenthaltes per SMS eine aktuelle Liste von präferierten Roamingnetzwerken auf die SIM Karte zu schreiben. Zusätzlich zu den neuen Informationen auf der SIM Karte werden auch noch neue Algorithmen auf der Endgeräteseite benötigt, um die SUCI und die Ciphering Keys zu berechnen. Diese können entweder auf der SIM Karte implementiert werden oder im Endgerät. Ein Feld auf der SIM Karte bestimmt dabei, ob der Algorithmus in der SIM Karte oder im Endgerät ausgeführt werden soll.

Um rückwärtskompatibel zu sein, ist es auch möglich, auf das 5G Netzwerk mit SIM Karten zuzugreifen, die keinen Public Key eines 5G Netzwerkes und andere 5G Parameter gespeichert haben. In diesem Fall wird eine ‚Null‘ Verschlüsselung für die Erzeugung der SUCI verwendet. Das bedeutet, dass die SUPI nicht verschlüsselt wird³¹.

Die folgende Liste gibt einen Überblick, welche Möglichkeiten es für die Verschlüsselung der SUPI in Abhängigkeit der verwendeten SIM Karte gibt:

- Alte 3G/4G SIM Karten, mit denen noch kein Update durchgeführt wurde: Zugang zum 5G Kernnetzwerk ist möglich, aber Funktionen wie die Verschleierung der SUPI können nicht verwendet werden.
- Alte 3G/4G SIM Karten, die mit einem Over-the-Air Update den Public Key des Heimatnetzwerkes und andere Parameter erhalten haben: Das verschleiern der SUPI ist möglich und wird im Endgerät durchgeführt.
- Neue 5G SIM Karten: Das Erzeugen der SUCI sowie der Session Keys für die Verschlüsselung der gesamten Kommunikation kann auf der SIM Karte erfolgen.

Zusätzlich zur Authentifizierung des Endgerätes durch das Netzwerk, authentifiziert das Endgerät im Gegenzug auch das Netzwerk. Diese Funktionalität wurde schon mit 3G UMTS eingeführt und wurde auch im LTE Kernnetz weitergeführt. Ein Nachteil in früheren Netzen war jedoch, dass die Authentifizierungsprozedur während des Roamings komplett im besuchten Netzwerk durchgeführt wurde. Dazu war es notwendig, dass das Heimnetz alle Authentifizierungsparameter und Schlüssel an das besuchte Netzwerk sendet, ohne zu wissen, ob sich der Teilnehmer dort auch tatsächlich aufhält. In 5G Kernnetzwerken ist dies nicht mehr der Fall, hier wird die Authentifizierungsprozedur immer zwischen dem Endgerät und dem Heimatnetzwerk durchgeführt. Damit ist sicher gestellt, dass das Visited Network erst die Schlüssel für das Ciphering erhält, wenn sichergestellt ist, dass sich ein Endgerät auch tatsächlich dort befindet. Zusätzlich zu den Schlüsseln erhält das besuchte Netzwerk auch die SUPI erst nach der erfolgreichen Authentifizierung.

In bisherigen Netzarchitekturen wurde die IMSI auch für die Benachrichtigung der Endgeräte in Paging Nachrichten verwendet, falls ein Endgerät sich nicht auf eine Paging Nachricht mit seiner temporären ID meldet. In 5G Netzwerken ist dies nun nicht länger erlaubt.

Mit einer weiteren Neuerung, die mit dem 5G Kernnetz eingeführt wurde, gibt es nun auch mehrere Authentifizierungsmöglichkeiten. Eine davon ist die 5G AKA Authentifizierung, eine Weiterentwicklung der LTE Authentifizierungsprozedur. Diese wird für Endgeräte mit SIM Karte verwendet. Standardisiert wurden jedoch nun auch EAP-basierte Authentifikationsprozeduren für Geräte ohne SIM Karten mit Pre-Shared Keys sowie Benutzername/Password Authentifizierung. Diese könnten z. B. für Fabriknetzwerke (Campus Networks) interessant werden.

Um die Sicherheit im internationalen Roaming weiter zu verbessern, wurde in den 5G Spezifikation der Security Edge Protection Proxy (SEPP) eingeführt. SEPPs dienen als Gateway zu anderen 5G Netzwerken und müssen sich zunächst gegenseitig authentifizieren, bevor Signalisierungsnachrichten ausgetauscht werden können. Zusätzlich verbergen sie die Topologie des dahinterliegenden Netzwerkes. Authentifizierungsinformation der Teilnehmer, die in diversen Nachrichten enthalten sind, werden verschlüsselt (confidentiality protected), damit diese von keinem Netzelement im IP Exchange Network (IPX) zwischen den Netzbetreibern mitgelesen werden können. Andere Teile der Nachrichten werden mit einer Checksumme versehen (integrity protected) und können somit nur von autorisierten IPX Knoten modifiziert werden.

2.8.10 Der 5G Kern und unterschiedliche RAN Optionen

In den vorangegangenen Abschnitten wurde davon ausgegangen, dass das 5G Kernnetz mit einem reinen 5G Zugangsnetzwerk verbunden ist. Dies wird auch als 5G New Radio Option 2 bezeichnet. Im Standard wurden zusätzlich weitere Optionen beschrieben, wie eine Kombination von 5G NR und LTE Zellen an ein 5G Kernnetz angeschlossen werden können. In der Praxis kommen diese heute jedoch nicht zum Einsatz.

In den nächsten Jahren werden in den meisten Netzwerken ein 4G LTE- und ein 5G-Kernnetz parallel betrieben werden. LTE-only Endgeräte werden das LTE Radio Netzwerk und das LTE Kernnetzwerk verwenden. 5G Option 3 Non-Standalone (NSA) Endgeräte werden das LTE Radionetzwerk in Kombination mit der 5G Luftschnittstelle als Speed Booster zusammen mit dem LTE Kernnetzwerk verwenden. Neuere Geräte, die 5G Standalone (SA) Option 2 beherrschen, können zusätzlich in manchen Gebieten auch nur die 5G Luftschnittstelle im Zusammenspiel mit dem 5G Kernnetzwerk verwenden. Je nach Verfügbarkeit von 5G Option 2 an einem Ort wechseln solche Endgeräte dann zwischen dem 4G und 5G Kernnetzwerk.

2.8.11 Zusammenspiel der 5G und 4G Kernnetzwerke

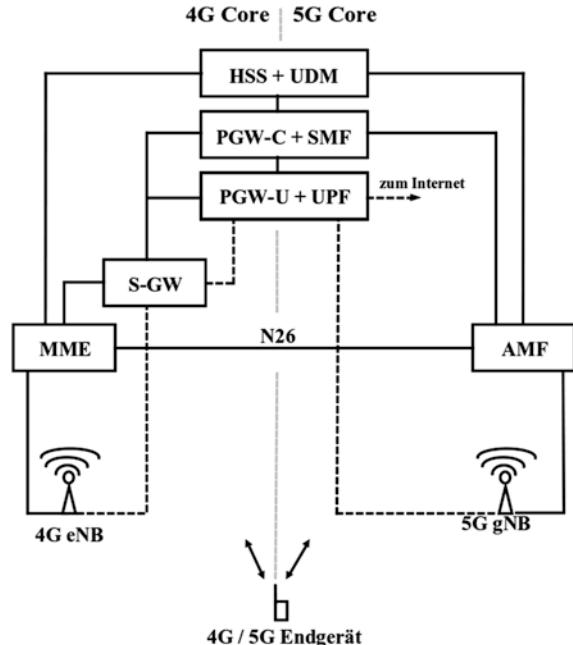
Erreicht das Endgerät die Grenze der 5G NR Netzabdeckung, oder verliert es das Signal komplett, werden Prozeduren eingeleitet, um in eine andere Radionetztechnologie zu wechseln. Dies ist notwendig, da in der Praxis heute die 5G NR Standalone

Netzabdeckung kleiner als die LTE Netzabdeckung ist. Das ist zum Beispiel der Fall, wenn LTE einen niedrigeren Frequenzbereich als 5G NR verwendet oder falls an einem Standort gar kein 5G Kanal vorhanden ist. Auch für den Fall, dass 5G NR an einem Standort vorhanden ist, jedoch nur mit einer im Vergleich zu LTE geringerer Bandbreite, macht es Sinn, nach LTE bzw. 5G NR Option 3 mit einer LTE Ankerzelle zu wechseln. Erreicht ein Endgerät dann wieder ein Gebiet in dem 5G SA vorhanden ist, müssen ebenfalls Prozeduren vorhanden sein, damit das Endgerät wieder dorthin wechseln kann. 3GPP Release 15 definiert den Wechsel (Mobility) zwischen 5G NR und LTE Kernnetzwerk in TS 23.501, Abschn. 4.3³², sowie die dafür notwendigen Prozeduren in TS 23.502, Abschn. 4.11.1³³.

Der Wechsel zu und von GSM und UMTS wurde bewusst nicht standardisiert. Hier wurde angenommen, dass beim Rollout von 5G NR schon eine großflächige und dichte LTE Abdeckung besteht und somit als alleinige Grundlage dienen kann. Nachdem ein Endgerät mit dem LTE Kernnetz verbunden ist, können jedoch weitere Handover nach GSM oder auch UMTS, falls noch vorhanden, durchgeführt werden. Somit muss das 5G Kernnetz nur Prozeduren beherrschen, um mit einem 4G Kernnetzwerk zu kommunizieren.

Abb. 2.24 zeigt, wie die 4G und 5G Kernnetzwerke miteinander verbunden werden, um einen Wechsel der Radionetzwerke zu ermöglichen (inter-RAT Mobility). Grundlage dafür sind Netzkomponenten, bzw. Funktionen, die sowohl 4G, als auch 5G Prozeduren beherrschen:

Abb. 2.24 Verbindung zwischen LTE und 5G Kernnetz für inter-RAT Mobility



- Die Teilnehmerdatenbank: Eine gemeinsame Datenbank ist nötig, damit sich ein Teilnehmer im 4G und 5G Netzwerk einbuchen kann. Deshalb müssen die 4G Home Subscriber Server (HSS) Datenbank und die 5G Unified Data Management (UDM) Function entweder über eine nicht spezifizierte Schnittstelle miteinander kommunizieren oder gemeinsam implementiert sein.
- Der Mobility Anker für eine Verbindung: In 4G endet der Bearer eines Teilnehmers am PDN-GW, der auch als PGW abgekürzt wird. Das bedeutet, dass der PGW das einzige Netzwerkelement ist, das sich niemals ändert, obwohl sich der Teilnehmer beliebig durch das Netzwerk bewegen kann. Dies ist notwendig, damit das Endgerät immer unter der gleichen IP Adresse aus dem Internet erreicht werden kann und TCP Verbindungen nicht abbrechen. Auf der 5G Seite wird diese Aufgabe von der User Plane Function (UPF) wahrgenommen. Kontrolle über diese Verbindung hat die Session Management Function (SMF). Das bedeutet, dass für den Wechsel zwischen 4G und 5G Kernnetzwerk der Kontrollteil des 4G PGW mit der 5G SMF kombiniert werden muss. Für die User Plane muss der 4G PGW mit der 5G UPF kombiniert werden.

Zusätzlich zu diesen kombinierten Funktionen wird für einen schnellen Wechsel des Kernnetzes noch die N26 Schnittstelle zwischen der 4G Mobility Management Entity (MME) und der 5G Access Management Function (AMF) benötigt. Ob N26 im Netzwerk vorhanden ist, teilt die AMF dem Endgerät in der Registration Accept Nachricht im Parameter mit dem Namen IKWN26 mit. Ist dieser Parameter auf 0 gesetzt, ist die N26 Schnittstelle vorhanden. Dies sollte in den meisten Netzen in der Praxis der Fall sein. Netze ohne N26 Schnittstelle zu LTE setzen den Parameter auf 1.

Im RRC-IDLE State, in dem keine Radioverbindung zum Netzwerk existiert, sucht das Endgerät selbstständig nach anderen Radiotechnologien. Um die Suche zu erleichtern, erhält es dazu die wichtigsten Parameter des LTE Netzwerkes über die 5G System Information Nachrichten. Wird das 5G Netzwerk dann zu schwach und hat das Endgerät ein 4G Netzwerk gefunden, führt es den Wechsel selbstständig durch. Wenn ein Endgerät im 4G RRC-IDLE Zustand dann in einen Bereich kommt, in dem auch ein 5G SA Radionetzwerk vorhanden ist, wechselt es wieder automatisch zurück. Obwohl diese Prozedur vom Endgerät autonom ausgeführt wird, kann das Netzwerk das Verhalten des Endgerätes über Parameter in den LTE System Information Nachrichten beeinflussen. Details zu den System Information Nachrichten folgen später.

Abb. 2.25 zeigt die wichtigsten Schritte eines Kontextransfers von 4G nach 5G aus dem RRC-IDLE Zustand wie in 3GPP TS 23.502, 4.11.1.3.3 beschrieben. Im ersten Schritt verbindet sich das Endgerät mit dem 5G Zugangsnetwork und startet eine NAS Registration Prozedur mit der AMF im Kernnetzwerk. Da dies keine neue Registrierung, sondern ein Kontext Transfer ist, wird das Registration Type Information Element in der Nachricht auf „Mobility Registration Update“ gesetzt. Außerdem wird der LTE Globally Unique Temporary Identifier (GUTI) in eine 5G-GUTI umgewandelt und dem Netzwerk mitgeteilt, dass das Endgerät aktuell im LTE Kernnetz registriert ist. Die

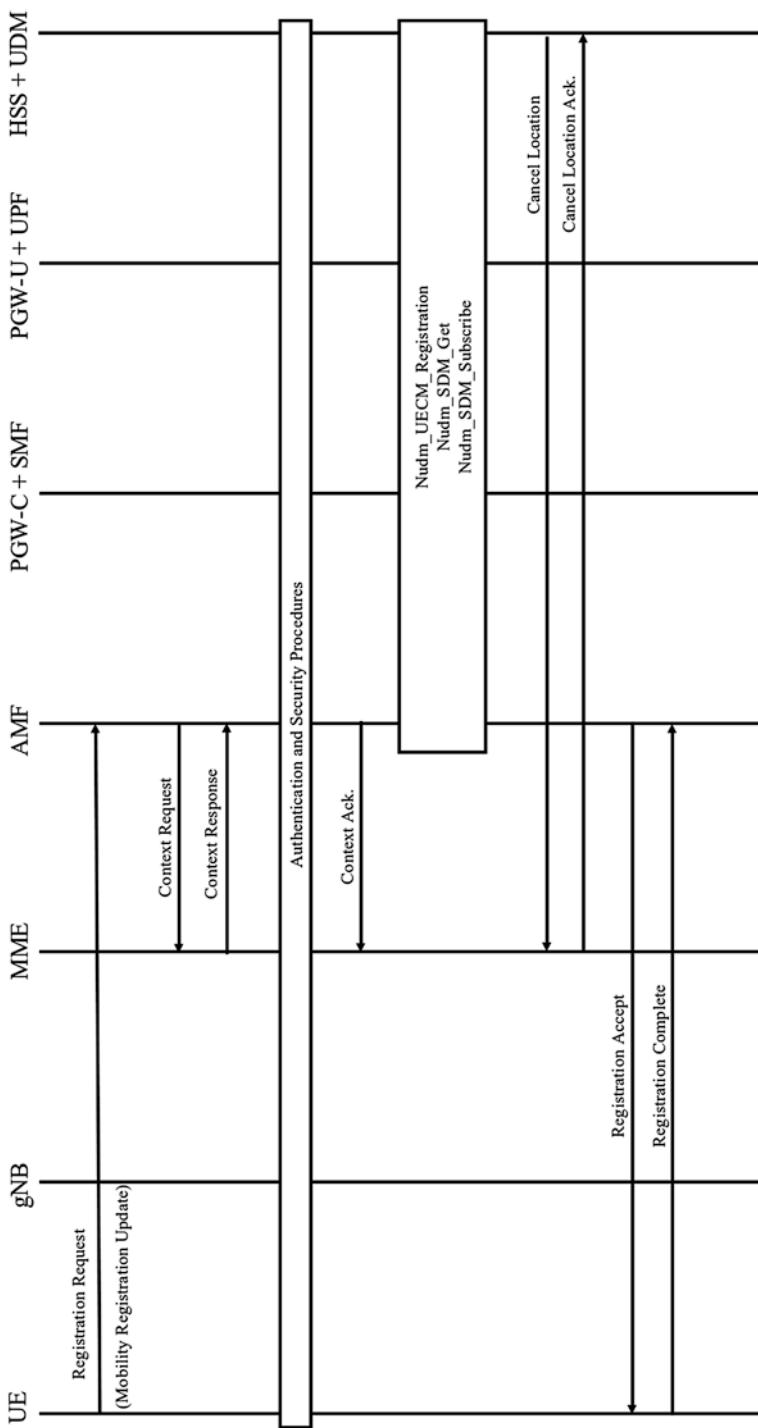


Abb. 2.25 Transfer des Endgerätekontext von 4G nach 5G aus dem RRC-Idle Zustand

GUTI identifiziert dabei nicht nur den Teilnehmer, sondern enthält auch Informationen, welche MME oder AMF das Endgerät zuvor verwendet hat. Die AMF kontaktiert daraufhin die LTE MME über die N26 Schnittstelle und fordert den Registrierungs- und Sessionzustand des Endgeräts bei der MME an. Mit den so erhaltenen Informationen können sich dann AMF und Endgerät gegenseitig authentifizieren und die Verschlüsselung kann aktiviert werden. Danach wird der Endgerätekontext vom kombinierten PGW-C+SMF auf die SMF Seite transferiert und alle bereits aufgebauten Bearer des Endgerätes werden vom kombinierten PGW-U+UPF mit dem 5G gNB verbunden. Der gNB baut dann die dazugehörigen Radio Bearer zum Endgerät auf. Die AMF kontaktiert außerdem die Teilnehmerdatenbank (HSS + UDM) und registriert sich als „Serving Function“. Außerdem registriert sie sich für Push Nachrichten für den Fall, dass sich die Daten des Nutzers in der Teilnehmerdatenbank ändern. Somit wird die MME ID in der Teilnehmerdatenbank gelöscht, die AMF ID hinzugefügt und eine NAS Cancel Location Nachricht wird zur MME geschickt. Diese kann dann den Kontext des Endgeräts in ihrer lokalen Datenbank löschen. Im letzten Schritt wird dann die ursprüngliche Registration Request Nachricht durch die AMF mit einer Registration Accept Nachricht bestätigt. Die Prozedur endet mit der vom Endgerät ausgehenden Registration Complete Nachricht zur AMF.

Kommt ein Endgerät an den Rand der 5G Abdeckung während einer Radioverbindung aufgebaut ist (RRC-Connected Zustand), gibt es mehrere Möglichkeiten das Radionetzwerk zu wechseln. Eine Möglichkeit für das Netzwerk ist, eine RRC Connection Release Nachricht an das Endgerät zu senden und das Endgerät mit dieser Nachricht aufzufordern, im RRC-IDle Zustand dann sofort in das LTE Netzwerk zu wechseln (Release with Redirect). Die Datenübertragung wird dadurch kurzzeitig unterbrochen, da das Endgerät zunächst eine 4G LTE Zelle suchen muss und dort dann eine Tracking Area Update Prozedur durchführt, um den Radiobearer erneut aufzubauen. Während eine kurze Unterbrechung für viele Applikationen unkritisch ist, ist diese Prozedur während der Nutzung von Real Time Applikation wie der Sprachtelefonie weniger gut geeignet. Eine bessere Variante ist deshalb der inter-RAT Handover von 5G nach 4G. Diese Prozedur ist deutlich komplexer als ein Release mit Redirect, da bestehende Sessions in Echtzeit umgeleitet werden müssen. Details zu dieser Prozedur, die nun nachfolgend beschrieben wird, finden sich in TS 23.502, 4.11.1.2.1.

Ausgelöst wird ein inter-RAT Handover durch den gNB, wenn das Endgerät bei schlechtem Signalpegel keine geeignete gNB Nachbarzelle findet, jedoch geeignete LTE Zellen gefunden wurden. Der gNB fordert daraufhin einen inter-RAT Handover bei der AMF an. Die AMF startet dann die Handover Prozedur mit einer Anfrage an die SMF nach den Session Kontext Informationen des Teilnehmers. Danach kontaktiert die AMF dann die LTE MME über das N26 Interface mit einer Relocation Request Nachricht. Die LTE MME kontaktiert daraufhin das LTE SGW mit einer Create Session Request Nachricht, damit entsprechend Ressourcen für den Nutzdatentunnel im LTE Netzwerk bereitstehen. Zusätzlich kontaktiert die LTE MME auch die LTE Zielzelle mit einer Handover Request Nachricht. Diese Nachricht enthält alle für den Aufbau des Nutzdatenbearers

benötigten Informationen. Nachdem der LTE eNB einen Kontext für das Endgerät angelegt hat, sowie der Nutzdatentunnel bereitsteht, antwortet die LTE MME der AMF mit einer Acknowledge Nachricht. Weiterhin wird ein „Indirect Data Tunnel“ zwischen 5G gNB und dem 4G eNB angelegt, über den alle IP Datenpakete während des Handovers weitergeleitet werden. Dies minimiert Auswirkungen des Handovers auf laufende Datentransfers.

Nachdem der indirekte Tunnel aufgebaut ist, startet die Ausführungsphase des Handovers. Die AMF sendet dazu eine Handover Command Nachricht an den gNB, der diese dann an das Endgerät weitergibt. Nachdem das Endgerät informiert wurde, wird der indirekte Tunnel für die Nutzdaten zwischen gNB und eNB für die Weiterleitung der Daten verwendet. Nachdem der Handover auf der Luftschnittstelle nach 4G durchgeführt wurde, sendet das Endgerät eine Handover Complete Nachricht an den LTE eNB, der dies dann als Handover Notify Nachricht an die LTE MME weiterleitet.

Im Kernnetz informiert die LTE MME dann die AMF über den erfolgreichen Handover über die N26 Schnittstelle. Diese leitet dann die letzte Phase des Handover ein. In dieser Aufräumphase wird der indirekte Tunnel abgebaut und alle Daten fließen direkt vom Internet über die UPF/PGW, den SGW und den eNB zum Endgerät.

Außerdem wird auch der Kontext des Endgeräts auf der 5G Seite des Kernnetz und aus dem gNB gelöscht. Im letzten Schritt führt das Endgerät dann eine Tracking Area Update Prozedur auf der LTE Seite durch. Diese Prozedur informiert die MME und die HSS Datenbank über die aktuelle Tracking Area des Endgerätes, welche später für das Paging des Endgerätes im RRC-IDLE Zustand bei Eintreffen von neuen IP Paketen aus dem Internet notwendig ist.

An dieser Stelle sei angemerkt, dass die Cell Reselection und Handover Prozeduren nur zwei Beispiele sind. Im Standard sind weitere Prozeduren spezifiziert, wie z. B. Connected Mode Handover und Idle Mode Cell Reselection in umgekehrter Richtung. Es ist jedoch davon auszugehen, dass die zwei oben beschriebenen Methoden die am häufigsten benutzten inter-RAT Prozeduren sind, solange die LTE Netzarbeitung besser als die 5G NR Netzarbeitung in einem Netzwerk ist.

2.8.12 Das 5G Kernnetz und SMS

Bei der Spezifikation des 5G Kernnetzes wurde beschlossen, keine Schnittstellen zum 2G und 3G Kernnetz zu spezifizieren. Man ging davon aus, dass dies nicht mehr länger notwendig war. Einzige Ausnahme ist der Short Message Service (SMS), der ursprünglich für GSM spezifiziert wurde, und seither über alle Evolutionsstufen des Kernnetzes verfügbar ist. Diese Ausnahme wurde gemacht, da SMS trotz deutlich sinkender Nutzung noch immer ein zentraler Dienst ist. 3GPP TS 23.501, 6.2.13 spezifiziert dazu eine Schnittstelle zum SMS Service Center und eine SMS Function (SMSF). Die SMSF hat ein API (Nsmsf) zur Kommunikation mit der AMF, das in TS 29.540³⁴ spezifiziert ist. Wie auch eine ähnliche Komponente im LTE Kernnetzwerk stellt die SMSF

die Kommunikation zwischen der Non-Access Stratum Signalisierung und dem SMS Service Center her. Das bedeutet, dass SMS Nachrichten in N1 Signalisierungsnachrichten zwischen Endgerät und AMF eingebettet werden können.

In der Praxis benötigen die meisten 5GC-fähigen Endgeräte die SMS Unterstützung über NAS jedoch nicht, da SMS ein Teil des IMS Sprachdienstes (VoLTE, VoNR) ist. Somit werden für solche Endgeräte SMS Nachrichten per SIP Protokoll übertragen. Somit wird SMS über 5G NAS Signalisierung nur für spezielle Endgeräte im Internet der Dinge Umfeld benötigt, die über keine Telefoniefunktion verfügen.

2.8.13 Das Cloud Native 5G Kernnetz

In den vorherigen Unterkapiteln wurde das 5G Kernnetz aus einer operationellen und funktionalen Sichtweise beschrieben, unabhängig davon, wie die Netzkomponenten und Dienste strukturiert sind. Dieses Unterkapitel zeigt nun, wie sich die Kernnetzarchitektur in 3GPP über die Zeit von Generation zu Generation gewandelt hat.

In 2G und 3G Netzwerken werden alle Netzfunktionen als Netzwerkelemente beschrieben, die in physischer Hardware implementiert wurden. Typischerweise wurde für jeden Netzelementtyp eine speziell dafür entwickelte Hardware verwendet.

Der 2G/3G SGSN beispielsweise, der die meisten Funktionen der AMF, SMF und des UPFs beinhaltete, wurde als einzelne Hardwarekomponente angesehen. Die meisten Netze nutzten mehrere SGSNs aus Redundanz und Kapazitätsgründen. Jeder SGSN bestand aus einem oder mehreren Racks mit speziell für diesen Zweck entwickelten Hardwarekomponenten. Manche dieser Komponenten, meist Einschubkarten, kümmerten sich dabei um die Netzwerkkommunikation, andere Karten waren für den Massenspeicher verantwortlich, und wieder andere Karten hatten CPUs, Speicher und spezielle Hardwarebeschleuniger für die eigentliche SGSN Software. Eine Backplane verband die Karten im und zwischen den Racks und ein proprietäres Betriebssystem und Software wurde in allen Teilen des Systems verwendet.

Bei Einführung von LTE änderte sich diese Architektur grundlegend. Statt proprietärer Hardware wurde nun vermehrt auf Standard Intel x86 Server und standardisierte Netzkomponenten gesetzt. Die Komponenten der Signaling Plane, die hauptsächlich Rechenleistung benötigten, sowie die User Plane Komponenten, die hauptsächliche IP Nutzdatenpakete weiterleiten, wurden in der 3GPP Spezifikation einzeln spezifiziert und konnten somit unabhängig voneinander entwickelt und eingesetzt werden. Ein Beispiel ist die LTE MME und das S-GW. Die Hauptaufgabe der MME ist das Control Plane Signaling. Dies benötigt im Wesentlichen Rechenleistung und diese Komponente konnte somit auf Standard x86 Servern, wie sie auch in Rechenzentren verwendet werden, implementiert werden. Das S-GW auf der anderen Seite ist einem gewöhnlichen IP Router sehr ähnlich, der spezielle Hardware für das effiziente Routing von IP Paketen verwendet. Um einen solchen Router als S-GW zu verwenden, ist nur wenig zusätzliche Software nötig. Diese kümmert sich dann um die IP basierten GTP

Tunnel und muss von der MME Befehle empfangen können, diese aufzubauen, abzubauen und zu modifizieren. Dies war in der Telekommunikationsindustrie ein wichtiger Schritt nach vorne, da man somit auf spezielle und somit sehr teure Hardware, die nur in kleinen Stückzahlen produziert wurde, verzichten konnte. Stattdessen werden nun in großer Stückzahl produzierte und somit wesentlich günstigere Server und Router verwendet, die sich zusätzlich auch schnell weiterentwickeln. Dies ist wichtig, um mit dem ständig wachsenden Datenverkehr schritthalten zu können.

In einem weiteren Schritt wurde dann dazu übergegangen, die Kernnetzsoftware nicht mehr direkt auf einem physischen Server auszuführen, sondern in virtuelle Maschinen zu verlagern. Als Betriebssystem kommt hier heute vorwiegend Linux zum Einsatz. Vor allem für die Signalling Plane und die Datenbankfunktionen ist dies sehr einfach möglich, da die LTE MME und das HSS hauptsächlich Rechenleistung und Hauptspeicher benötigen. Der Vorteil dieses Ansatzes ist, dass das Netzwerk einfacher skalierbar und unabhängig von speziellen Hardwarelieferanten wird. Statt eines dedizierten Racks und Server Blades für eine bestimmte Netzfunktion zu verwenden, können die virtuellen Maschinen nun unterschiedliche Netzfunktionen beinhalten und auf beliebigen Servern ausgeführt werden. Die Verwendung von virtuellen Maschinen trennt zusätzlich auch die Software von der Hardware. Somit können sich Software und Hardware unabhängig voneinander entwickeln und die Software muss nicht mehr zusammen mit der Hardware vom gleichen Hersteller gekauft werden. Weitere Details hierzu finden sich im LTE Kapitel.

Die folgende Liste gibt nochmals einen Überblick, wie sich Mobile Kernnetze von 2G bis 4G LTE entwickelt haben:

- Proprietäre Hardware wurde für eine spezifische Netzkomponente entwickelt. Software und Hardware sind integriert.
- Verwendung von standardisierter Hardware für Rechenzentren mit x86 Prozessoren. Software läuft direkt auf der Hardware.
- Einführung von virtuellen Maschinen, klare Trennung von Hardware und Software.

Das 5G Netzwerk könnte auf alle oben beschriebenen Arten implementiert werden. Bei der Standardisierung der ersten Version des 5GC gab es jedoch bereits eine weitere fundamentale Entwicklung in der Internetwelt, die dann auch für das 5GC verwendet wurde:

Wie bereits zuvor beschrieben, teilt die Spezifikation den 5G Core strikt in Control Plane und User Plane ein. Dies wird auch als Control-User Plane Separation (CUPS) bezeichnet. Statt Netzknoten wie bisher werden nun „Funktionen“ spezifiziert und die Kommunikation zwischen diesen Funktionen ist nun „Service basiert“ (Service based). Das bedeutet, dass die Kommunikation zwischen den Funktionen nun Zustandslos (stateless) ist und das HTTP Protokoll für den Zustandslosen Anfrage/Antwort Mechanismus verwendet wird. Daten in den Anfragen und Antworten werden im JSON (JavaScript Object Notation) Format übertragen. Dies unterscheidet sich deutlich von früheren

Kernnetz Spezifikationen, da bisher immer ein zustandsorientiertes (stateful) Protokoll verwendet wurde.

Diese Änderungen sind jedoch nur ein Mittel zum Zweck. Schon in der ersten Version der 3GPP 5G Core Spezifikation wurde eine „Cloud Native“ Implementierung des Kernnetzes ermöglicht, die auf Containern basiert. Diese werden nicht in virtuellen Maschinen ausgeführt, sondern direkt auf physischen Servern mit einer Container Orchestration Software³⁵. Trotz direkter Ausführung auf der Hardware trennen Container die Software von der darunterliegenden Hardware, ohne dafür eine virtuelle Maschine zu benötigen, die sämtliche Hardware eines Servers emulieren. Dies hat folgende Vorteile gegenüber den bisherigen Ansätzen:

Für Softwareentwicklungsteams sind Container eine ideale Entwicklungsplattform, da Container nach einer Änderung immer wieder neu auf beliebigen Servern erzeugt werden können. Dies ist möglich, da ein Container die komplette Software inklusive aller Softwarebibliotheken bündelt, die zum ausführen der Applikation notwendig ist. Software, die in anderen Containern auf dem gleichen Server ausgeführt werden, können andere Versionen der gleichen Softwarebibliotheken verwenden. Somit wird erreicht, dass ein Container, der auf einem Testsystem eines Entwicklers läuft, später auch im Produktiveinsatz auf einem anderen Server ohne Änderungen funktioniert.

Der „Cloud Native“ Ansatz beschreibt nicht nur eine neue Methode der Softwareentwicklung, sondern teilt auch die Funktionalität einer großen Applikation in kleine Microservices auf. Jeder Microservice läuft in einem eigenen Container und mehrere Containerinstanzen werden verwendet, um die Rechenkapazität eines Dienstes zu steigern. Das nachfolgende Beispiel aus der Internetwelt soll diesen Ansatz verdeutlichen:

Für eine Wordpress Installation einer Firmenwebseite wird heute typischerweise eine virtuelle Maschine (VM) verwendet. Nachdem die VM erzeugt werden ist, wird üblicherweise zunächst ein Linux Betriebssystem installiert. Danach wird ein Webserver wie Apache oder Nginx installiert und konfiguriert, sowie eine Datenbanksoftware wie Maria-DB für die Konfigurationsinformationen und den Inhalt der Webpräsenz. Und schließlich wird die Wordpress PHP Software in der VM installiert.

Wird stattdessen ein Container basierter Ansatz gewählt, muss kein Betriebssystem installiert werden, da der Container alle Dienste des Linuxbetriebssystems außerhalb des Containers verwendet. Für den Webserver und die Datenbank werden zwei von einander unabhängige Container verwendet. Da der Webserver und die Datenbanksoftware Open Source sind, stellen die dahinter stehenden Projekte fertige Container Images bereit, die schon die komplette Software inklusive der benötigten Softwarebibliotheken enthalten. In diesem Beispiel würde die Web Server Container Vorlage (Template) durch die Wordpress Installation über einen weiteren Layer erweitert werden. Während die komplette Software in Containern installiert ist, werden alle Daten, die während des Betriebs geändert werden, außerhalb der Container auf dem Server gespeichert. In diesem Beispiel wären das die Datenbankdateien, die alle Webseiten sowie Konfigurationsinformationen enthalten. Diese werden auf dem Hostsystem

abgelegt und dem Datenbank Container über einen Dateipfad zur Verfügung gestellt. Auf diese Weise kann jederzeit ein Softwareupdate des Containers erfolgen, ohne dass die Daten davon berührt werden. Im letzten Schritt wird ein virtuelles Netzwerk zwischen den zwei Containern konfiguriert, damit die Wordpress Installation in einem Container auf die Datenbank im anderen Container zugreifen kann. Somit spielt es dann auch keine Rolle, ob die zwei Container auf dem gleichen, oder auf unterschiedlichen Servern ausgeführt werden.

Ein entscheidender Vorteil des Containeransatzes ist, dass Container bei der Entwicklung und während dem Rollout auf die gleiche Weise erzeugt werden und alle verwendeten Softwarebibliotheken enthalten. Die einzige Abhängigkeit eines Containers ist somit zum Linux Kernel. Bei der Kernel API wird jedoch sehr stark auf Rückwärtskompatibilität geachtet und kann somit als unveränderlich angesehen werden.

Wird Software in Container verpackt und in viele unabhängige Microservices aufgeteilt, ist es wichtig, ein Container Management System zu verwenden. Diese Aufgabe wird auch als „Orchestration“ bezeichnet. Orchestration ist ein wichtiger Teil des gesamten Systems, da es ab einem gewissen Punkt schwierig wird, Microservices und die Interaktion zwischen den Containern manuell zu administrieren. Zu automatisierende Managementaufgaben sind z. B. die Konfiguration, welche Container mit welchen anderen Containern kommunizieren dürfen, sowie das System Monitoring und die Fehlersuche. Zusätzlich können mit einer automatischen Orchestrierung die Container über viele Server in einem Cluster verteilt werden. Dies erzeugt Redundanz und erhöht die Kapazität des Gesamtsystems. In der Praxis hat sich für diese Aufgabe das Open Source Orchestration System Kubernetes durchgesetzt³⁶.

2.9 Die 5G Standalone Luftschnittstelle

Am Anfang dieses Kapitels wurde die 5G Luftschnittstelle im Zusammenhang mit dem 5G NR Non-Standalone (NSA) Modus beschrieben. Mit der Einführung von 5G NR Standalone (SA) benötigen dann SA-fähige Endgeräte keine LTE Ankerzelle mehr. Somit werden über die 5G Luftschnittstelle nun auch alle Signalisierungs- und Managementoperationen abgewickelt, und das Endgerät muss die 5G Non-Access Stratum (NAS) Protokolle des 5G Kernnetzes beherrschen. Zusätzlich müssen gNB und Endgeräte auch das 5G Radio Resource Control (RRC) Protokoll implementieren. Details sind in 3GPP TS 38.331³⁷ beschrieben. Grundsätzlich hat das 5G RRC Protokoll viele Gemeinsamkeiten mit dem 4G LTE RRC Protokoll, das in 3GPP TS 36.331³⁸ beschrieben ist. Es enthält jedoch eine Anzahl von Erweiterungen und Verbesserungen, die nun nachfolgend beschrieben werden.

2.9.1 Der RRC Inactive Zustand

Während es auf der LTE Luftschnittstelle mit RRC-IDLE und RRC-CONNECTED zwei Zustände geben kann, wurde mit 5G ein Dritter hinzugefügt: RRC-INACTIVE. Dieser Zustand befindet sich zwischen RRC-IDLE und RRC-CONNECTED. Die Idee dabei ist, den Zustand der 5G Luftschnittstelle vor dem Kernnetz zu verbergen und so die Anzahl der Tunnel Auf- und Abbauten zwischen dem gNB und dem Kernnetz zu reduzieren. Heute führen Endgeräte wie Smartphones und Tablets durch die sporadische Kommunikation vieler Applikationen ständig einen Wechsel zwischen RRC-IDLE und RRC-CONNECTED aus. Dies ist auch der Fall, wenn der Bildschirm abgeschaltet ist und auch wenn das Gerät über längere Zeit nicht verwendet wird. Auch in diesen Fällen senden Applikationen im Hintergrund sporadisch Datenpakete, um z. B. TCP Verbindungen offen zu halten. Typischerweise werden solche „Keep-Alive“ Pakete im Abstand von nur wenigen Minuten gesendet und entsprechend oft wird somit auch der RRC Zustand geändert, um möglichst wenig Energie zu verbrauchen. Somit muss dann jedes Mal der Nutzdatentunnel zwischen gNB und Kernnetz aufgebaut und dann nach kurzer Zeit wieder abgebaut werden.

Um dies zu reduzieren, kann der gNB das Endgerät anweisen, in den RRC-INACTIVE Zustand zu wechseln. Dies geschieht mit einer RRC-Release Nachricht, die ein „Suspension Configuration“ Setup enthält. Danach wird die Radioverbindung abgebaut und somit Energie gespart. Erhalten bleibt jedoch die Verbindung zur AMF im Kernnetzwerk und der Nutzdatentunnel zur UPF. Wenn neue Daten für das Endgerät vom Netzwerk ankommen, werden diese einfach an den gNB weitergeleitet, zu dem aktuell der Nutzdatentunnel aufgebaut ist. Der gNB startet dann eine „RAN-based“ Paging Prozedur mit allen gNBs, die in der gleichen RAN Notification Area (RNA) gruppiert sind. In dieser RNA kann sich das Endgerät bewegen, ohne das Netzwerk über den Wechsel des gNB zu informieren. Für das Kernnetz ist dieses Paging transparent. Wenn das Endgerät auf ein RAN-based Paging mit einer RRC Resume Request Nachricht von einem gNB in der RNA antwortet, zu dem der Nutzdatentunnel nicht aufgebaut ist, wird die Signalisierungsverbindung dorthin transferiert. In einem weiteren Schritt wird dann auch der Nutzdatentunnel dorthin verlegt.

Bewegt sich das Endgerät in eine Zelle außerhalb der RAN Notification Area, muss es eine RRC Verbindung aufbauen und eine RNA Update Nachricht senden. Der gNB kontaktiert dann den vorherigen gNB, der daraufhin alle Kontextinformationen für diesen Nutzer sowie den Nutzdatentunnel an den neuen gNB übergibt. Danach kann die RRC Verbindung wieder in den RRC-INACTIVE Zustand wechseln und das Endgerät kann sich dann in der neuen RNA ohne Kommunikation mit dem Netzwerk bewegen.

2.9.2 System Information Nachrichten

Wie auch in den früheren Radionetzwerkgenerationen werden Informationen, die Endgeräte für das Auffinden der Zellen, für die Random Access Prozedur und für das Cell Reselection benötigen, in System Information Broadcast (SIB) Nachrichten übertragen. Auf der 5G Luftschnittstelle werden die meisten SIBs nur für den 5G Standalone Modus benötigt, und deren Struktur gleicht den bereits von LTE bekannten Nachrichten. Sowohl bei LTE, als auch auf der 5G Luftschnittstelle werden SIBs periodisch ausgestrahlt. Auf der 5G Luftschnittstelle gibt es jedoch auch die Möglichkeit, SIBs nicht periodisch, sondern nur auf Anfrage des Endgeräts zu senden. Die folgenden SIBs wurden für die 5G Luftschnittstelle spezifiziert:

- **MIB (Master Information Block):** Enthält Informationen, wo SIB-1 gefunden werden kann und ob die Zelle generell gesperrt (barred) oder nur für Wartungspersonal zugänglich sein soll (Operator Use).
- **SIB 1:** Konfigurationsparameter der aktuellen Zelle wie z. B. der Mobile Country Code und der Mobile Network Code (MCC, MNC), der niedrigste noch erlaubte Signalpegel, mit dem die Zelle noch verwendet werden darf, die Cell ID, der Tracking Area Code, etc. Dieser SIB wird immer ausgestrahlt, da alle Endgeräte diese Informationen für den Aufbau einer Radio Verbindung (Random Access Prozedur) benötigen.
- **SIB 2:** Enthält Parameter für intra-Frequency, inter-Frequency und inter-RAT Cell Reselection, sowie Reselection Informationen wie minimale Signalpegel für den Zugriff, bei welchem Signalpegel begonnen werden soll, nach Nachbarzellen zu suchen, etc.
- **SIB 3:** Detaillierte Parameter für intra-Frequency Cell Reselection
- **SIB 4:** Inter-Frequency NR Cell Reselection Informationen, wie die genutzten Frequenzbänder und Kanalnummern, minimale Signalpegel und Bandprioritäten. Die Bandpriorität wird der Signalstärke vorgezogen, da es oft vorteilhaft ist, Endgeräte in Zellen auf höheren Bändern mit größerer Kanalbandbreite aber niedrigem Signalpegel zu belassen, anstatt diese in Zellen auf niedrigeren Bändern mit stärkerem Signalpegel zu lenken.
- **SIB 5:** Inter-RAT Cell Reselection Parameter mit Prioritäten für die unterschiedlichen Radiotechnologien. Auf diese Art kann ein Endgerät angewiesen werden, bei Verlust der 5G Netzarbeitung LTE zu bevorzugen, auch wenn ein 2G oder 3G Kanal besser empfangen werden kann.
- **SIB 6, 7 und 8:** Informationen über ETWS und CMAS Public Warning Broadcasts für Erdbeben, Überschwemmungen und andere Katastrophenfälle. Diese werden nicht in allen Ländern verwendet.
- **SIB 9:** Uhrzeitinformationen in UTC, local und GPS Zeit (optional).

2.9.3 Messkonfiguration, Events und Handover

Nachbarzellenmessungen im gleichen Band, in anderen Frequenzbändern, sowie die Suche nach Zellen einer anderen Radiotechnologie werden in gleicher Weise wie bereits im LTE Kapitel beschrieben, durchgeführt. Deshalb folgt hier nur eine kurze Zusammenfassung. Endgeräte müssen aus mehreren Gründen das Downlinksignal messen und Feedback an den gNB geben:

- Damit Kanäle (Carrier) zu einer Verbindung hinzugefügt und entfernt werden können (Carrier Aggregation).
- Für das rechtzeitige Einleiten eines Handovers zu Nachbarzellen.
- Für die Durchführung eines Handovers zu Zellen in anderen Bändern, z. B. in ein niedrigeres Frequenzband, wenn sich das Endgerät weiter von der Zelle entfernt.
- Um in ein anderes Radionetzwerk zu wechseln, wenn das Endgerät den 5G SA Abdeckungsbereich verlässt.

Wie in LTE werden Messungen mit RRC Reconfiguration Nachrichten an das Endgerät übermittelt und bestehen aus drei Teilen. Measurement Objekte beschreiben eine zu messende Carrier Konfiguration. Die Beschreibung enthält die Radio Technologie (NR, LTE), die Carrier Frequency (ARFCN), sowie die Kanalbandbreite. In der Report Konfiguration wird dann der Messtyp (Event Type) konfiguriert, der auf unterschiedliche Measurement Objekte angewandt werden kann. Für 5G NR sind folgende Events in 3GPP TS 38.331 spezifiziert:

- **Event A1:** Das Signal der Serving Cell übersteigt einen konfigurierten Grenzwert (Threshold Value).
- **Event A2:** Das Signal der Serving Cell unterschreitet einen konfigurierten Grenzwert.
- **Event A3:** Das Signal der Nachbarzelle wird stärker als das Signal der aktuellen Zelle.
- **Event A4:** Das Signal der Nachbarzelle wird besser als ein konfigurierter Wert.
- **Event A5:** Das Signal der Serving Cell wird schlechter als ein Grenzwert, die Signalstärke der Nachbarzelle wird besser als ein Grenzwert.
- **Event A6:** Das Signal der Nachbarzelle wird um einen Schwellwert besser als das Signal der Serving Cell.

Für LTE inter-RAT Messungen, die typischerweise an der Grenze der NR Abdeckung konfiguriert werden, sind folgende Events definiert:

- **Event B1:** Das Signal einer inter-RAT Nachbarzelle wird besser als ein konfigurierter Schwellwert.

- **Event B2:** Das Signal einer Serving Cell wird schlechter als Schwellwert 1 und das Signal einer inter-RAT Nachbarzelle wird besser als Schwellwert 2.

Vergleicht man die Namen der Events mit denen von LTE, so fällt auf, dass diese in LTE und 5G NR identisch sind. B1 und B2 Events können jedoch nur für LTE Zellen verwendet werden. GSM und UMTS inter-RAT Messungen wurden in 3GPP Release 15 nicht spezifiziert, da keine Handover in diese Radionetze möglich sein sollen. Schwellwerte (Thresholds) können in der Eventkonfiguration für folgende Parameter gegeben werden:

- **RSRP:** Die Reference Signal Received Power.
- **RSSI:** Die Received Signal Strength Indication.
- **RSRQ:** Die Reference Signal Received Quality.

Details hierzu sind im LTE Kapitel zu finden.

Der dritte Teil einer Messkonfiguration sind die eigentlichen „Measurements“. Diese kombinieren Measurement Objekte mit Measurement Konfigurationen. Auf diese Weise kann eine Measurement Konfiguration mit mehreren Measurement Objekten verknüpft werden.

Sendet ein Endgerät einen Measurement Report, reagiert das Netzwerk dann typischerweise mit Aktionen wie dem Hinzufügen oder Entfernen eines Kanals einer Carrier Aggregation Konfiguration, oder einem Handover zu einer NR oder inter-RAT Nachbarzelle. Dieser Prozeduren sind mit den im LTE Kapitel beschriebenen Funktionen identisch.

2.10 Network Slicing

Ein in der Öffentlichkeit sehr häufig diskutiertes 5G Konzept ist Network Slicing. 3GPP TS 38.300³⁹ gibt hierzu einen Architekturüberblick. In der Praxis ist Network Slicing eine 5G Ende-zu-Ende Quality of Service Architektur und erweitert Funktionalitäten, die für LTE schon in 3GPP Release 13 und 14 spezifiziert wurden. Um Network Slicing zu verwenden, müssen Endgerät, Zugangsnetzwerk und Kernnetz die dazu notwendigen Funktionen unterstützen.

Um das Ziel von Network Slicing besser zu verstehen, ist es hilfreich, die historische und aktuelle Nutzung von Mobilfunknetzwerken zu betrachten: Das 4G LTE Zugangsnetzwerk wurde ursprünglich ausschließlich für den schnellen mobilen Zugang zum Internet ausgelegt. Die LTE Luftschnittstelle konnte somit für genau diese Anwendung optimiert werden. Dadurch eignet sich LTE jedoch nicht für Geräte und Anwendungen, die nur sehr wenig Daten übertragen und sehr energieeffizient sein müssen. Aus diesem Grund wurde dann später die Narrow-Band Internet of Things (NB-IoT) Luftschnittstelle spezifiziert. NB-IoT kann in einen LTE Kanals eingebettet werden, oder am Rand

des Kanals verwendet werden, oder auch für sich alleine stehen. Falls NB-IoT innerhalb eines LTE Kanals verwendet wird, werden einige Subcarrier auf der Frequenzachse aus dem LTE Resource Grid herausgenommen. Der LTE Scheduler umgeht dann diese Subcarrier und NB-IoT ist somit für normale LTE Geräte nicht sichtbar. NB-IoT Geräte haben andererseits ein Radiomodul mit limitierter Bandbreite und sehen somit nur den NB-IoT Kanal. Trotzdem verwenden LTE und NB-IoT Geräte den gleichen Kanal. Man könnte also sagen, dass LTE für Breitbandnutzung ein Radionetzwerk „Slice“ darstellt und der NB-IoT Kanal einen anderen Radionetzwerk „Slice“. Die Bezeichnung „Slice“ wurde hier in Anführungszeichen gesetzt, da dieser Begriff in den 4G LTE Spezifikation nicht existiert. Trotz Nutzung des gleichen Kanals haben die zwei „Slices“ sehr unterschiedliche Layer 2 Eigenschaften in Bezug auf Bandbreite, Datenrate und Kanalkonfiguration.

Netzbetreiber können LTE und NB-IoT Kunden mit dem gleichen Kernnetzwerk bedienen, es werden jedoch typischerweise zwei unabhängige Kernnetzwerke (MME/S-GW + PGW) verwendet. Dies könnte man aus Kernnetsicht als zwei „Network Slices“ ansehen. Die Nutzung unterschiedlicher Kernnetzwerke kann in diesem Zusammenhang sehr nützlich sein, da das Verhältnis zwischen Signalisierung und Nutzdaten von NB-IoT Kunden sich deutlich von LTE Kunden unterscheidet. Die zwei unterschiedlichen Kernnetzwerke können somit unterschiedlich optimiert werden.

Viele Netzbetreiber möchten in Zukunft nicht nur den traditionellen Internetzugangsmarkt, sondern auch verstärkt neue Anwendungsfälle abdecken. Deswegen wurde das 5G Zugangsnetzwerk so gestaltet, dass unterschiedliche Layer 2 Konfigurationen auf dem gleichen Kanal auf der Luftschnittstelle möglich sind, und auch in Zukunft möglichst keine problematischen Kompromisse aus Gründen der Rückwärtskompatibilität notwendig sind. In 3GPP TS 38.300 ist festgelegt, dass ein Endgerät mindestens 8 Slices unterstützen soll, während das Netzwerk nicht auf diese Zahl festgelegt ist. Innerhalb eines Slice werden dann die gleichen Quality of Service (QoS) Maßnahmen verwendet, die es heute auch schon bei LTE gibt und somit können unterschiedliche Datenpakete mit unterschiedlichen Prioritäten versehen werden.

Das Network Slice Konzept hatte auch einen großen Einfluss auf das Design des 5G Kernnetzwerk. Es wurde so gestaltet, dass ein bestimmtes Radionetzwerk Slice mit einem eigenen Core Network Slice verbunden werden kann. Wenn sich ein Endgerät mit dem 5G Radio Netzwerk verbindet, entscheidet der gNB, zu welchem Core Network Slice die Signalisierungsnachricht gesendet werden soll. Dies ist analog zu LTE, falls zusätzlich NB-IoT auf einem Träger verwendet wird. In Abhängigkeit ob das Endgerät auf LTE oder NB-IoT zugreift, werden die Daten zum LTE oder zum NB-IoT Kernnetz (MME/S-GW) weitergeleitet. Mit 5G Network Slicing wurde dieses Konzept weiter flexibilisiert. Die Entscheidung, mit welchem Kernnetz kommuniziert werden soll, wird nun nicht mehr anhand der Verwendung einer bestimmten Layer 2 Luftschnittstelle getroffen. Stattdessen kann nun das Endgerät dem gNB bei der Verbindungsaufnahme mitteilen, mit welchem Slice ein Bearer verbunden werden soll. Außerdem ist es nun

möglich, nicht nur wie bei LTE mit nur einem Kernnetz verbunden zu sein, sondern mit bis zu 8 Core Network Slices gleichzeitig.

In Kernnetzwerk ist das Endgerät immer mit einer einzigen AMF Instanz verbunden, auch wenn es Zugang zu mehreren Slices hat. Für jeden Slice teilt die AMF dann eine Session Management Function (SMF) und andere Netzelemente entsprechend zu.

Network Slicing bietet eine große Flexibilität. In der Praxis benötigen jedoch Endgeräte dies für den normalen Internetzugang mit gleicher Priorität für alle Datenpakete nicht. Somit könnten diese Mechanismen vor allem in Fabrik (Campus) Netzwerken und anderen Umgebungen sinnvoll sein, in denen manche Endgeräte und Applikationen eine höhere Priorität als andere haben sollen.

2.11 Fragen

1. Welche grundsätzlichen Konzepte verwendet die 5G Non-Standalone (NSA) Architektur?
2. Welche Unterschiede gibt es zwischen der 5G TDD und FDD Luftschnittstelle?
3. Was ist ein Split-Bearer?
4. Warum werden zwei Sender im Endgerät für einen 5G Option 3 Split-Bearer benötigt?
5. Was ist ein Bandwidth Part (BWP)?
6. Was ist ein CORESET?
7. Für was wird Dynamic Spectrum Sharing (DSS) beim „Reframing“ verwendet?
8. Warum können der 4G und der 5G Teil eines NSA Bearers unabhängig voneinander zu einer anderen Zelle (Handover) übergeben werden?
9. Welches Konzept steckt hinter der Service Oriented Architecture des 5G Kernnetzes.
10. Was ist der Unterschied zwischen Registrierung und Session Establishment?
11. Was ist der Unterschied zwischen RRC-IDLE und RRC-Inactive Zustand?
12. Beschreibe das Konzept des 5G Network Slicing. Welche Vorteile gibt es für den mobilen Internetzugang?

Lösungen sind auf der Website zum Buch unter <http://www.cm-networks.de> zu finden.

Anmerkungen

1. 3GPP, NG-RAN; Architecture description, TS 38.401.3GPP, NG-RAN; Architecture description, TS 38.401.
2. 3GPP, System architecture for the 5G System (5GS), TS 23.501.
3. M. Sauter, 3GPP 5G NR – What's the 'g' in gNB all about – Part 2, <https://blog.wirelessmoves.com/2016/11/3gpp-5g-nr-whats-the-g-in-gnb-all-about-part-2.html>, November 2016.
4. 3GPP, NR; Multi-connectivity; Overall description; Stage-2, TS 37.340.

5. The Common Public Radio Interface Forum, CPRI specification page, <http://www.cpri.info/spec.html>.
6. M. Sauter, Paranormal 5G Numerology, <https://blog.wirelessmoves.com/2018/01/paranormal-5g-numerology.html>, January 2018.
7. A. Zaidi, R. Baldemair, In the race to 5G, CP-OFDM triumphs!, <https://www.ericsson.com/en/blog/2017/5/in-the-race-to-5g-cp-ofdm-triumphs>, May 2017.
8. 3GPP, NR; Physical layer procedures for control, TS 38.213, Abschn. 11.1.
9. M. Sauter, 5G TDD Inter-Operator Network Synchronization, <https://blog.wirelessmoves.com/2020/02/5g-tdd-inter-operator-network-synchronization.html>, February 2020.
10. J. Ryu, 5G/NR – Carrier Bandwidth Part, http://www.sharetechnote.com/html/5G/5G_CarrierBandwidthPart.html.
11. Halbherd Bastion, Radio Frequency Technologies, <https://halberdbastion.com/technology/cellular/5g-nr/5g-frequency-bands/n257-28-ghz>.
12. IEEE ComSoc, GSA Report: Spectrum Above 6 GHz & related FCC Activity, <https://techblog.comsoc.org/2019/12/05/gsa-report-spectrum-above-6-ghz-related-fcc-activity>, December 2019.
13. M. Sauter, 5G – How do mmWave Antennas Look Like?, <https://blog.wirelessmoves.com/2018/11/5g-how-do-mmwave-antennas-look-like.html>, November 2018.
14. M. Sauter, 5G EN-DC: Flow Control Between 4G and 5G, <https://blog.wirelessmoves.com/2018/08/5g-en-dc-flow-control-between-4g-and-5g.html>, August 2018.
15. 3GPP, General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access, TS 23.401, Abschn. 4.3.2a, 5.3.2.1, 5.7.1, 5.11.3.
16. 3GPP, Multi-connectivity; Overall description; Stage-2, TS 37.340.
17. Niviuk, NR Frequency band calculator, https://www.sqimway.com/nr_band.php.
18. 3GPP, Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification, TS 36.331.
19. M. Sauter, How Does An LTE eNB Signal A Co-Located 5G Cell?, <https://blog.wirelessmoves.com/2018/04/how-does-an-lte-enb-signal-endc.html>, April 2018.
20. Yilmaz O and Teyeb O. LTE-NR tight-interworking and the first steps to 5G, <https://www.ericsson.com/en/blog/2017/11/lte-nr-tight-interworking-and-the-first-steps-to-5g>, November 2017
21. Wikipedia, Small form-factor pluggable transceiver, https://en.wikipedia.org/wiki/Small_form-factor_pluggable_transceiver.
22. Ericsson, Ericsson Microwave Outlook, <https://www.ericsson.com/4a8c1f/assets/local/reports-papers/microwave-outlook/2019/ericsson-microwave-outlook-report-2019.pdf>, October 2019.
23. 3GPP, System architecture for the 5G System (5GS), TS 23.501.
24. 3GPP, Procedures for the 5G System (5GS), TS 23.502.
25. 3GPP, Non-Access-Stratum (NAS) protocol for 5G System (5GS), TS 24.501.
26. 3GPP, System architecture for the 5G System (5GS), TS 23.501, Abschn. 5.9.

27. B. Aboba et al., The Network Access Identifier, RFC 4282.
28. 3GPP, Numbering, addressing and identification, TS 23.003.
29. 3GPP, 5G System; Policy and Charging Control signalling flows and QoS parameter mapping; Stage 3, TS 29.513.
30. 3GPP, Security architecture and procedures for 5G System, TS 33.501.
31. P. Nakarmi, D. Castellanos, Does the switch to 5G security require a new SIM card?, <https://www.ericsson.com/en/blog/2020/1/5g-security-sim-card>, January 2020.
32. 3GPP, System architecture for the 5G System (5GS), TS 23.501, Abschn. 4.3.
33. 3GPP, Procedures for the 5G System (5GS), TS 23.502, Abschn. 4.11.1.
34. 3GPP, 5G System; SMS Services; Stage 3, TS 29.540.
35. 5G Americas, 5G and the cloud, <https://www.5gamericas.org/5g-and-the-cloud>, December 2019.
36. Wikipedia, Kubernetes, <https://en.wikipedia.org/wiki/Kubernetes>.
37. 3GPP, 5G; NR; Radio Resource Control (RRC); Protocol specification, TS 38.331.
38. 3GPP, LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification, TS 36.331.
39. 3GPP, NR; Overall description; Stage-2, TS 38.300.



Voice over LTE und NR (VoLTE, VoNR)

3

Wie später in Kap. 5 gezeigt, ist in 2G GSM Netzwerken die Sprachtelefonie komplett im Netz integriert. Auch bei UMTS war dies teilweise noch der Fall. Mit LTE entschloss man sich jedoch, das Netzwerk und den Sprachdienst, von wenigen Ausnahmen abgesehen, komplett unabhängig voneinander zu halten. Aus diesem Grund kam auch die Beschreibung der LTE und 5G NR Netzarchitektur in den vorangegangenen Kapiteln ganz ohne die Behandlung der Sprachtelefonie aus.

Die Sprachtelefonie ist jedoch nach wie vor ein wichtiger Dienst, und die in Kap. 1 vorgestellte CS-Fallback Lösung stellte nur eine Zwischenlösung auf dem Weg zum mobilen „All-IP“ Netzwerk dar, in der auch die Sprachtelefonie in LTE über das Internet Protokoll abgewickelt wird. Dieses Ziel wurde nach einigen Jahren dann in den meisten Netzwerken mit VoLTE (Voice over LTE), das auf dem SIP (Session Initiation Protocol) basierten 3GPP IP Multimedia Subsystem (IMS) und dem GSMA IR.92 Profil¹ aufsetzt, erreicht. Für 5G NR wurde der VoLTE Dienst mit nur wenigen Änderungen übernommen und wird dort als Voice over NR bezeichnet. In diesem Kapitel wird der Begriff VoLTE verwendet, bezieht sich jedoch nicht nur auf LTE, sondern auch auf die Sprachtelefonie im 5G NR Netzwerk. Nur für die Beschreibung von 5G NR spezifischen Erweiterungen wird der Begriff VoNR verwendet.

Auch heute haben GSM Netzwerke in der Praxis noch eine bessere Flächenabdeckung als LTE und somit ist weiterhin ein Rückfall auf die klassische leitungsvermittelnde Sprachtelefonie während eines Gesprächs notwendig. Diese Funktionalität, die Single Radio Voice Call Continuity (SRVCC) genannt wird, wird ebenfalls beschrieben.

Zusätzlich zur Sprachtelefonie über LTE (VoLTE) bieten manche Netzbetreiber ihren Sprachdienst mit Voice over Wifi (VoWifi) auch in privaten und öffentlichen WLAN Netzwerken an. Dieser Dienst verwendet das gleiche IMS Netzwerk wie VoLTE und Gespräche können zwischen LTE und Wifi übergeben werden. Dies ist z. B. ein großer Vorteil gegenüber anderen Sprachdiensten, bei denen ein Gespräch unterbrochen wird oder sogar abbricht, wenn ein Endgerät seine Internetverbindung zwischen LTE und Wifi wechselt.

3.1 Das Session Initiation Protocol (SIP)

Ein Telefoniedienst hat zwei grundsätzliche Aufgaben. Möchte ein Nutzer einen anderen Nutzer anrufen, muss der Dienst den Zielteilnehmer finden und ihn über den eingehenden Anruf informieren. Die zweite Aufgabe besteht darin, zwischen den Nutzern eine direkte oder indirekte Verbindung (Session) aufzubauen. Im Falle der Sprachtelefonie wird in dieser Session dann in beiden Richtungen ein Sprachdatenstrom übertragen. In der Praxis hat sich sowohl bei großen Netzbetreibern als auch für Telefonanlagen in Firmen das Session Initiation Protocol (SIP) für diese Aufgabe durchgesetzt. Eine Open Source Implementation einer SIP Telefonanlage ist beispielsweise die Asterisk Platform².

SIP ist ein sehr generisches Protokoll und eignet sich somit für das Herstellen einer Verbindung zwischen zwei oder mehr Teilnehmern, über die dann die unterschiedlichsten Arten von Informationen fließen können. Nachfolgend wird hauptsächlich der Einsatz von SIP für das Herstellen einer Sprachverbindung betrachtet. Details finden sich in der IETF RFC 3261³ Spezifikation, die, wie auch die 3GPP Spezifikationen, frei im Internet zugänglich ist.

Wie in Abb. 3.1 gezeigt, bildet der Kern eines SIP basierten Telefoniesystems der SIP Registrar und der SIP Proxy Dienst. Um erreichbar zu sein und auch um abgehende Gespräche führen zu können, muss sich ein Endgerät nach dem Einschalten zunächst am SIP System registrieren. Die SIP Software Komponente auf dem Endgerät wird als User Agent (UA) bezeichnet. Auf der Netzwerkseite ist der SIP Registrar für das Registrieren und die Authentifizierung von Teilnehmern verantwortlich. Abb. 3.2 zeigt,

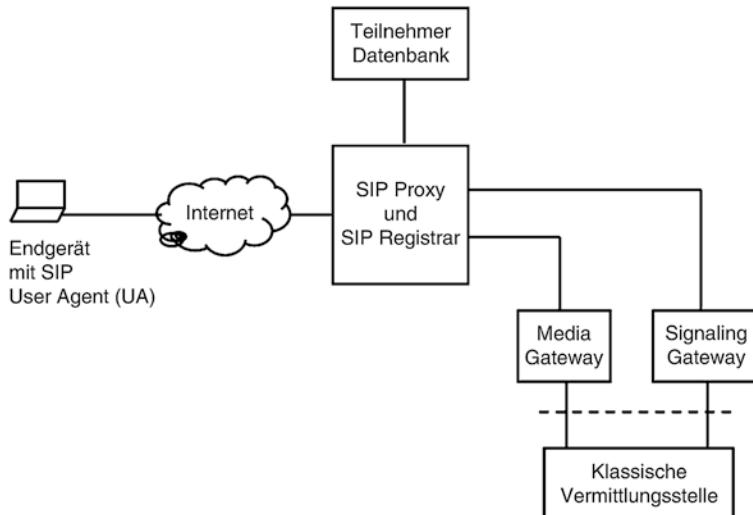


Abb. 3.1 SIP Infrastruktur

wie eine Registrierung in der Praxis durchgeführt wird. Zu Beginn sendet das Endgerät eine Anfrage an den DNS Server, um die IP-Adresse des SIP Registrar Servers zu ermitteln, dessen Domain Name zusammen mit der SIP-ID des Teilnehmers im Endgerät konfiguriert wurde. Danach sendet der User Agent eine erste ‚Register‘ Nachricht an den Registrar. Dieser sucht dann in der Teilnehmerdatenbank nach der SIP-ID des Teilnehmers, sowie dessen Authentifizierungsinformationen und fordert daraufhin den UA zu Authentifizierung mit einer ‚SIP 401 Unauthorized‘ Nachricht auf. Wie zuvor für andere Systeme beschrieben, basiert die Authentifizierung auf einem gemeinsamen Schlüssel/Passwort und einem Algorithmus, der auf beiden Seiten auf eine Zufallszahl angewandt wird. Da nur die Zufallszahl und das Ergebnis der Berechnung übertragen wird, nicht jedoch der gemeinsame Schlüssel, kann der Teilnehmer auch über eine unsichere Verbindung authentifiziert werden. Das Ergebnis der Berechnung sendet der User Agent an den Registrar Server mit einer neuen ‚Register‘ Nachricht zurück. War die Antwort korrekt, antwortet der Registrar Server mit einer ‚SIP 200 OK‘ Nachricht und der Teilnehmer ist im System angemeldet. Der Registrar speichert außerdem die

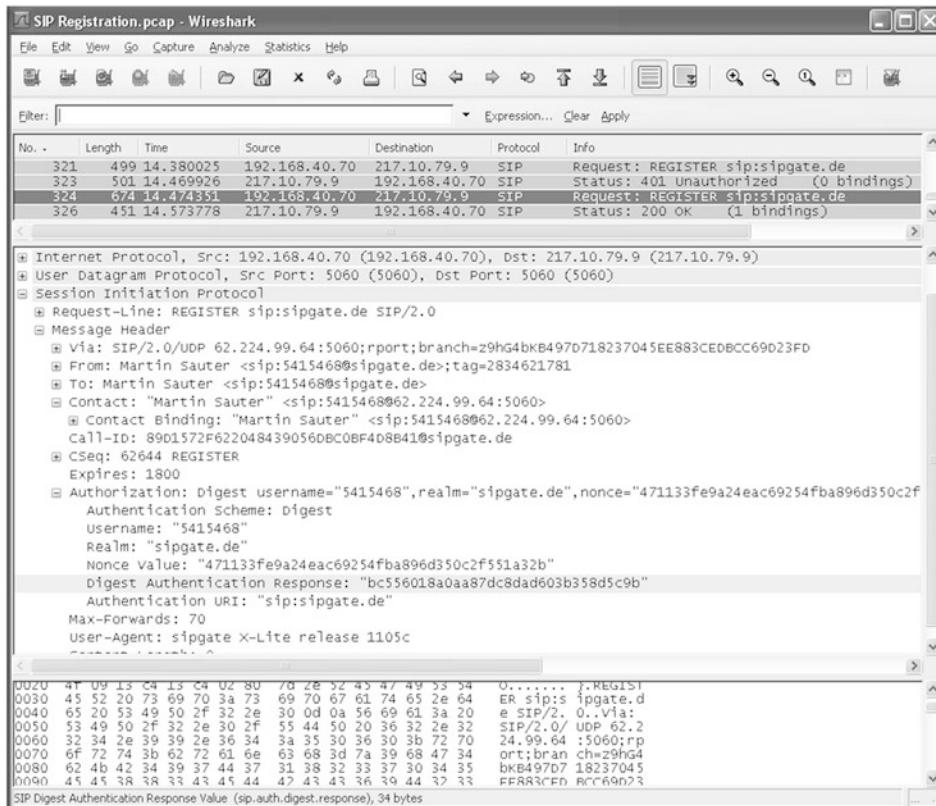


Abb. 3.2 Eine SIP Register Nachricht

IP-Adresse und den verwendeten UDP Port in seiner Teilnehmerdatenbank, damit er ankommende Verbindungen an das Endgerät weiterleiten kann.

An dieser Stelle sei angemerkt, dass sich die Zahlen in den Antwortnachrichten (401, 200, etc.) an den Zahlencodes der Antwortnachrichten des Hypertext Transfer Protocol (HTTP) orientieren, das für die Übertragung von Webseiten zwischen Webserver und Webbrower verwendet wird.

Abb. 3.2 zeigt eine SIP Register Nachricht nach einer „401 Unauthorized“ Antwort, die mit Wireshark aufgezeichnet wurde⁴. Zu sehen ist im mittleren Teil die SIP-ID des Teilnehmers (5415468) zusammen mit der SIP Domain (@sipgate.de), die zusammen den SIP Universal Resource Identifier (SIP URI) bilden. Im mittleren Teil der Nachricht ist dann die vom Netzwerk gesendete Zufallszahl (Nonce) zu sehen und das mit dem geheimen Schlüssel berechnete Ergebnis (Digest Authentication Response).

Weiterhin ist in Abb. 3.2 zu sehen, dass der Registrar Server in der „200 OK“ Status Nachricht dem User Agent meldet, dass „1 Binding“ vorhanden ist. Dies bedeutet, dass sich bisher nur dieser User Agent für diese Adresse registriert hat. Mit SIP ist es jedoch auch möglich, mehrere Endgeräte gleichzeitig auf eine sogenannte Public User ID zu registrieren. Bei ankommenden Anrufen werden dann beide Geräte benachrichtigt.

Nachdem der User Agent registriert ist, kann dieser dann jederzeit eine Sprachverbindung zu einem anderen Teilnehmer aufbauen und auch Gespräche von anderen Teilnehmern empfangen. Der Gesprächsaufbau wird in Abb. 3.3 gezeigt und läuft wie folgt ab:

Da der Nutzer zwar die SIP-ID des anderen Gesprächsteilnehmers kennt, nicht aber die IP-Adresse des anderen Endgeräts, läuft die Signalisierung über einen SIP Proxy Server im Netzwerk. Diesem teilt das Endgerät über eine SIP Invite Nachricht mit, dass eine Verbindung zu einem anderen Teilnehmer hergestellt werden soll. Der SIP Proxy authentifiziert den Teilnehmer dann zunächst mit einer „408 Authentication Request“ Antwort und sucht nach korrekter Antwort dann den Zielteilnehmer. Ist der Zielteilnehmer Kunde eines anderen Netzanbieters, wird die Invite Nachricht zum SIP Proxy des anderen Anbieters (SIP Proxy B) weitergeleitet. Dieser sucht dann in seiner Datenbank die IP Adresse des Endgerätes, das sich für diese SIP-ID registriert hat und leitet dann die Invite Nachricht weiter. Damit in entgegengesetzter Richtung ebenfalls wieder alle SIP Proxies durchlaufen werden, fügt jeder Proxy seine IP Adresse in die SIP Nachricht ein, bevor er diese weiterleitet.

Nachdem das Endgerät die Invite Nachricht empfangen hat, antwortet es mit einer „100 Trying“ Nachricht und trifft alle Vorbereitungen das Gespräch anzunehmen. Nachdem es zur Rufannahme bereit ist, sendet es eine „180 Ringing“ Nachricht und signalisiert somit dem Anrufer, dass der Zielteilnehmer über den eingehenden Anruf benachrichtigt wurde. Akzeptiert der Zielteilnehmer den Anruf, sendet das Endgerät eine „200 OK“ Nachricht über die zwei Proxy Server zum Anrufer und beide Endgeräte schalten dann den Sprachkanal auf ihre jeweiligen Lautsprecher und Mikrofone. Welcher Codec verwendet wird, bestimmen die Endgeräte untereinander, indem die Invite Nachricht eine Codec Liste des Endgeräts des rufenden Teilnehmers enthält. Das Endgerät des

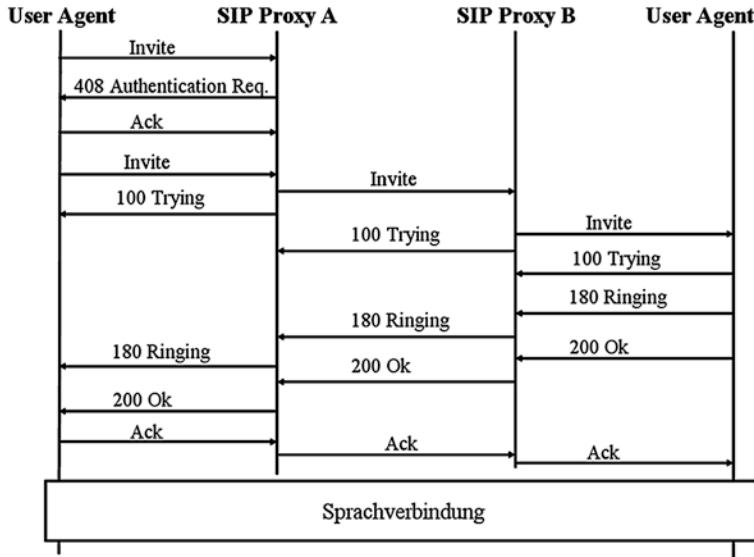


Abb. 3.3 SIP Gesprächsaufbau

Zielteilnehmers wählt aus dieser Codec Liste einen geeigneten Codec aus und teilt dies während dem Rufaufbau dann der Gegenstelle mit.

Während alle Signalisierungsnachrichten immer über die SIP Proxy Server laufen, kann der Sprachkanal direkt zwischen den Endgeräten aufgebaut werden. In der Praxis kommt es jedoch oft vor, dass die Endgeräte hinter Network Address Translation (NAT) Routern sitzen und somit nur lokale IP Adressen und UDP Ports verwenden, die im NAT Gateway dann in die zugeteilte globale IP Adresse und einen neuen UDP Port übersetzt wird. Ein direkter Austausch von Sprachpaketen ist somit nur möglich, wenn das Endgerät diese Adressumsetzung bemerkt und der Gegenstelle die globale IP Adresse und UDP Port mitteilen kann, an die der Sprachdatenstrom gesendet werden soll. Da diese für den User Agent nicht sichtbar sind, sendet dieser während des Registrierungsprozesses Probe Nachrichten an einen STUN (Session Traversal Utilities for NAT) Server im Internet. Dieser empfängt die Pakete nicht mit der dem User Agent bekannten lokalen IP Adresse und Port sondern mit der im NAT Gateway erzeugten neuen Kombination aus globaler IP Adresse und neuem UDP Port. Diese Information gibt der STUN Server an den User Agent zurück und dieser kann daraufhin ermitteln, wie IP Adresse und Port geändert wurden. Leider gibt es in der Praxis sehr unterschiedliche NAT Lösungen und es ist somit nicht immer möglich, eine eindeutige Regel für die IP-Adressen und Portänderungen zu finden. Deshalb verwenden SIP Provider auch oft Media Gateways. Jedes Endgerät sendet seinen Sprachdatenstrom zum Media Gateway und erhält auch Sprachdaten der Gegenrichtung nicht vom anderen Teilnehmer, sondern

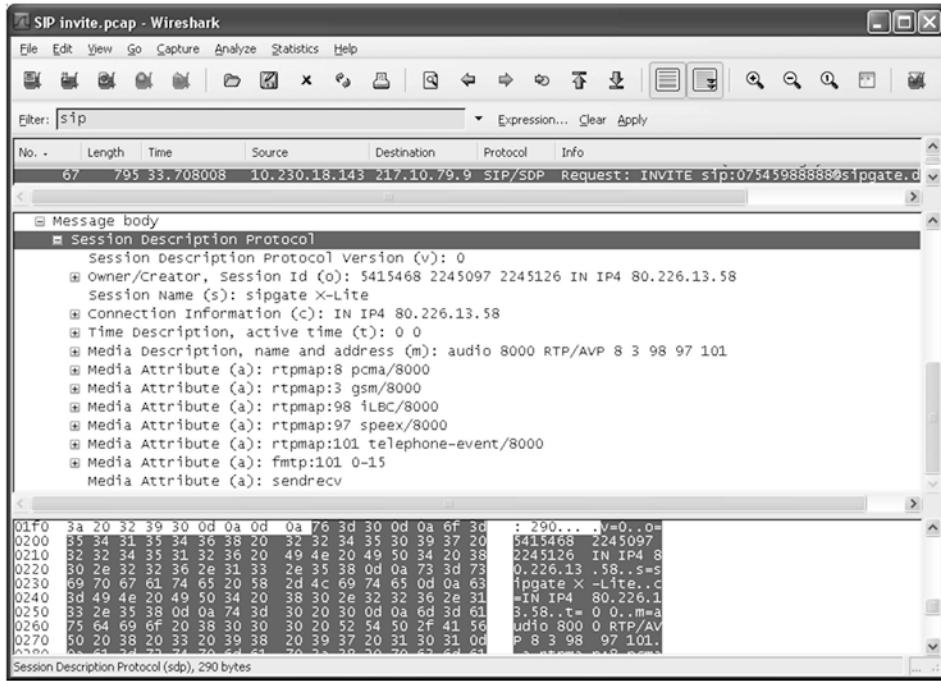


Abb. 3.4 Codec Liste im SDP Abschnitt einer SIP Invite Nachricht

vom Media Gateway. Somit muss statt zwei nur ein NAT Gateway überwunden werden, was in der Praxis problemlos möglich ist.

Unabhängig von direkter oder indirekter Sprachverbindung über ein Media Gateway wird das Real Time Transport Protocol (RTP) für die Sprachdatenübertragung verwendet, das in RFC 3550 spezifiziert ist⁵. Übliche Codecs sind in der Praxis entweder der aus der Festnetzwelt bekannte G.711 Codec oder, wenn von beiden Endgeräten unterstützt, der G.722 Wideband Codec mit einer besseren Sprachqualität. Beide Codecs wandeln die Sprachdaten in einen 64 kbit/s Datenstrom, der in 20 ms Sequenzen unterteilt in IP Paketen über UDP übertragen wird. Mit diesem Overhead ist die effektive Datenrate dann etwa 100 kbit/s pro Richtung. An dieser Stelle sei angemerkt, dass der G.722 Wideband Codec leider nicht mit dem im Mobilfunk verwendeten G.722.2 Wideband Codec kompatibel ist, da dieser nur eine Übertragungsrate von 12,65–23,85 kbit/s verwendet. Wideband Codec Gespräche zwischen Festnetz und Mobilfunknetzwerken sind somit nur mit einem Codec Wandler zwischen den Netzwerken möglich.

Abb. 3.4 zeigt, wie die Sprachcodecliste in einer Invite Nachricht zum Zielteilnehmer übertragen wird. In der SIP Nachricht ist diese Liste Teil des Session Description Protocol (SDP) Abschnitts, der in RFC 4566 spezifiziert ist⁶.

SIP Proxies können nicht nur Zielteilnehmer ausfindig machen und Nachrichten an andere Proxies oder Teilnehmer weiterleiten, sondern auch eigenständig modifizieren. Befindet sich ein Zielteilnehmer schon in einem anderen Gespräch und weist deshalb die Invite Nachricht mit einer Busy Nachricht ab, kann der SIP Proxy diese Nachricht verwerfen und stattdessen eine neue Invite Nachricht generieren, die dann zu einem zentralen Anrufbeantwortersystem geschickt wird.

Um auch mit Teilnehmern kommunizieren zu können, die über ein altes leitungsvermittelndes Festnetz erreichbar sind, können SIP Proxies eine Gesprächsanforderung, wie in Abb. 3.1 gezeigt, über ein Signaling Gateway und ein Media Gateway in ein nicht-SIP Netzwerk weiterleiten. Das Signaling Gateway übersetzt dazu die SIP Nachrichten in leitungsvermittelnde Signalisierung und umgekehrt, während das Media Gateway die Umwandlung zwischen Sprachdaten in IP Paketen und Sprachdaten in leitungsvermittelten Zeitschlitzten vornimmt.

3.2 Das IP Multimedia Subsystem (IMS) und VoLTE

3.2.1 Architekturüberblick

Für die Verwendung in Mobilfunknetzwerken wurde das SIP System erheblich von 3GPP erweitert und wird als IP Multimedia Subsystem (IMS) bezeichnet. Abb. 3.5 zeigt die zentralen Komponenten des IMS. Der Kern des Systems ist die Serving Call Session Control Function (S-CSCF), die die Rolle des SIP Registrars und SIP Proxys übernimmt. Um mit der zentralen Datenbank, dem Home Subscriber Server (HSS) zu

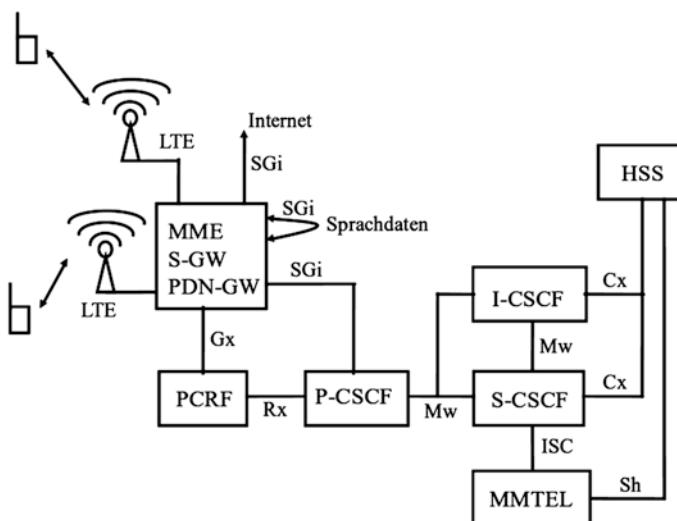


Abb. 3.5 Das IMS System im Überblick

kommunizieren, verwendet die S-CSCF die Cx-Schnittstelle und das Diameter Protokoll, das in RFC 6733 beschrieben wird⁷.

Da es mehrere S-CSCF in einem großen IMS System geben kann, wird eine Verteilungsfunktion für eingehende Anfragen benötigt. Diese Funktion übernimmt die Interrogating-CSCF (I-CSCF), die auch über das Cx-Interface mit dem HSS verbunden ist, da sie Teilnehmerinformationen benötigt, um Nachrichten an die zuständige S-CSCF weiterzuleiten. An der Grenze zum IMS System sitzt der Proxy-CSCF (P-CSCF), der eine wichtige Rolle spielt, da er nicht nur als SIP Proxy agiert, sondern auch den Nutzer gegenüber dem IMS System repräsentiert. Dies ist notwendig, da die Verbindung zwischen Netzwerk und mobilem Endgerät unterbrochen werden kann, wenn das Endgerät z. B. die Netzabdeckung verliert. In diesem Fall kann das Endgerät kein SIP ‚Bye‘ schicken, eine Session kann also nicht korrekt beendet werden. Diese Aufgabe übernimmt in diesem Szenario die P-CSCF, nachdem sie vom LTE Netzwerk darüber informiert wurde, dass der Kontakt zum Teilnehmer unterbrochen wurde. Zwischen den einzelnen CSCF Elementen wird die Mw-Schnittstelle und das SIP Protokoll für die Kommunikation verwendet. Ein Endgerät kommuniziert mit der P-CSCF über das LTE Netzwerk und den PDN-GW über die schon existierende SGi-Schnittstelle, die auch für den direkten Internetzugang verwendet wird.

Zusätzlich hat die P-CSCF auch die Aufgabe, die Quality of Service (QoS) für eine Sprachverbindung sicherzustellen. Zu diesem Zweck wird dem Endgerät ein logischer Dedicated Radio Bearer (DRB) während des Verbindungsbaus zugeteilt, auf dem dann die Sprachdatenpakete übertragen werden. Dieser Bearer wird im Netzwerk und auf der Luftschnittstelle bevorzugt, um, solange dies die Radiobedingungen zulassen, eine konstante Verzögerungszeit zu garantieren und Paketverluste zu vermeiden. Zu diesem Zweck ist die P-CSCF, wie in Abb. 3.5 gezeigt, mit der Policy and Charging Rules Function (PCRF) über die Rx-Schnittstelle verbunden, über die die QoS Anforderungen des IMS Systems für den Sprachpfad in Kommandos für das LTE Netzwerk übersetzt werden, die dann über die Gx-Schnittstelle weitergegeben werden.

Des Weiteren gibt es im IMS System Application Server (AS), mit denen das System erweitert werden kann. Application Server können z. B. den Aufbau und die Aufrechterhaltung einer Session kontrollieren, in dem sie vom S-CSCF erhaltene SIP Nachrichten modifizieren. Um flexibel zu sein, können mehrere Application Server für die Weiterleitung in einem Nutzerprofil eingetragen sein. Für VoLTE wird ein Application Server verwendet, der die MMTel (Multimedia Telephony) Spezifikation aus 3GPP TS 22.173 umsetzt, und ist somit zuständig für die typischen zusätzlichen Telefoniefunktionen (Supplementary Services) wie Anrufweiterleitung, Konferenzgespräche, Gespräch halten, Unterdrückung der Telefonnummer des Anrufers, etc.⁸

3.2.2 IMS Registrierung

Ähnlich wie ein einfaches zuvor beschriebenes SIP Endgerät verwendet ein IMS VoLTE Endgerät eine SIP „Register“ Nachricht, um sich beim IMS System nach dem Einschalten anzumelden. Das Endgerät könnte dazu die bereits vorhandene Internetverbindung nutzen. Dies wird aber in der Praxis nicht gemacht. Stattdessen fragt das Endgerät nach dem Einschalten nach einem separaten Default Bearer für VoLTE wie in Abb. 3.6 gezeigt. Dieser separate Bearer hat seine eigenen IPv4 und/oder IPv6 Adressen und das Netzwerk informiert das Endgerät in der „Activate Default EPS Bearer Context Request“ Nachricht (vgl. Kap. 1) über die IP Adressen der vorhandenen P-CSCFs. Der standardisierte APN Name, den das Endgerät und das Netzwerk dazu verwenden, lautet „ims“. Das Konzept eines Default Bearers ist zwar etwas abstrakt, man kann sich diesen jedoch als virtuelles Netzwerkinterface vorstellen, von denen mehrere verwendet werden, um über die Luftschnittstelle mit dem Netzwerk zu kommunizieren. Da der IMS Bearer seine eigene IP-Adresse hat, können alle IP Pakete für das IMS System über dieses virtuelle Netzwerkinterface geleitet werden, während alle anderen IP Pakete das andere virtuelle Netzwerkinterface für den Internetzugang verwenden. Apps auf mobilen

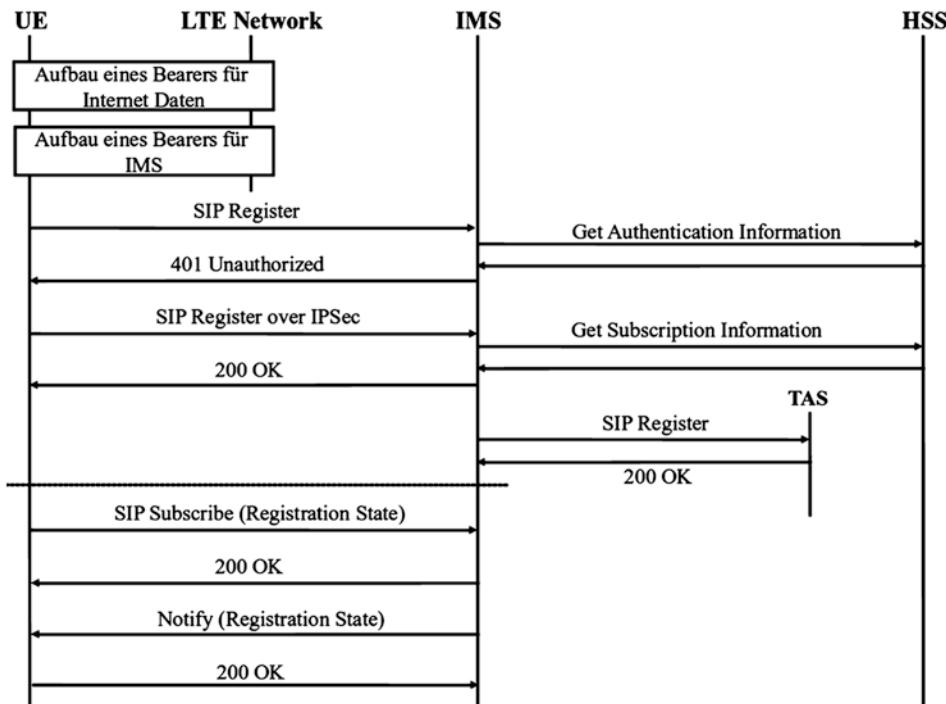


Abb. 3.6 Die IMS Registrierungsprozedur

Endgeräten können ohne erweiterte Zugriffsrechte nur auf das virtuelle Netzinterface für den Internetzugang zugreifen.

Nachdem der IMS LTE Bearer aufgebaut ist, können SIP Nachrichten gesendet und empfangen werden. Zusätzlich zur zuvor beschriebenen SIP Registrierungsprozedur sind im IMS System zusätzliche Aktionen für eine vollständige Registrierung notwendig. Um flexibel zu sein, brauchen im Endgerät keine Konfigurationsinformationen über eine oder mehrere P-CSCF vorhanden zu sein, da das Endgerät deren IP-Adressen während des IMS Default Bearer Activation Prozesses erhält. Zusätzlich informiert das Netzwerk das Endgerät während des ersten Verbindungsbaus zum Netzwerk, sowie bei Tracking Area Updates, ob das LTE Netzwerk den IMS Service unterstützt. Ist dies nicht der Fall, wird der im LTE Kapitel beschriebene CS-Fallback Mechanismus verwendet, um ein GSM oder UMTS Netzwerk für ein Telefonat zu verwenden. Hier kommt dann das klassische leitungsvermittelnde Telefoniesystem zum Einsatz.

Wie in anderen Teilen des Netzwerkes wird die IMSI (International Mobile Subscriber Identity) verwendet, um den Nutzer zu identifizieren. Die IMSI ist auf der SIM Karte gespeichert und es ist somit ebenfalls nicht nötig, eine Nutzer-ID für VoLTE im Endgerät zu konfigurieren.

Um die Übertragung von SIP Nachrichten zwischen Endgerät und der P-CSCF zu schützen, wird während der SIP Registrierung ein Security Context aufgebaut, für dessen Aufbau der geheime Schlüssel Ki und der Authentifizierungsmechanismus auf der SIM Karte verwendet werden, die auch in GSM, UMTS, LTE und 5G NR Anwendung finden. Um sicherzustellen, dass Nachrichten nicht absichtlich oder durch Übertragungsfehler verändert wurden, wird eine Integritätschecksumme an jede Nachricht angefügt. Optional können SIP Nachrichten zwischen Endgerät und der P-CSCF auch verschlüsselt werden.

Abb. 3.6 zeigt, wie der IMS Registrierungsprozess abläuft. Wie im SIP Festnetzbeispiel am Anfang des Kapitels sendet das Endgerät als erstes eine SIP „Register“ Nachricht. Wichtige Parameter in dieser Nachricht sind die IMSI, die unterstützten Authentifizierungs- und Verschlüsselungsalgorithmen, die Modellbezeichnung des Endgerätes, Softwareversion, Seriennummer (IMEI, International Mobile Equipment Identity) und welche IMS Services unterstützt werden. Typischerweise unterstützen VoLTE Geräte den MMTEL Dienst für Sprachtelefonie, sowie SMS und IMS. Außerdem enthält die SIP Nachricht den Namen des IMS Systems, das kontaktiert werden soll. Der Name setzt sich unter anderem aus dem Mobile Country Code (MCC) und dem Mobile Network Code (MNC) des Heimnetzwerkes des Teilnehmers zusammen, z. B. „ims.mnc002.mcc262.3gppnetwork.org“. Dieser Name wird von der P-CSCF verwendet, um über eine DNS Anfrage die IP Adresse des IMS Systems zu ermitteln.

Nachdem die S-CSCF die SIP „Register“ Nachricht erhalten hat, verwendet sie die IMSI des Teilnehmers, um dessen Nutzereintrag im Home Subscriber Server (HSS) zu finden, und sendet dann ein „401 Unauthorized“ an das Endgerät mit einer Authentifizierungsaufforderung zurück. Die Authentifizierungsaufforderung wird dann vom Endgerät an die SIM Karte weitergereicht, und erhält dann die Parameter, mit dem ein IPSec Security Context zwischen dem Endgerät und der P-CSCF aufgebaut werden kann. Falls

vom Netzwerk verlangt, werden alle weiteren SIP Nachrichten dann auch verschlüsselt. Andernfalls wird der IPSec Tunnel nur für die erforderliche Integritätssicherung der Nachrichten verwendet. An dieser Stelle sei angemerkt, dass IPSec nur im Mobilfunk zwischen dem Endgerät und der P-CSCF verwendet wird und nicht Teil der eigentlichen SIP Spezifikation ist. Festnetz SIP Systeme verwenden üblicherweise keine IPSec Verschlüsselung und Integritätssicherung für SIP Nachrichten.

Das Endgerät Antwortet auf die SIP ‚401 Unauthorized‘ mit einer neuen SIP ‚Register‘ Nachricht und den erforderlichen Authentifizierungsinformationen. Wenn das Netzwerk den Teilnehmer damit erfolgreich authentifizieren kann, schickt es ein SIP ‚200 OK‘ zurück. Zusätzlich fordert die S-CSCF eine Kopie des IMS Teilnehmerprofils vom HSS an, das unter anderem die Information enthält, welche Application Server über die erfolgreiche Registrierung informiert werden sollen. Bei VoLTE ist dies üblicherweise der Telephony Application Server (TAS), der die MMTel Funktionen implementiert.

Im letzten Schritt abonniert das Endgerät sogenannte ‚Registration Notification Events‘. Über diese Events kann das IMS System das Endgerät z. B. über Vorgänge wie einen Server Shutdown informieren und die SIP Registrierung beenden. Der Event Mechanismus wird auch verwendet, wenn sich ein anderes Gerät für die gleiche Public User Identity (Telefonnummer) registriert. Auf diese Weise ist es möglich, einen Multi-SIM Service anzubieten, mit dem der Kunde mehrere Endgeräte mit der gleichen Telefonnummer verwenden kann.

3.2.3 Der VoLTE Gesprächsaufbau

Nach erfolgreicher Registrierung können dann eingehende oder ausgehende Telefongespräche geführt werden. Dieses Unterkapitel gibt nun einen Überblick über diesen Prozess, in weiteren Unterkapiteln werden dann weitere Details betrachtet. Abb. 3.7 und 3.8 zeigen, welche Nachrichten nötig sind, um eine Session (Telefongespräch) zwischen zwei IMS Endgeräten aufzubauen. Die Rolle des MMTel Servers wird in den Abbildungen nicht gezeigt, da in diesen Szenarien keine wesentlichen Parameter durch die MMTel in den Nachrichten geändert werden. Dies wäre nur bei Prozeduren wie der Gesprächsweiterleitung, z. B. zur Sprachbox, bei Nichterreichbarkeit nötig.

Im ersten Schritt sendet das Endgerät eine SIP ‚Invite‘ Nachricht an den SIP Proxy, in diesem Fall an die P-CSCF. Die P-CSCF bestätigt dem Endgerät zunächst die SIP Nachricht mit einer SIP ‚100 Trying‘ Nachricht und leitet die ‚Invite‘ Nachricht dann zur S-CSCF weiter. Hier wird die Nachricht dann ausgewertet und zu einer I-CSCF weitergeleitet, die dann ermittelt, welche S-CSCF für den Zielteilnehmer verantwortlich ist. Die Nachricht wird dann entsprechend weitergeleitet. Die S-CSCF des Zielteilnehmers kann sich entweder im gleichen Netzwerk befinden oder im Netzwerk eines anderen Mobilfunkbetreibers. Im zweiten Fall werden Border Gateway Controller (BGC) verwendet, um die zwei eigenständigen IMS Netzwerke miteinander zu verbinden. Nachdem eine SIP Nachricht beim S-CSCF des Zielteilnehmers eingegangen ist, wird dann

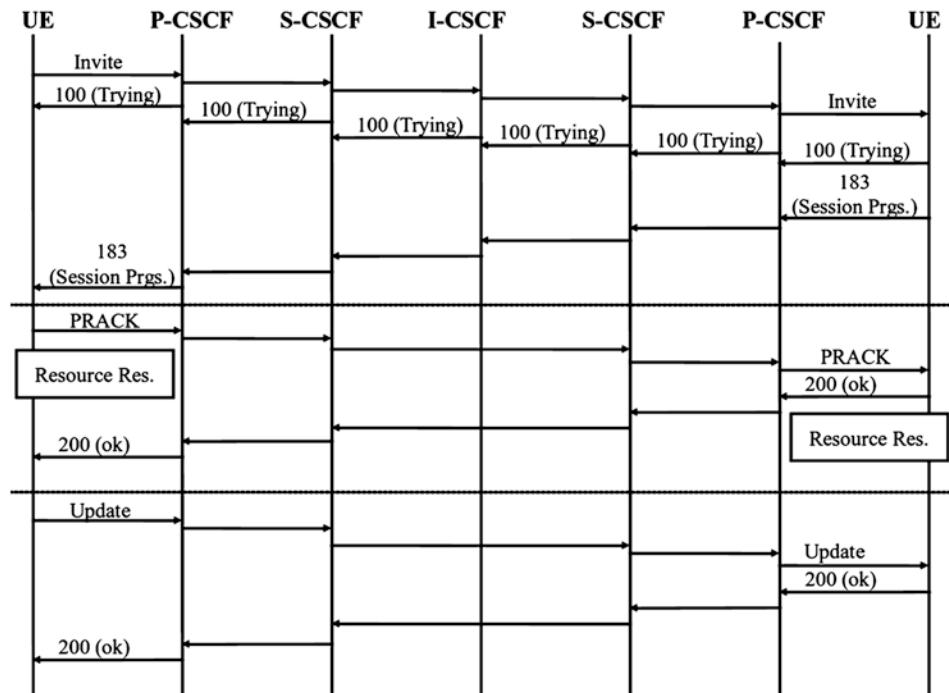


Abb. 3.7 Der VoLTE Gesprächsaufbau, Teil 1

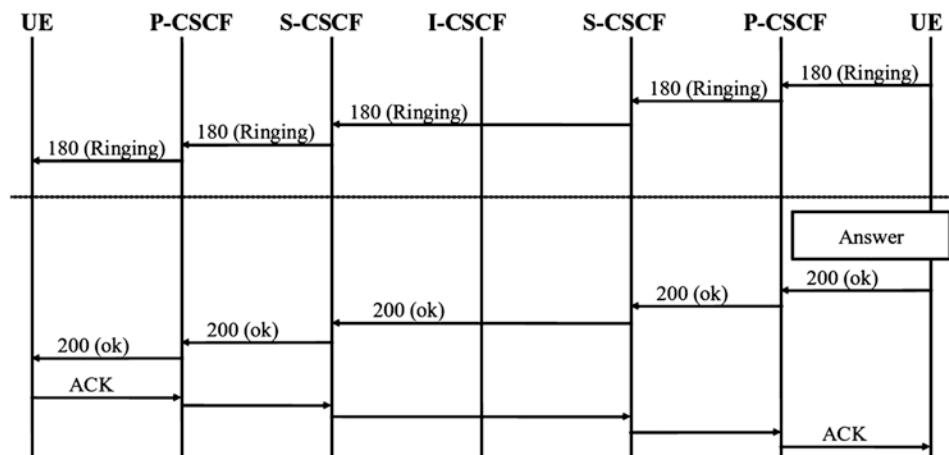


Abb. 3.8 Der VoLTE Gesprächsaufbau, Teil 2

die P-CSCF des Zielteilnehmers ermittelt und die SIP Nachricht dorthin weitergeleitet. Die P-CSCF versucht dann, die ‚Invite‘ Nachricht an das Endgerät weiterzuleiten. Falls dies gelingt, antwortet das Endgerät des Zielteilnehmers zuerst mit einer SIP ‚100 Trying‘ Nachricht und dann mit einer SIP ‚Session Progress‘ Nachricht an die P-CSCF, die dann über alle IMS Knoten zum Endgerät auf der anderen Seite weitergeleitet wird.

Unter anderem enthalten die SIP ‚Invite‘ und SIP ‚Session Progress‘ Nachrichten den am Anfang des Kapitel beschriebenen Session Description Protocol (SDP) Teil, der alle Sprachcodes auflistet, die das Endgerät unterstützt. Dies ist notwendig, um einen Sprachcodec zu wählen, für den dann geeignete Quality of Service Eigenschaften auf der Luftschnittstelle sichergestellt werden. Diese Aufgabe fällt dem P-CSCF zu, der das LTE Netzwerk über die PCRF anweist, einen Dedicated Bearer für die Sprachpakete aufzubauen. Dies ist möglich, da die P-CSCF die SIP Nachrichten nicht nur weiterleitet, sondern auch analysiert und entsprechend handelt.

Nachdem Ressourcen für den Sprachkanal zugewiesen wurden, sendet das Endgerät dann wiederum über alle SIP Router eine SIP ‚180 Ringing‘ Nachricht zur anderen Seite, während die IP Sprachdatenpakete direkt zwischen den Endgeräten ausgetauscht werden können. Falls sich beide Endgeräte im gleichen Netzwerk befinden, können die Sprachpakete direkt am PDN-Gateway in das Netzwerk zurückgesendet werden (vgl. Abb. 3.5). An dieser Stelle sei angemerkt, dass Sprachpakete üblicherweise über ein Sprachgateway geleitet werden, um einen schnelleren Handover des Gesprächs von LTE nach UMTS oder GSM durchführen zu können, falls der Teilnehmer den LTE Netzbereich verlässt. Dies ist Teil der 3GPP Release 10 Single Radio Voice Call Continuity (SRVCC) Funktion, die nachfolgend noch genauer beschrieben wird.

Da die 3GPP IMS Spezifikationen sehr viele Optionen enthalten, deren komplette Umsetzung das System sehr komplex machen würden, legten sich die Netzbetreiber mit der GSMA IR.92⁹ Spezifikation auf ein Subset fest. Dies stellt sicher, dass der IMS Sprachdienst zwischen den Netzen und zwischen Endgeräten und Netzen kompatibel ist. Wie alle 3GPP und RFC Spezifikationen ist auch IR.92 frei im Internet zugänglich. Für sich selbst gesehen ist die IR.92 Spezifikation nicht für eine Einführung in IMS und VoLTE geeignet, sie enthält jedoch viele Referenzen auf relevante 3GPP Spezifikationen, die für das tiefere Verständnis von VoLTE sehr nützlich sind.

3.2.4 LTE Bearer Konfiguration für VoLTE

Da VoLTE eine Schnittstelle zum LTE Netzwerk hat, kann das System mit dem Transportnetzwerk kommunizieren und somit für Gespräche eine möglichst kleine Latenzzeit und konstante Bandbreite anfordern. Dies hilft auch dem eNodeB, die Übertragung von VoIP Paketen über die Luftschnittstelle zu optimieren. Für VoLTE werden dazu während der Gesprächsaufbauphase folgende Optimierungen aktiviert:

In LTE wird für die Sprachdatenpakete ein zusätzlicher Dedicated Bearer aktiviert, mit dem die QoS für die Sprachdatenpakete garantiert wird. Ein Dedicated Bearer wird

vom Netzwerk aufgebaut, wenn ein Service im Netzwerk eine Priorisierung von IP Paketen einrichten möchte, die zu einem bestimmten Datenstrom gehören, der zwischen zwei konfigurierbaren IP Adressen und/oder TCP/UDP Ports ausgetauscht wird. Der Dedicated Bearer wird dann auf der Luftschnittstelle wie folgt umgesetzt:

Unacknowledged Radio Bearers for Voice

Für den Sprachdatenstrom wird ein Unacknowledged Mode Data Radio Bearer (UM-DRB) verwendet und in einer RRC Connection Reconfiguration Nachricht, wie nachfolgend gezeigt, konfiguriert. Der UM-DRB wird auf Layer 3 des RLC Protokolls auf der Luftschnittstelle umgesetzt. Auf dieser Schicht werden normalerweise fehlende Pakete automatisch wiederholt (AM, Acknowledged Mode). Für Sprachdaten wird hier jedoch der Unacknowledged Mode konfiguriert, da es keinen Sinn macht, fehlende Sprachdatenpakete zu wiederholen, da diese zu spät beim Empfänger eintreffen würden und somit nutzlos sind. Neben dem Dedicated Bearer, der als UM-DRB für die Sprachdatenpakete konfiguriert ist, sind natürlich der Signaling Bearer (SRB) und die anderen Default Bearer (DRBs für IMS Signalisierung und Internetzugang) weiterhin aktiv.

Der UM-DRB für den Sprachdatenstrom wird von IMS System während der Gesprächsaufbauphase durch eine Anforderung an das Transportnetzwerk aufgebaut. Dieser ist Teil des Dedicated Bearers für die IP Sprachdatenpakete, die zwischen zwei IP Adressen und UDP Ports ausgetauscht werden. Dieser Datenstrom wird dann auf einem Radio Bearer übertragen, für den keine RLC Fehlerkorrektur verwendet wird. Alle anderen IP Pakete, die dem IP und UDP Filter nicht entsprechen, werden über die normalen Default Bearer geleitet, die im AM-DRB Modus ohne garantierte Verzögerung und Bandbreite arbeiten.

An dieser Stelle sei angemerkt, dass der Dedicated Bearer für die Sprachdatenpakete keine eigene IP Adresse besitzt. Dieser teilt die IP Adresse mit dem dazugehörigen Default Bearer und es entscheidet nur die Kombination aus Quell- und Ziel- IP Adresse, sowie Quell- und Ziel UDP Port, ob ein Datenpaket über den AM-DRB oder UM-DRB übertragen wird. Für die VoLTE Applikation ist dies jedoch völlig transparent, da das Trennen und Zusammenführen der IP-Pakete Aufgabe des IP- und des Radioprotokollstacks ist.

Paketverlust und Garantierte Bitrate

Um sicherzustellen, dass die Paketverlustrate für den Unacknowledged Mode Bearer unterhalb von 1 % bleibt, werden die Übertragungscharakteristiken (Sendeleistung, Modulation, Kodierung, ...) entsprechend angepasst. Dieser Paketverlust ist für VoLTE noch akzeptabel, da der Sprachcodec diesen noch ohne Probleme ausgleichen kann.

Zusätzlich wird der UM-DRB Bearer für die Sprachdatenpakete mit einer garantierten Bitrate konfiguriert und alle Netzelemente stellen sicher, dass zu jeder Zeit genug Bandbreite zur Verfügung steht, um Daten mit dieser Datenrate zu übertragen. Auf der Luftschnittstelle kann dies z. B., wie in Kap. 1 beschrieben, durch Semi-Persistentes Scheduling erreicht werden. Bei dieser Übertragungsart kann das Endgerät zu festen

Zeitpunkten eine vereinbarte Menge an Daten übertragen, ohne dass es beim Netzwerk für Ressourcen anfragen oder auf Ressourcenzuteilungen warten muss. Diese Art des Schedulings für Sprachdaten ist auch für das Netzwerk vorteilhaft, da hier der Overhead für ständige Uplinkzuweisungen entfallen kann.

Für ein VoLTE Gespräch wird der Dedicated Bearer immer vom Netzwerk aufgebaut und das Endgerät mit einer ‚Activate Dedicated EPS Bearer Context Request‘ Nachricht informiert. Der nachfolgende Ausschnitt aus einer solchen Nachricht zeigt die wichtigsten Parameter und deren Werte. Die Werte in Klammern zeigen die Werte für den Default Bearer, der für den Internetzugang verwendet wird. Dedicated Bearer für den VoLTE Sprachpfad verwenden den Quality of Service Identifier (QCI) 1 statt QCI 9, der für den Internet Default Bearer verwendet wird. QCI 1 ist in 3GPP standardisiert und weist die Basisstation (eNodeB) an, Pakete in diesem Bearer bei Überlast für maximal 100 ms statt 300 ms wie für einen QCI 9 Bearer zu puffern. Wie zuvor beschrieben bedeutet QCI 1 außerdem, dass auf dem RLC-Layer kein Acknowledgement aktiviert werden soll. Zusätzlich enthält die Dedicated Bearer Setup Nachricht an den eNodeB die Information, für diesen Bearer eine Bitrate von 40 kbit/s zu garantieren. Da der Dedicated Bearer nur für Sprachpakete verwendet werden soll, werden vier Traffic Flow Templates (TFT) übergeben, die als Filterregeln für den eNodeB und das Endgerät dienen. Diese Regeln legen fest, dass über den Dedicated Bearer nur Datenpakete zwischen zwei festgelegten IP Adressen übertragen werden sollen, sowie für diese IP Adressen nur zwischen zwei definierten UDP Portnummern. Über eine dieser Portkombinationen wird der RTP (Realtime Transport Protocol) Sprachdatenstrom übertragen, während über die Andere RTCP (Realtime Control Protocol) Pakete übertragen werden, die für das Ende zu Ende Management der Sprachverbindung nötig sind.

```

EPS Quality of Service
    Quality of Service Class Identifier: QCI 1 [QCI 9]
    Maximum bit rate for uplink: 40 kbps [default = 0]
    Maximum bit rate for downlink: 40 kbps [default = 0]
    Guaranteed bit rate for uplink: 40 kbps [default = 0]
    Guaranteed bit rate for downlink: 40 kbps [default = 0]
    Traffic Flow Templates für den Downlink
        Remote IPv6 address: 2304:724:610:4221::8
        Local UDP Port: 1254, Remote Port: 60002
        Local UDP Port: 1255, Remote Port: 60003
    Traffic Flow Templates für den Uplink
        Remote IPv6 address: 2304:724:610:4221::8
        Remote UDP Port: 60002
        Remote UDP Port: 60003
Ausgetauschter QoS
    Precedence Class: Normal priority [default = low priority]
    Mean Throughput: Best effort

```

```

Traffic Class: Conversational [default = background]
Transfer Delay: 100 ms [default = 300 ms]
Reliability: Unacknowledged GTP/LLC/RLC
[default = Unack. GTP/LLC, Ack RLC]

```

3.2.5 Dedicated Bearer Setup mit Preconditions

Eine Möglichkeit, ein Gespräch erst dann zu beginnen, nachdem die Quality of Service Einstellungen im gesamten Netzwerk hergestellt wurden, ist die Verwendung des SIP ‚Precondition‘ (Vorbedingungs-) Mechanismus während des Gesprächsaufbaus. 3GPP TS 23.490 Abschn. 3.3.1 zeigt den dazugehörigen SIP Message Flow¹⁰. RFC 3312 beschreibt zusätzlich die Grundlagen des Precondition Mechanismus¹¹. Der nachfolgende gekürzte Message Flow, der schon in Abb. 3.7 und 3.8 dargestellt ist, zeigt, welche SIP Nachrichten Precondition Information enthalten. Diese sind mit einem * gekennzeichnet. Die Zahlen am Anfang jeder Zeile entsprechenden Nachrichtennummern, die in 3GPP 23.490 verwendet werden.

```

--> Endgerät zum Netzwerk
<-- Network zum Endgerät
1 --> * Invite
2 <-- 100 Trying
16 <-- * 183 Session Progress
17 --> PRACK (Provisional Acknowledgement)
24 <-- 200 OK
25 --> * Update
32 <-- * 200 OK
36 <-- 180 Ringing
40 <-- 200

```

Der Precondition Aufbau beginnt mit der ersten SIP Nachricht (Invite) in der das Endgerät dem Netzwerk und der Gegenstelle mitteilt, dass es den Precondition Mechanismus für den Aufbau des Sprachpfades unterstützt. In der Praxis wird dies durch folgende Header Statements und Attributes (a=) im Session Description (SDP) Teil der SIP Nachricht mitgeteilt:

```

Supported: precondition
a=curr:qos local none
a=curr:qos remote none
a=des:qos mandatory local sendrecv
a=des:qos none remote sendrecv
a=sendrecv

```

Zu diesem Zeitpunkt des Gesprächsaufbaus gibt es weder eine lokale (local) Ressourcen Reservierung (Quality of Service, QoS) noch auf der Gegenseite (remote). Das Endgerät teilt dem Netzwerk und der Gegenseite jedoch mit, das QoS auf seiner Seite eine Voraussetzung ist (mandatory). Von der Gegenseite (remote) hat das Endgerät noch keine Informationen und setzt den Wert deswegen auf ‚none‘.

Darauf folgt Nachricht Nummer 16, eine SIP ‚Session Progress‘ Nachricht mit Informationen des Netzwerks und der Gegenseite mit den folgenden Header und QoS Zeilen:

```
Require: precondition
[...]
a=curr:qos local none
a=curr:qos remote none
a=des:qos mandatory local sendrecv
a=des:qos mandatory remote sendrecv
a=conf:qos remote sendrecv
a=sendrec
```

In dieser Nachricht teilt das Netzwerk dem Endgerät in den Header und ‚remote‘ Zeilen mit, dass es die Verwendung von ‚Preconditions‘ fordert. Zu diesem Zeitpunkt hat noch keiner der beiden Teilnehmer einen Dedicated Bearer (QoS) für die Verbindung. Die ‚a=conf:...‘ Zeile ist auch wichtig, da mit dieser die Gegenstelle eine Bestätigung vom lokalen Endgerät einfordert, die gesendet werden soll, sobald der Dedicated Bearer beim ihm aufgebaut ist. Erst bei Erhalt dieser Bestätigung wird die Gegenstelle den Teilnehmer über den eingehenden Anruf benachrichtigen.

In Nachricht 25 informiert das lokale Endgerät die Gegenstelle, dass es eine LTE RRC Nachricht für einen Dedicated Bearer erhalten hat (nicht im Message Flow oben gezeigt, da keine SIP Nachricht) und somit QoS auf dieser Seite der Verbindung nun aktiviert ist (wie in der conf(irm) Zeile in Nachricht 16 gefordert):

```
a=curr:qos local sendrecv
a=curr:qos remote none
a=des:qos mandatory local sendrecv
a=des:qos mandatory remote sendrecv
a=sendrecv
```

Und schließlich antwortet die Gegenstelle in Nachricht 32, dass auch bei ihr Ressourcen für QoS zugewiesen wurden und das Telefongespräch nun beginnen kann.

```
a=curr:qos local sendrecv
a=curr:qos remote sendrecv
a=des:qos mandatory local sendrecv
a=des:qos mandatory remote sendrecv
```

Die Gegenstelle informiert dann den Nutzer über den eingehenden Anruf in Nachricht 36 ,180 Ringing‘. Nachdem der entfernte Teilnehmer den Anruf annimmt, sendet sein Endgerät eine SIP ,200 OK‘ Nachricht und die zwei Teilnehmer können über den durch QoS garantierten Dedicated Bearer auf beiden Seiten kommunizieren.

Falls das Netzwerk keine Precondition Informationen in Nachricht 16 einfügt (,183 Session Progress‘), wird die Ende-zu-Ende Verbindung ohne den Precondition Mechanismus aufgebaut. Das bedeutet, dass die Nachrichten 25 und 32 nicht gesendet werden, nachdem der Dedicated Bearer, der trotzdem aufgebaut wird, aktiv ist. Nachdem in den ersten Jahren der Precondition Mechanismus typischerweise genutzt wurde, verzichten heute manche Netzbetreiber auf diese Signalisierung und beschleunigen somit den Gesprächsaufbau.

3.2.6 Header Compression und DRX

Die größte Ineffizienz eines VoIP Datenstroms ist der Overhead des IP Protokollstacks in jedem Paket. Um diesen Overhead zu reduzieren, kann Robust Header Compression (RoHC) für die Übertragung der Sprachpakete über die Luftschnittstelle verwendet werden. Dieser Mechanismus wurde schon in Kap. 1 als Teil des PDCP Header Compression Mechanismus beschrieben.

Außerdem ist es wichtig, den Energieverbrauch in Zeiten während eines Anrufs zu reduzieren, in denen keine Daten gesendet oder empfangen werden. Dies kann mit dem DRX (Discontinuous Reception) Mechanismus erreicht werden. Dies ist vor allem für Sprachverbindungen wichtig, da die Zeit zwischen zwei IP Paketen, die Sprachdaten enthalten, sehr lange ist. Wird der Transceiver im Endgerät in diesen Zeiten abgeschaltet, verlängert dies die Laufzeit bis zum nächsten Ladevorgang.

RoHC und DRX werden in einer RRCConnectionReconfiguration Nachricht während des Dedicated Bearer Aufbaus aktiviert. Der nachfolgende Ausschnitt einer solchen Nachricht zeigt, wie diese Parameter im Vergleich zu einem normalen Default Bearer für den Internetzugang konfiguriert sind. In diesem Beispiel wird DRX nach 4 Millisekunden aktiviert, verglichen mit einigen hundert Millisekunden, die üblicherweise für einen Default Bearer für den Internetzugang konfiguriert sind. Die Zeit, die das Radiomodul dann deaktiviert bleiben kann, wurde auf 40 Millisekunden konfiguriert, also für ein viel längeres Intervall als für den Internet Default Bearer. Dies macht Sinn, da die Pause zwischen zwei Sprachdatenpaketen sehr lange und vorhersehbar ist. Aus diesem Grund kann DRX auch sehr schnell aktiviert werden, muss aber auch rechtzeitig wieder beendet werden, um die Verzögerung der Sprachdaten trotzdem möglichst gering zu halten.

```
rrcConnectionReconfiguration
[...]
pdcpc-Config
```

```
[...]
headerCompression: RoHC
  maxCID: 2
    profile 1: Used [Not used]
    profile 2: Used [Not used]
    profile 3: Not Used [Not used]
[...]
rlc-Config: um-Bi-Directional
  um-Bi-Directional
    ul-Unacknowledged-RLC [ul-Acknowledged-RLC]
    dl-Unacknowledged-RLC [dl-Acknowledged-RLC]
[...]
drx-Config
  onDurationTimer: psf6 [psf4]
  drx-InactivityTimer: psf4 [psf200]
  drx-RetransmissionTimer: psf4 [psf16]
  longDRX-CycleStartOffset: sf40 [sf80]
```

Für die VoLTE Applikation am oberen Ende des Protokollstacks sind diese Vorgänge transparent. Alle Optimierungen werden vom IMS Dienst im Netzwerk über Schnittstellen zwischen den IMS Komponenten und dem LTE Kern- und Zugangsnetzwerk aktiviert und gesteuert. Theoretisch könnten auch Internet-basierte Sprachdienste diese Schnittstellen nutzen, falls diese vom Netzbetreiber offengelegt werden. Dies findet aber in der Praxis nicht statt.

3.2.7 Sprachcodecs und Aushandlung der Bandbreite

In der Vergangenheit wurde in leitungsvermittelnden Netzwerken nur ein standardisierter Sprachcodec verwendet und es war somit keine Aushandlung des Sprachcodecs notwendig. In SIP Netzwerken unterstützen Endgeräte jedoch üblicherweise mehrere Codecs mit sehr unterschiedlicher Sprachqualität und Bandbreitenbedarf (Codec Rates). Aus diesem Grund müssen Endgeräte beim Verbindungsauflauf sich gegenseitig über die unterstützten Codecs informieren und einen Codec auswählen, den beide Seiten unterstützen. In VoLTE Netzwerken können Codecs zusätzlichen ratenadaptiv sein und die Bandbreite für die Sprachdaten kann vom Netzwerk auf ein Maß begrenzt werden, mit dem manche Coding Raten nicht unterstützt werden.

Codecs – Narrowband – Wideband – Super Wideband

In VoLTE werden aktuell hauptsächlich zwei Sprachcodecs verwendet, der Adaptive Multi-Rate Narrowband (AMR-NB) und der AMR Wideband (AMR-WB) Codec. Aktuelle Highend Geräte und manche VoLTE Netzwerke unterstützen auch Super-Wideband Sprachdatenübertragung mit der Enhanced Voice Services (EVS)

Codecfamilie, mit der nochmals eine deutliche bessere Sprach- und Musikqualität bei gleicher Datenrate möglich ist. Zwischen zwei VoLTE Geräten wird in der Praxis mindestens AMR-WB verwendet. Viele 2G und 3G Netzwerke wurden mittlerweile auch für die Nutzung des AMR-WB Codec ausgestattet und es ist möglich, den gegenüber AMR-NB sehr viel besseren AMR-WB Codec für Gespräche zwischen VoLTE Endgeräten Geräten und Netzwerken zu verwenden, die noch kein VoLTE unterstützen oder zu Geräten, die sich aktuell nicht in einem Gebiet mit LTE Abdeckung befinden.

Manche Netzbetreiber unterstützen auch einen Wideband Sprachcodec in deren IP-basierten Festnetzen. Auch für Verbindungen zu solchen Teilnehmern wird daher AMR-WB bevorzugt. Für solche Verbindungen wird jedoch ein Media Gateway an der Grenze zwischen den beiden Netzwerken benötigt, da der Festnetz Wideband Codec (G.722) zu breitbandig (64 kbit/s) und somit nicht zum AMR-Wideband Codec (G.722.2) kompatibel ist, der in Mobilfunknetzwerken verwendet wird und nur eine Datenrate von 12,65–23,85 kbit/s benötigt.

Wenn AMR-WB nicht von beiden Endgeräten unterstützt wird, wird AMR-NB als kleinster gemeinsamer Nenner mit einer Datenrate von 12,2 kbit/s als Codec verwendet.

Adaptive Codecs

Sprachcodecs, die heute im Mobilfunk verwendet werden, sind ratenadaptiv. AMR-WB kann einen Sprachdatenstrom mit einer Datenrate von 6,6 bis 23,85 kbit/s übertragen. Am unteren Ende der Spanne ist die Sprachqualität eingeschränkt, während eine Sampling Rate von 16.000 Hz bei 23,85 kbit/s eine sehr natürlich klingende Wiedergabe der Sprache ermöglicht. In der Praxis wird AMR-WB mit einer Datenrate von 12,65 kbit/s in 2G und 3G Netzwerken verwendet und nutzt somit eine sehr ähnliche Datenrate wie die älteren Narrowband Codecs und passen somit in bereits existierende leitungsvermittelnde Kanäle. In VoLTE Netzwerken wird AMR-WB entweder mit einer Datenrate von 12,65 oder 23,85 kbit/s verwendet.

Die Codecrate kann während eines Gesprächs alle 20 ms geändert werden, also in jedem Sprachframe der in ein IP Paket verpackt wurde. Im Header wird dem Empfänger mitgeteilt, welche Datenrate verwendet wird. Abb. 3.9 zeigt, wie die Datenrateninformation in einem IP/UDP/RTP (Realtime Transport Protocol) Sprachpaket übertragen wird.

Codec Auswahl

Während eines Gesprächs haben die Endgeräte jederzeit die Möglichkeit, die Datenrate des Codecs zu ändern. Der Codec selber wird jedoch üblicherweise nur beim Verbindungsaufbau festgelegt und dann während des Gespräches nicht mehr verändert. Eine Ausnahme wird in der Praxis nur gemacht, wenn einer der beiden Teilnehmer während des Gesprächs von VoLTE in einen leitungsvermittelnden GSM oder UMTS Kanal wegen schlechten Übertragungsbedingungen wechseln muss.

```

▼Real-Time Transport Protocol
►[Stream setup by SDP (frame 8)]
 10... .... = Version: RFC 1889 Version (2)
 ..0. .... = Padding: False
 ...0 .... = Extension: False
 .... 0000 = Contributing source identifiers count: 0
 0... .... = Marker: False
 Payload type: AMR-WB (96)
 Sequence number: 54
 [Extended sequence number: 65590]
 Timestamp: 164800
 Synchronization Source identifier: 0x41815a05 (1098996229)

▼Adaptive Multi-Rate
 Payload decoded as RFC 3267 bandwidth-efficient mode
 0010 .... = CMR: AMR-WB 12.65 kbit/s (2)
 .... 0... = F bit: Last frame in this payload
 .... .001 0... .... = FT bits: AMR-WB 12.65 kbit/s (2) / Frame OK
 .1... .... = Q bit: Ok
 Frame Data (32 Bytes)

```

Abb. 3.9 AMR-WB Codec in einem RTP Paket

Am Anfang eines Gesprächs wird der Codec wie folgt ausgewählt: Im ersten Schritt teilt der Anrufer dem Netzwerk und der Gegenstelle in der SIP ‚Invite‘ Nachricht mit, welche Sprachcodecs er unterstützt. Im Session Description Protocol (SDP) Teil der Nachricht führt das Endgerät dazu alle unterstützten Codecs samt Parameter auf. Details werden in RFC 4566 beschrieben¹². Nachfolgend ein gekürztes Beispiel:

```

m=audio 42888 RTP/AVP 116 118
a=rtpmap:116 AMR-WB/16000/1
a=fmtp:116 mode-change-capability=2;max-red=0
a=rtpmap:118 AMR/8000/1
a=fmtp:118 mode-change-capability=2;max-red=0

```

In der ersten Zeile (media, m=) teilt das Endgerät der Gegenseite mit, dass es zwei Mediencodecs unterstützt und teilt ihnen die IDs 116 und 118 zu. Diese zwei Codecs werden dann in den darauffolgenden Zeilen genauer beschrieben. Ein weiterer wichtiger Parameter in der ersten Zeile ist die lokale UDP Portnummer (42888), zu dem der später eingehende Sprachdatenstrom gesendet werden soll. Die Attributzeilen (a=) beschreiben dann die Codecs hinter den IDs 116 und 118, in diesem Beispiel aus der Praxis AMR-WB und AMR-NB. Die Gegenstelle sucht sich dann einen dieser beiden Codecs aus und informiert das lokale Endgerät in einer SIP ‚183 Session Progress‘ Nachricht. Auch das Netzwerk prüft die Codec Listen und löscht alle Codec Einträge in der SIP ‚Invite‘ Nachricht, die nicht verwendet werden sollen. Dies kann z. B. der Fall sein, wenn Codecs enthalten sind, die nicht für den Mobilfunk geeignet sind, z. B. wegen zu hoher Bandbreitenanforderungen.

Aushandlung der Bandbreite

Heute verwendet VoLTE die adaptiven AMR-NB, AMR-WB und EVS Super-Wideband Codecs und kann somit Sprache in vielen unterschiedlichen Datenraten und sehr unterschiedlicher Sprachqualität übertragen. Im Falle von AMR-WB werden Sprachdaten mit einer Übertragungsrate von 6,6 bis 23,65 kbit/s gesendet. Manche Netze limitieren in der Praxis die Datenrate von AMR-WB auf 12,65 kbit/s und 12,2 kbit/s bei AMR-NB, da sich die Sprachqualität bei Verdoppelung der Übertragungsrate nicht mehr wesentlich verbessert. Netzbetreiber können, falls dies notwendig ist, auf diese Weise den Bandbreitenbedarf für Sprachtelefonie im Netzwerk begrenzen, was bei sehr vielen gleichzeitigen Verbindung durchaus vorteilhaft sein kann.

Beim Gesprächsaufbau informiert das Endgerät das Netzwerk und die Gegenstelle über den benötigten Bandbreitenbedarf im SDP Teil der SIP ‚Invite‘ Nachricht im ‚Bandwidth Information‘ Parameter. Das nachfolgende Beispiel zeigt, wie in diesem Parameter die ‚Application Specific (AS)‘ maximale Datenrate von 49 kbit/s signalisiert wird.

b=AS:49

Die signalisierte Datenrate enthält auch den Overhead für die IP, UDP und RTP Header. Eine Datenrate von 49 kbit/s wird z. B. für einen AMR-WB Datenstrom mit bis zu 23,85 kbit/s benötigt, falls IPv6 auf den IP Layer verwendet wird. In IMS Netzwerken die noch auf IPv4 basieren, werden 41 kbit/s benötigt. Der Unterschied zwischen den benötigten Bandbreiten mag zunächst groß erschienen, ist in der Praxis jedoch deutlich geringer, da durch Robust Header Compression (RoHC) in beiden Fällen in etwa die gleiche Bandbreite benötigt wird. Weitere Details zum Bandbreitenbedarf der unterschiedlichen Sprachcodecs finden sich in 3GPP TS 26.114¹³. In Tab. K.6 sind dort Beispiele für den Bandbreitenbedarf des AMR-WB Codecs in Kombination mit IPv6 im ‚Bandwidth Efficient (not octet aligned)‘ Modus mit einer Frame Länge von 20 ms pro IP Paket zu finden.

Während des Gesprächsaufbaus teilt dann das Netzwerk dem Endgerät einen LTE Dedicated Bearer zu, der dann diese angeforderte oder auch eine niedrigere Datenrate unterstützt. Falls das Netzwerk die Datenrate z. B. auf 12,65 kbit/s begrenzen möchte, limitiert es den Datendurchsatz auf dem Dedicated Bearer auf 40 kbit/s und teilt dies auch dem Endgerät in der LTE ‚Establish Dedicated Bearer‘ NAS Nachricht mit, aus der nachfolgend ein Ausschnitt gezeigt wird:

```
Quality of Service Class Identifier: QCI 1
Maximum bit rate for uplink: 40 kbps
Maximum bit rate for downlink: 40 kbps
Guaranteed bit rate for uplink: 40 kbps
Guaranteed bit rate for downlink: 40 kbps
```

Nachdem das Endgerät diese Nachricht empfangen hat, muss es eine SIP ‚Update‘ Nachricht mit einer ‚b=AS:38‘ Zeile im SDP Teil senden, und informiert so das andere Endgerät, dass das Netzwerk die Bandbreite auf seiner Seite begrenzt. Die Gegenstelle führt die gleichen Aktionen durch und schlussendlich wird für den Sprachpfad die kleinste zugewiesene Bandbreite verwendet.

3.2.8 Freiton, Ring-Back Melodien und Early-Media

Das Ziel von VoLTE ist, ein kompletter Ersatz für die leitungsvermittelnde Mobilfunktelefonie zu sein. Eine Frage die sich dabei stellt ist, wie ein Media Stream gesendet werden kann, bevor der angerufene Teilnehmer das Gespräch annimmt. Dies ist notwendig, da der rufende Teilnehmer ein Freiton oder eine Ring-Back Musik oder Ansage hören soll, während das Telefon des angerufenen Teilnehmers klingelt. Dies wird in VoLTE als ‚Early-Media‘ bezeichnet und ist in der Praxis wie folgt implementiert:

Die erste Option ist, dass das rufende Gerät einen gespeicherten Freiton abspielt, sobald es von der Gegenstelle eine SIP ‚180 Ringing‘ Nachricht erhält. In der Praxis ist jedoch der Freiton von Land zu Land verschieden und das Endgerät muss somit länderräumig unterschiedliche Sounddateien abspielen.

Die zweite Option ist, dass das Netzwerk den Freiton dem Anrufer zuspielt, bis vom gerufenen Teilnehmer eine SIP ‚200 OK‘ Nachricht empfangen wird. Dies wird als ‚Early-Media‘ Streaming bezeichnet. Auf diese Weise ist es möglich, nicht nur das normale Freizeichen abzuspielen, sondern auch eine Musikdatei, die der Nutzer statt des Freitons hinterlegt hat.

In der Praxis werden beide Varianten verwendet. Damit Early-Media verwendet werden kann, muss das anrufende Endgerät beim Gesprächsaufbau dem Netzwerk mit einer SIP ‚P-Early-Media: supported‘ Header Zeile mitteilen, dass es diese Funktion unterstützt. Will das Netzwerk dann einen Freiton oder einen Musiktitel abspielen und so die Zeit überbrücken, bis der gerufene Teilnehmer das Gespräch annimmt, teilt es dies dem anrufenden Endgerät in der SIP ‚180 Ringing‘ Nachricht über eine Header Zeile mit dem Inhalt ‚P-Early-Media: sendonly‘ mit. Danach beginnt das Netzwerk, den Freiton oder den Musiktitel auf dem virtuellen Sprachkanal zu senden. Falls Early-Media nicht verwendet wird, enthält die SIP ‚180 Ringing‘ Nachricht eine Headerzeile mit dem Inhalt ‚P-Early-Media: inactive‘. Weitere Details zu diesem Thema sind in RFC 3960¹⁴, GSMA IR.92 Abschn. 2.2.8 und in 3GPP TS 24.628¹⁵ spezifiziert.

3.2.9 Verwendung von Ports

Anwendungen, die mit einem Server im Netzwerk kommunizieren, verwenden meist nur eine einzelne TCP oder UDP Verbindung von einem Port mit zufälliger Nummer zu einem ‚Well-Known‘ Port (z. B. 443 für HTTPS) auf der Serverseite. Über diese

einzelne Verbindung führen sie dann eine Authentifizierung durch und aktivieren die Verschlüsselung. Dies ist beim VoLTE System wesentlich aufwendiger gelöst, hier werden drei Verbindungen verwendet und die Verwendung von TCP und UDP kann sogar im Wechsel erfolgen.

Ein Endgerät, das sich mit dem IMS Netzwerk über die P-CSCF verbinden will, sendet zunächst eine nicht-verschlüsselte SIP „Register“ Nachricht von einem zufälligen Port zum „Well-Known“ SIP Port 5060. Die IMS reagiert darauf mit einer SIP „401 Unauthorized“ Nachricht von Port 5060, um den Authentifizierungsprozess zu starten. Die Nachricht enthält ein „Authentication Challenge“, sowie eine UDP/TCP Portnummer, an welche die nachfolgend verschlüsselte Nachricht gesendet werden soll. Dieser Port wird als „port-s“ (server) bezeichnet. Die Nachricht enthält auch einen „port-c“ (client) Parameter, der später verwendet wird, falls das IMS System das Endgerät ohne vorherige Anfrage kontaktieren will. Dies sind die einzigen zwei Nachrichten, die über Port 5060 ausgetauscht werden.

Danach sendet das Endgerät eine neue SIP „Register“ Nachricht, dieses Mal verschlüsselt und als Antwort auf die Security Challenge an den TCP Port, der im „port-s“ Parameter vom Netzwerk vorgegeben wurde. Die Register Nachricht des Endgeräts wiederholt die „port-c“ und „port-s“ Parameter, die es vom IMS System erhalten hat und gibt zusätzlich seine eigenen „port-c“ und „port-s“ Nummer bekannt. Die zweite Portkombination, also IMS „port-c“ und die „port-s“ Adresse des Endgerätes werden später verwendet, wenn das IMS System das Endgerät unaufgefordert kontaktieren will. War die zweite SIP „Register“ Nachricht korrekt, antwortet das IMS Netzwerk mit einer SIP „200 OK“ Nachricht und das Endgerät ist im VoLTE System registriert.

Wie zuvor schon erwähnt, verwendet das Endgerät immer seine „port-c“ (client) UDP/TCP Portnummer und die „port-s“ (server) Nummer, wenn es dem IMS System eine Nachricht schicken will. Die Antwort auf eine solche Nachricht wird auf der gleichen Verbindung gesendet. Wenn das Netzwerk dagegen eine Nachricht senden möchte, die nicht mit einer zuvor vom Endgerät erhaltenen Nachricht zusammenhängt, verwendet der Server seinen „port-c“ und sendet die Nachricht an den „port-s“ des Endgeräts. Außerdem ist es auch möglich, TCP und UDP zu mischen. Für kleine SIP Nachrichten werden UDP Pakete verwendet und es wird auf TCP gewechselt, wenn eine SIP Nachricht zu groß für ein einzelnes UDP Paket wird. Details hierzu sind in 3GPP TS 33.203, Kap. 7 zum Thema „Access Security“ zu finden¹⁶.

3.2.10 Filterung von Nachrichten und Asserted Identities

Abb. 3.7 und 3.8 lassen zunächst den Eindruck entstehen, dass SIP Nachrichten mehr oder weniger transparent zwischen den zwei Endgeräten ausgetauscht werden. In VoLTE Systemen ist das jedoch nicht der Fall. In der Praxis ist es die Aufgabe diverser Netzwerkelemente, Nachrichten zu modifizieren, bevor sie dann an den nächsten Netzwerknoten und schließlich bis zum anderen Endgerät weitergeleitet werden. Ein gutes

Beispiel ist die SIP ‚Invite‘ Nachricht, mit der ein Gespräch aufgesetzt wird. Zunächst wird die SIP Nachricht vom Endgerät über den IPSec Tunnel zur P-CSCF gesendet. Dort endet zunächst der IPSec Tunnel. Die P-CSCF fügt dann in der ‚Invite‘ Nachricht eine SIP ‚P-Asserted-Identity‘ Headerzeile hinzu. Dies ist nötig, da das Endgerät theoretisch jede beliebige Identity als seine Telefonnummer angeben kann. Um Missbrauch vorzubeugen, fügt das Netzwerk diese Header Zeile ein, die dann dem angerufenen Teilnehmer signalisiert, dass die Identität (Telefonnummer) vom Netzwerk eingefügt wurde und dieser Information somit vertraut werden kann.

In der S-CSCF und der TAS wird die ‚Invite‘ Nachricht dann komplett neu erzeugt, bevor sie weiter zum Zielteilnehmer gesendet wird. Informationen, die z. B. entfernt werden, sind Herstellerinformationen, Modellname des Endgerätes, Software Versionsnummern und die International Mobile Equipment Identity (IMEI) des Endgeräts. Diese Informationen sind im ‚User-Agent‘ Header, sowie im ‚Contact‘ Header und in den SDP ‚originator, o=‘ Parametern enthalten.

Üblicherweise werden in VoLTE Netzwerken die Sprachpakete nicht direkt zwischen zwei Endgeräten ausgetauscht, sondern werden über Media Gateways geleitet. Aus diesem Grund werden deshalb auch die IP-Adressen und Portnummern für den Mediastream vom Netzwerk geändert. Zusätzlich werden auch nur Sprachcodec- und Bandbreiteninformationen an den Zielteilnehmer weitergeleitet, die im Netzwerk auch unterstützt werden. Auch alle nicht bekannten Informationen werden, um Missbrauch zu verhindern, aus den SIP Nachrichten entfernt. Zusammenfassend lässt sich sagen, dass die SIP ‚Invite‘ Nachricht, sowie auch zahlreiche anderen Nachrichten, stark modifiziert werden und deutliche Unterschiede zwischen Sender und Empfänger aufweisen.

3.2.11 DTMF Töne

Eine weitere Funktion des Telefoniedienstes, die auch heute noch gebraucht wird, sind Dual-Tone Multi-Frequency (DTMF) Töne, da mit diesen über einen Sprachkanal Geräte wie Anrufbeantworter und Konferenzbrücken gesteuert werden. Im analogen Festnetz wurden DTMF Töne vom Endgerät erzeugt und als hörbarer Ton über den Sprachkanal übertragen. In GSM und UMTS Netzwerken werden DTMF Töne als Signalisierungsnachrichten an das Mobile Switching Center übertragen. Dort werden diese Nachrichten ausgewertet und das Media Gateway angewiesen, die hörbaren Töne in den Sprachkanal einzuspielen. In VOLTE Netzwerken werden DTMF Töne durch eine Mischung von In-Band Signalisierung und digitalen Nachrichten erzeugt. Statt einer SIP Nachricht zwischen den zwei Geräten für einen DTMF Ton zu senden, wird die DTMF Information, wie in Abb. 3.10 gezeigt, in den RTP (Realtime Transport Protocol) Headern von RTP Sprachpaketen gesendet. Da ein RTP Paket üblicherweise 20 ms Sprachdaten enthält, muss alle 20 ms eine DTMF Signalisierungsnachricht erzeugt werden und im RTP Header eingebettet werden. GSMA IR.92 referenziert für

```

► User Datagram Protocol, Src Port: 1424 (1424), Dst Port: 32566 (32566)
▼ Real-Time Transport Protocol
  ► [Stream setup by SDP (frame 17)]
    10... .... = Version: RFC 1889 Version (2)
    ..0.... = Padding: False
    ..0.... = Extension: False
    .... 0000 = Contributing source identifiers count: 0
    0... .... = Marker: False
    Payload type: telephone-event (110)
    Sequence number: 60738
    [Extended sequence number: 60738]
    Timestamp: 952360462
    Synchronization Source identifier: 0x003e070a (4065034)

▼ RFC 2833 RTP Event
  Event ID: DTMF Seven 7 (7)
  0... .... = End of Event: False
  ..0.... = Reserved: False
  ..00 1010 = Volume:10
  Event Duration: 960

```

Abb. 3.10 Ein RTP Frame mit einer eingebetteten DTMF Signalisierungsnachricht. Diese Nachricht wird durch Drücken der Taste „7“ auf dem Endgerät erzeugt

die Implementierung 3GPP TS 26.114 Annex G, wo wiederum auf RFC 4733 für Details verwiesen wird¹⁷.

Wenn RTP DTMF Signalisierungsnachrichten auf der Gegenseite ankommen, ist es die Aufgabe des Endgeräts, die dazu gehörenden hörbaren Töne zu erzeugen. Falls die Gegenseite kein VoLTE Endgerät ist, müssen die Töne vom Media Gateway im Netzwerk umgewandelt und in den Sprachdatenstrom eingefügt werden.

Aus Implementationssicht macht es Sinn, DTMF Töne als Nachrichten im Sprachdatenstrom zu senden. Da der hörbare Ton erst möglichst nahe am Empfänger erzeugt wird, sind Qualitätseinbußen, die durch eventuelle Transcoder im Netz verursacht werden können, so gering wie möglich. Außerdem stellt die Übertragung im Sprachdatenstrom sicher, dass die DTMF Töne ohne Verzögerungen übertragen werden.

3.2.12 SMS über IMS

Nicht-VoLTE fähige LTE Endgeräte senden und empfangen traditionelle SMS Nachrichten über den LTE Signalisierungskanal und das SGs Interface wie in Kap. 1 beschrieben. VoLTE Endgeräte können dagegen SMS Nachrichten über SIP Nachrichten senden und empfangen. Abb. 3.11 zeigt, wie eine SMS in einer SIP „Message“ Nachricht gesendet wird. Die SMS Nachricht wird dabei in der SIP Nachricht gekapselt und ist für

- ▶ User Datagram Protocol, Src Port: 5122 (5122), Dst Port: 5063 (5063)
- ▼ Session Initiation Protocol (MESSAGE)
 - ▶ Request-Line: MESSAGE sip:+447700330000@ims.btxa.co.uk;user=phone SIP/2.0
 - ▶ Message Header
 - ▼ Message Body
 - ▼ GSM A/I/F RP – RP-DATA (MS to Network)
 - Message Type RP-DATA (MS to Network)
 - ▶ RP-Message Reference
 - ▶ RP-Originator Address
 - ▶ RP-Destination Address - (447700330000)
 - ▶ RP-User Data
 - ▼ GSM SMS TPDU (GSM 03.40) SMS-SUBMIT
 - 0... = TP-RP: TP Reply Path parameter is not set in this SMS SUBMIT/DELIVER
 - .0. = TP-UDHI: The TP UD field contains only the short message
 - .0. = TP-SRR: A status report is not requested
 - ...1 0... = TP-VPF: TP-VP field present - relative format (2)
 -0.. = TP-RD: Instruct SC to accept duplicates
 -01 = TP-MTI: SMS-SUBMIT (1)
 - TP-MR: 1
 - ▶ TP-Destination-Address - (4477127233140)
 - ▶ TP-PID: 0
 - ▶ TP-DCS: 0
 - TP-Validity-Period: 63 week(s)
 - TP-User-Data-Length: (11) depends on Data-Coding-Scheme
 - ▼ TP-User-Data
 - SMS text: All is well

Abb. 3.11 Eine SMS Nachricht, die per SIP übertragen wird

das IMS System transparent. Das bedeutet, dass die SMS den exakt gleichen Aufbau wie bei einer Übermittlung über GSM, UMTS oder über LTE und das SGs Interface hat.

Auf der Netzwerkseite bildet das IP-Short-Message-Gateway (IP-SM-GW) die Brücke zwischen dem SMS Service Center im alten leitungsvermittelnden Signalisierungsnetzwerk und der IMS IP Welt. Details hierzu finden sich in 3GPP TS 23.204¹⁸. Da das Format der SMS Nachricht für die Übertragung per SIP nicht geändert wird, können alle Anwendungen, die auf SMS Kommunikation aufbauen, ohne Änderung auf höheren Schichten weiterverwendet werden. Ein Beispiel für eine solche Anwendung ist die Übertragung von Binärdaten auf die SIM Karte, um dort Informationen zu ändern (vgl. Kap. 5).

3.2.13 Konfiguration der Anrufweiterleitung und XCAP

Der traditionelle Telefoniedienst bietet eine Anzahl von Zusatzfunktionen (Supplementary Services) wie z. B. folgende Arten der Gesprächsweiterleitung:

- Sofortige Anrufweiterleitung (Call Forward Immediately)
- Anrufweiterleitung bei besetzt (Call Forward Busy)
- Anrufweiterleitung wenn der Nutzer das Gespräch nicht annimmt (Call Forward No Reply)

- Anrufweiterleitung wenn das Gerät ausgeschaltet ist oder nicht erreicht werden kann (Call Forward Not Reachable)

Im IMS System sind solche Supplementary Services im Telephony Application Server (TAS) implementiert und können vom Nutzer über sein Endgerät konfiguriert werden. Das bedeutet, dass der Nutzer Telefonnummern konfigurieren kann, zu denen ein eingehendes Gespräch für die unterschiedlichen Fälle weitergeleitet werden soll. Außerdem kann er über das Endgerät die unterschiedlichen Weiterleitungsarten aktivieren, deaktivieren und im Falle von ‚Call Forward No Reply‘ die Zeit festlegen, nach der ein Anruf weitergeleitet werden soll. Für diese Konfigurationsaufgaben verwendet VoLTE das XML Configuration Access Protocol (XCAP), dessen Aufgaben in GSMA IR.92 Abschn. 2.3 beschrieben sind.

XCAP Nachrichten verwenden das HTTP Protokoll über einen LTE Default Bearer oder über einen GSM/UMTS PDP Kontext. In der Praxis verwenden Netzbetreiber dazu nicht den IMS Default Bearer, der für die SIP Signalisierung verwendet wird, sondern entweder den Internet Default Bearer oder einen dritten Default Bearer mit eigener IP Adresse. Der Grund dafür ist, dass manche Netzbetreiber den IMS Default Bearer nur im LTE Netzwerk unterstützen, nicht jedoch in deren GSM und UMTS Radionetzwerken. Außerdem kann oftmals auch im Roaming Fall nicht auf den IMS Bearer zugegriffen werden.

GSMA IR.92 definiert mehrere Mechanismen für die Authentifizierung des Datenaustauschs über HTTP, wie z. B. ein Challenge/Response Mechanismus mit Authentifizierung über den geheimen Schlüssel Ki auf der SIM-Karte. Über HTTP wird der Challenge/Response Mechanismus wie folgt ausgeführt: Zunächst sendet das Endgerät einen XCAP Request ohne Authentifizierungsinformationen. Das Netzwerk sendet daraufhin eine ‚HTTP 401 Unauthorized‘ Nachricht, in der auch ‚Challenge‘ Informationen enthalten sind. Das Endgerät gibt die Challenge Informationen dann an die SIM-Karte weiter, in der dann die dazugehörige Antwort berechnet wird. Mit dieser Antwort sendet dann das Endgerät die XCAP Anfrage erneut.

Noch bevor eine XCAP Anfrage zur TAS Supplementary Service Datenbank über HTTP geschickt werden kann, muss das Endgerät über eine DNS Anfrage die IP Adresse des Servers ermitteln. Der Name des Servers ist wie folgt strukturiert und enthält den Mobile Country Code (MCC) und Mobile Network Code (MNC) des Heimnetzwerkes:

xcap.ims.mncXXX.mccXXX.pub.3gppnetwork.org

Nachdem dem Endgerät die IP Adresse des Servers bekannt ist, sendet es eine HTTP ‚GET‘ Anfrage, um zunächst einmal die aktuelle Gesprächsumleitungskonfiguration zu erhalten. Der Inhalt der Anfrage ist standardisiert und enthält, wie nachfolgend gezeigt, die Telefonnummer des Teilnehmers und den IMS Identifier des Heimnetzwerkes:

```
GET/simservs.ngn.etsi.org/users/sip:+443393144238@ims.vodafone.co.uk/simservs.xml/~~/simservs/communication-diversion HTTP/1.1
```

Der TAS antwortet auf die Anfrage dann mit einer XML formatierten Liste der Parameter und deren Werte für alle Gesprächsweiterleitungsformen. Der folgende Ausschnitt der XML Antwort zeigt die aktuelle Konfiguration der „Call Forward No Reply“ Weiterleitung:

```
<communication-diversion active="true">
  <NoReplyTimer>25</NoReplyTimer>
  [...]
  <cp:ruleset>
    [...]
    <cp:rule id="cfnry">
      <cp:conditions>
        <rule-deactivated/>
        <no-answer/>
      </cp:conditions>
      <cp:actions>
        <forward-to>
          <target>tel:+493397788990</target>
        </forward-to>
      </cp:actions>
    </cp:rule>
    [...]
  </cp:ruleset>
</communication-diversion>
```

In diesem Beispiel ist der No-Reply Timer auf 25s gesetzt und die Telefonnummer, zu der ein ankommender Anruf weitergeleitet soll, ist +493397788990. Außerdem kann man sehen, dass diese Regel aktuell deaktiviert ist (rule-deactivated).

Um Einstellungen auf dem TAS Server zu ändern, sendet das Endgerät eine HTTP „PUT“ Anfrage, die wie folgt aufgebaut ist:

```
PUT/simservs.ngn.etsi.org/users/sip:+443393144238@ims.telekom.de/simservs.xml/~~/simservs/communication-diversion?xmlns(cp=urn:ietf:params:xml:ns:common-policy) HTTP/1.1
```

Im „Body“ Teil der HTTP „PUT“ Anfrage befindet sich die gleiche XML Struktur wie oben für die „GET“ Anfrage gezeigt. Um diese Weiterleitung zu aktivieren wird einfach der „<rule-deactivated>“ Parameter weggelassen.

3.2.14 Single Radio Voice Call Continuity

Da zu erwarten ist, dass auch in den nächsten Jahren GSM noch eine bessere Flächenversorgung als LTE haben wird, ist es nötig, ein bestehendes Gespräch am Rand der LTE Abdeckung nach GSM oder UMTS zu übergeben. Von UMTS zu GSM war dies noch recht einfach möglich, da hier eine leitungsvermittelte Verbindung in eine andere leitungsvermittelte Verbindung übergeben wird. Da ein VoLTE Gespräch jedoch auf IMS und IP basiert, muss bei einem solchen Handover ein Voice over IP Gespräch in eine leitungsvermittelte Verbindung übergeben werden. Da ein Endgerät nicht gleichzeitig in LTE und UMTS/GSM aktiv sein kann (Single Radio), muss eine VoLTE Handover Lösung auch dies berücksichtigen. In 3GPP wurde dazu ein Verfahren namens Single Radio Voice Call Continuity (SRVCC) in 3GPP 23.237¹⁹ spezifiziert und über die Jahre weiterentwickelt. Die nachfolgende Beschreibung basiert auf der SRVCC Lösung, wie sie in 3GPP Release 10 beschrieben ist. Netzbetreiber können aber in der Praxis auch andere SRVCC Versionen einsetzen, die im Prinzip jedoch sehr ähnlich funktionieren.

Abb. 3.12 zeigt, welche IMS Komponenten bei einem VoLTE Handover nach GSM oder UMTS involviert sind. Wie auch zuvor sind die P-, I- und S-CSCF Server für den Aufbau und Erhalt der VoLTE Verbindung zuständig. Zusätzlich zum MMTEL Application Server gibt es in SRVCC-fähigen IMS Netzwerken einen Service Centralization and Continuity Application Server (SCC-AS), der schon während der IMS Registrierung und während des Gesprächsaufbaus an der SIP Signalisierung beteiligt wird und somit im Handoverfall schon alle benötigten Informationen hat, die Verbindung schnell in das leitungsvermittelnde Netzwerk zu übergeben. Dieses wird in Abb. 3.12 durch den MSC-Server (MSC-S) und dem Media Gateway (MGW) repräsentiert, die in später im Kap. 5 noch näher beschrieben werden.

Um die Verbindung möglichst schnell übergeben zu können, werden die IP Sprachdatenpakete nicht mehr direkt zwischen den Endgeräten ausgetauscht, sondern für beide Teilnehmer des Gesprächs jeweils über ein Access Transfer Gateway (ATGW) geführt. Dieses wird von der Access Transfer Control Function (ATCF) gesteuert, die Teil des P-CSCF ist.

Abb. 3.13 zeigt, wie eine Sprachverbindung während eines SRVCC Handovers im Netzwerk umgeleitet wird. Der Gesprächsaufbau verläuft zunächst wie zuvor beschrieben, SRVCC Release 10 Netzwerke schalten jedoch an beiden Enden der Verbindung das Access Transfer Gateway in den Sprachpfad. Da das Endgerät während des Attach Prozesses dem Netzwerk mitgeteilt hat, dass es SRVCC unterstützt, kennt der eNodeB die Fähigkeit des Endgeräts, eine Sprachverbindung in ein leitungsvermittelndes Netzwerk mitnehmen zu können und weist die MME an, die Verbindung in ein 2G oder 3G Netzwerk zu überführen. Diese sendet daraufhin einen „PS to CS Transfer Request“ an die MSC, die für die GSM oder UMTS Zielzelle zuständig ist. Die MSC reserviert daraufhin in der Zielzelle den benötigten leitungsvermittelten Kanal und weist außerdem die Access Transfer Control Function (ATCF) an, das ATGW so zu konfigurieren, dass

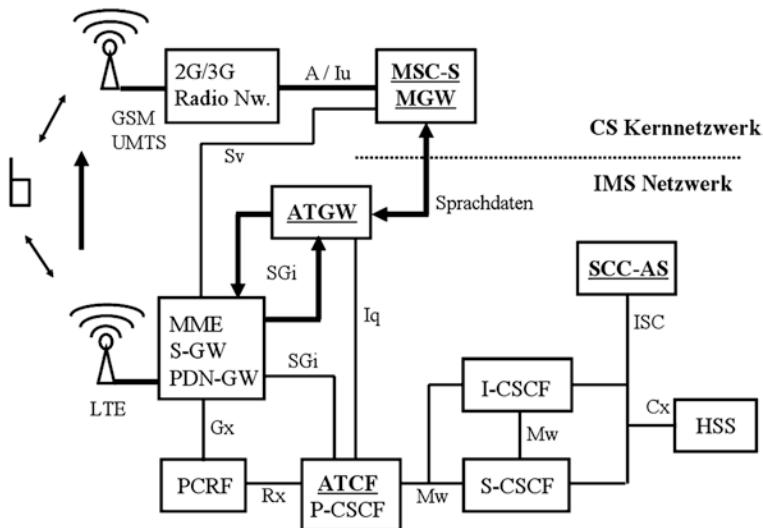


Abb. 3.12 Für SRVCC benötigte IMS und MSC Komponenten

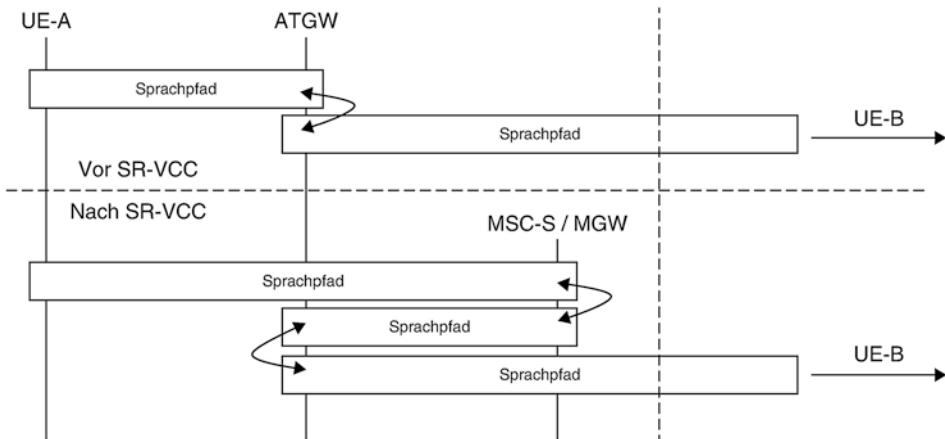


Abb. 3.13 Sprachverbindung vor und nach einem SRVCC Handover

in Kürze die Gesprächsdaten nicht mehr von und zur IP Adresse des Endgerätes übertragen werden sollen, sondern von und zur IP Adresse des MSC Media Gateways. Dies geschieht mit einer SIP „Invite“ Nachricht.

Nachdem die ATCF diese Nachricht mit einer SIP „200 OK“ Antwort beantwortet hat und somit alles für den Handover vorbereitet ist, schickt der MSC Server eine „PS to CS Response“ Nachricht an die MME zurück. Die MME löst dann mit einem Handover Kommando an das Endgerät den Netzwerkwechsel aus.

Der nachfolgende Ausschnitt aus einer solcher Nachricht zeigt die wichtigsten Handover Parameter, die vom Netzwerk zum Endgerät gesendet werden:

```

mobilityFromEUTRACommand
CS-Fallback Indicator: False
Purpose: Handover
Target RAT Type: GERAN
Target RAT Message Container
    GSM A-I/F DTAP - - - Handover Command
Protocol Discriminator: Radio Resources Management Messages
    DTAP Radio Resources Management Message Type: Handover Command
Cell Description
    NCC: 2
    BCC: 6
    BCCH ARFCN(RF channel number): 32
    TCH/F + FACCH/F and SACCH/F
    Timeslot: 6
    Training Sequence: 6
    Hopping Channel: No
    Single Channel : ARFCN 32
    Channel Mode: FR AMR-WB (Full Rate - AMR-Wideband)
    Cipher with algorithm A5/3

```

Nachdem das Endgerät die Handover Nachricht erhalten hat, wechselt es in die GSM oder UMTS Zelle und sendet seine Sprachdaten fortan in einer leitungsvermittelten Verbindung. Diese wird dann am MSC Media Gateway wieder in eine IP Verbindung umgewandelt und zum ATGW geschickt. Das ATGW informiert den ATCF über den geglückten Handover, der wiederum den Service Centralization and Continuity Application Server (SCC-AS) über den Transfer informiert. Dieser schickt dann zum Abschluss des Handovers eine SIP ‚Bye‘ Nachricht an den P-CSCF, damit dieser den Dedicated Bearer der Sprachverbindung im LTE Netzwerk freigeben kann.

Ein weiterer wichtiger Aspekt der SRVCC Prozedur ist, wie der Handover überhaupt ausgelöst wird. Typischerweise sendet das LTE Netzwerk eine Messkonfiguration zum Endgerät beim Aufbau der Verbindung und wird so über sich verschlechternde Radiobedingungen informiert. Wenn ein konfigurierter Grenzwert erreicht ist, weist der eNodeB das Endgerät an, nach GSM und UMTS Zellen während des laufenden VoLTE Gesprächs zu suchen und gefundene Zellen an ihn zu melden. Damit das Endgerät auf anderen Frequenzbändern nach Zellen suchen kann, muss der eNodeB die LTE Verbindung für periodische Sende- und Empfangspausen umkonfigurieren. Nachdem der eNodeB vom Endgerät eine Liste von gefundenen GSM und UMTS Zellen bekommen hat, wählt dieser dann eine geeignete Zelle für den Handover in. In der Praxis kann

beobachtet werden, dass die Suche nach GSM und UMTS Zellen etwa 2 bis 3s dauert. Während dieser Zeit führt der Nutzer das Telefongespräch auf der LTE Zelle fort. Das bedeutet, dass der Handoverprozess schon deutlich vor dem kompletten Verlust der LTE Abdeckung gestartet werden muss, also einige db vor der kritischen Signalstärkegrenze.

Ein Anruf kann sich zum Zeitpunkt, an dem ein SRVCC Handover nötig wird, in unterschiedlichen Zuständen befinden. Im einfachsten Fall ist der Anruf aufgebaut und der Zielteilnehmer hat das Gespräch schon angenommen. Ein Anruf kann auch in der ‚Alerting‘ Phase sein, d. h. das Endgerät des Zielteilnehmers klingelt, dieser hat aber noch nicht abgenommen. Wenn der Zielteilnehmer das Gespräch nicht sofort annimmt, kann diese Phase durchaus viele Sekunden dauern. Die Wahrscheinlichkeit, dass ein Gespräch sich also gerade in diesem Zustand befindet ist nicht gering. Ursprünglich wurde in den 3GPP Spezifikation nur die SRVCC Prozedur für angenommene Gespräche standardisiert, nicht jedoch für die ‚Alerting‘ Phase. Wenn der Teilnehmer sich in dieser Zeit aus der LTE Abdeckung bewegt, würde das Gespräch abbrechen. Aus diesem Grund wurde etwas später in 3GPP eine Erweiterung namens Alerting-SRVCC (aSRVCC) spezifiziert. Sowohl das Netzwerk als auch das Endgerät muss diese Erweiterung unterstützen. Endgeräte signalisieren dies an das IMS Netzwerk mit einem ‚+g.3gpp.srvcc-alerting‘ Tag im SIP ‚Contact‘ Header.

Weitere Zustände, in denen ein SRVCC Handover notwendig werden kann, sind:

- Konferenzgespräche
- Anruf halten (Call Hold)
- Der Anrufer spricht und ein weiterer Anruf ging ein, wurde aber noch nicht angenommen (Call Waiting).
- Der Anrufer spricht, während ein zweites Gespräch gehalten wird.

Um auch in diesen Zuständen einen SRVCC Handover zu ermöglichen, spezifizierte 3GPP auch SRVCC für ‚Mid-Call Services‘. Wenn Endgeräte auch diesen SRVCC Typ unterstützen, senden sie ein ‚+g.3gpp.mid-call‘ Tag im SIP ‚Contact‘ Header.

Ein weiterer, jedoch weniger wahrscheinlicher Zustand eines Anrufes, bei dem ein SRVCC durchgeführt werden sollte, ist die Zeit zwischen der SIP ‚Invite‘ Nachricht und dem SIP ‚180 Ringing‘. Für diese Phase wurde in 3GPP TS 23.237 der ‚Before-Ringing SRVCC‘ (bSRVCC) spezifiziert. Und schließlich wurde auch noch die Möglichkeit spezifiziert, einen Handover eines leitungsvermittelnden GSM oder UMTS Gespräch nach LTE/VoLTE durchzuführen. Dies wird als ‚Reverse SRVCC‘ (rSRVCC) bezeichnet.

In der Praxis verwenden LTE Netzbetreiber mit GSM oder UMTS Netzwerken üblicherweise zumindest den grundsätzlichen SRVCC Mechanismus für den Handover bestehender Gespräche. Unterstützung von SRVCC in weiteren Zuständen ist dagegen weniger weit verbreitet.

3.2.15 Wahl des Radionetzwerkes, T-ADS und VoLTE Interworking mit GSM und UMTS

Zusätzlich zur SRVCC Gesprächsübergabe eines VoLTE Telefonats ist es nötig, weiterhin leitungsvermittelte Telefonie zu unterstützen, da sich ein Teilnehmer auch für längere Zeit außerhalb der LTE Abdeckung befinden kann. Außerhalb der LTE Netzabdeckung deaktiviert ein Endgerät seinen VoLTE Stack und registriert sich als normales GSM oder UMTS Endgerät im Netzwerk. Für eingehende Gespräche muss das IMS Netzwerk deshalb wissen, dass der Teilnehmer sich nicht im LTE Netzwerk befindet und Gespräche zum Mobile Switching Center weiterleiten, bei dem der Teilnehmer gerade registriert ist. Die MSC benachrichtigt dann den Teilnehmer (paging) wie später in Kap. 5 beschrieben und baut einen leitungsvermittelnden Kanal auf. Diese Funktionalität wird als ‚Terminating-Access Domain Selection‘ (T-ADS) bezeichnet und ist in 3GPP TS 23.221, Abschn. 7.2b beschrieben²⁰.

Nach einer Reselection Prozedur zu GSM oder UMTS führt das Endgerät eine Location Area Update Prozedur durch und meldet sich dabei beim leitungsvermittelnden Mobile Switching Center an. Außerdem führt das Endgerät eine Routing Area Update Prozedur durch, um den paketvermittelnden Teil des GSM oder UMTS Netz von seiner Anwesenheit zu informieren. Bei beiden Registrierungen wird die Home Subscriber Server (HSS) Datenbank informiert. Während eines Routing Area Updates transferiert das Netzwerk auch die LTE Default Bearer (die IP Adressen des Endgerätes) von der LTE Mobility Management Entity (MME) zum GSM oder UMTS Serving GPRS Support Node (SGSN). Auf diese Weise ist es möglich, bestehende IP Verbindungen auch beim Wechsel zwischen den Radio Access Netzwerken mitzunehmen.

Manche Netzwerke transferieren alle Default Bearer von LTE nach GSM oder UMTS. Dies sind üblicherweise der Default Bearer für den Internetzugang, sowie der Default Bearer für Zugriff auf das IMS Netzwerk. Das bedeutet, dass das Endgerät weiterhin SIP Nachrichten senden und empfangen kann und auch einen VoLTE Call aufbauen könnte. In der Praxis ist dies jedoch nicht möglich, da GSM Netzwerke nicht genug Kapazität haben, um Telefonate über das paketorientierte GPRS Netzwerk zu übertragen. Über UMTS wären VoLTE Gespräche theoretisch möglich, Netzbetreiber haben jedoch üblicherweise keine Anpassungen im UMTS Radionetzwerk vorgenommen, um Dedicated Bearer für die Sprachpakete zu unterstützen. Aus diesem Grund überprüft die T-ADS Funktion für jeden eingehenden Anruf im Home Subscriber Server (HSS), in welchem Radionetzwerk sich der Teilnehmer gerade befindet. Falls der Teilnehmer im UMTS oder GSM Netzwerk registriert ist, wird das Gespräch zu einem Mobile Switching Center weitergeleitet und kein SIP ‚Invite‘ zum Endgerät geschickt. Auch die Endgeräte wissen, dass VoLTE Gespräche nur im LTE Netzwerk zugelassen sind und bauen von sich aus in GSM und UMTS nur leitungsvermittelnde Gespräche auf, auch wenn sie theoretisch ein SIP ‚Invite‘ schicken könnten.

Manche Netzbetreiber transferieren nur den Internet Default Bearer, wenn das Endgerät von LTE nach GSM oder UMTS wechselt. Der IMS Default Bearer für VoLTE

wird hingegen abgebaut. In diesem Fall wird dann die P-CSCF informiert und kann dann an Stelle des Endgeräts eine SIP Deregistrierung vornehmen. Der Nachteil dieser Prozedur ist jedoch, dass das Endgerät bei der Rückkehr zu LTE den VoLTE IMS Default Bearer neu aufbauen muss und sich danach auch wieder im IMS Netzwerk registrieren muss.

Das Zusammenspiel (Interworking) zwischen der MSC und dem VoLTE/IMS System benötigt eine enge Integration der zwei Systeme. Dies ist nicht nur für Gespräche und Erreichbarkeit notwendig, sondern auch für die Konfiguration der Supplementary Services. Für Kunden, die den VoLTE Dienst nutzen können, wird die Gesprächsweiterleitungsfunktion nicht mehr von der MSC und dem Home Location Register (HLR) verwaltet, sondern vom IMS Telephony Application Server (TAS) bereitgestellt. Supplementary Service Anfragen müssen deshalb nun vom Mobile Switching Center (MSC) zum TAS weitergeleitet werden, oder die Endgeräte müssen so konfiguriert werden, immer XCAP zu verwenden, falls die MSC keine Weiterleitung der Anfragen unterstützt. Teilnehmer von Netzen, die diese Weiterleitung nicht unterstützen, haben das Problem, dass sie mit einer für VoLTE aktivierten SIM-Karte in einem noch nicht VoLTE fähigen Endgerät keine Änderungen an der Konfiguration für Gesprächsweiterleitungen vornehmen können, da diese Endgeräte die XCAP Schnittstelle nicht unterstützen.

3.2.16 VoLTE Notrufe

Da VoLTE ein Dienst der Netzbetreiber ist, sind diese auch verpflichtet, den Notrufdienst über VoLTE in gleicher Weise wie in GSM und UMTS bereitzustellen. Notrufe werden vom Netzwerk automatisch zu einem Public Safety Answering Point (PSAP), also einem Notrufzentrum, weitergeleitet, das sich in der Nähe des Teilnehmers befindet. Das nächstgelegene Notrufzentrum wird anhand der Cell-ID der Zelle ermittelt, über die der Teilnehmer den Notruf aufbaut. In GSM und UMTS Netzwerken werden Notrufe vom Endgerät über standardisierte Nummern wie 112 (z. B. Europa) und 911 (z. B. Nordamerika) erkannt und ein Rufaufbau gestartet, der sich vom normalen Rufaufbau unterscheidet. Länderspezifische Notrufnummern können vom Netzbetreiber auf der SIM-Karte hinterlegt werden oder dem Endgerät während der Registrierung im Netzwerk mitgeteilt werden. Beim Aufbau eines Notrufs in GSM oder UMTS wird dem Netzwerk die gewählte Nummer nicht mitgeteilt, sondern statt einer normalen DTAP Call Setup Nachricht eine DTAP Emergency Call Setup Nachricht gesendet. Auf diese Weise ist der Aufbau eines Notrufs unabhängig von der gewählten Nummer. Des Weiteren gibt das Netzwerk Notrufen eine höhere Priorität und beendet bei Überlast andere Gespräche in der Zelle. Und schließlich erlauben GSM und UMTS Netzwerke auch Notrufe von Endgeräten mit SIM-Karten anderer Netzbetreiber, die sich ansonsten nicht in das Netzwerk einbuchen dürfen. Dies ist z. B. in ländlichen Gegenden wichtig, die nicht von allen Netzbetreibern versorgt werden. Für VoLTE gelten die gleichen Regeln.

In der Praxis gibt es für VoLTE LTE Netzwerke zwei Möglichkeiten, die Notruffunktionalität anzubieten. Die erste Möglichkeit, die nur als Übergangslösung dienen soll, ist die Verwendung des CS-Fallback (CSFB) Mechanismus nach GSM oder UMTS wie in Kap. 1 beschrieben. Dies bedeutet, dass das VoLTE Netzwerk Notrufe nicht selber weiterleiten kann. Die zweite Möglichkeit für VoLTE Netzwerke ist, die Notruffunktionalität selber anzubieten. Ob VoLTE Notrufe in einem Netzwerk erlaubt sind oder ob der CSFB Mechanismus verwendet werden soll, erfährt ein Endgerät über einen Parameter in der System Information Broadcast (SIB) 1 Nachricht, die VoLTE Endgeräte unabhängig ihres Registrierungsstatus im Netzwerk empfangen können. Falls VoLTE Notrufe unterstützt werden, enthält SIB 1 den ‚ims-EmergencySupport‘ Parameter, der auf ‚true‘ gesetzt ist. Falls der Parameter nicht vorhanden ist und das Gerät nicht im Netzwerk registriert ist, muss es ein GSM oder UMTS Netzwerk für einen Notruf verwenden.

Zusätzlich teilt das LTE Netzwerk einem Endgerät auch bei der LTE Attach und der Tracking Area Update Prozedur mit, ob VoLTE Notrufe erlaubt sind. Wenn VoLTE Notrufe unterstützt werden, enthalten die Attach Accept/Tracking Area Update Accept Nachrichten die ‚EPS Network Feature Support (IMS VoPS: IMS voice over PS session in S1 mode)‘ und ‚VoLTE emergency support (EMC BS: emergency bearer services in S1 mode supported‘ Parameter, die auf ‚1‘ gesetzt sind. ‚S1‘ bezeichnet dabei die LTE Schnittstelle zwischen eNodeB und MME.

Optional kann das Netzwerk in diesen Nachrichten auch eine Liste von Notrufnummern übergeben, die im Land dieses Netzwerkes gültig sind, sowie die Information, welche dieser Nummern mit der Polizei, Notarzt, Feuerwehr, etc. verbunden sind. Damit ist es möglich, nationale Notrufnummern zusätzlich zu den standardisierten Notrufnummern (112, 911) und den Nummern, die auf der SIM-Karte gespeichert sind, zu deklarieren. Das Netzwerk hat dann auch die Möglichkeit, je nach Notrufart eine andere Notrufzentrale zu kontaktieren.

Wenn ein Nutzer eine Notrufnummer wählt und das Netzwerk unterstützt VoLTE Notrufe, wird nicht die existierende VoLTE SIP Verbindung über den IMS Default Bearer verwendet, sondern ein zusätzlicher Default Bearer für den Notruf aufgebaut. Für diesen zusätzlichen Default Bearer muss kein APN Namen angegeben werden. Stattdessen wird der ‚Request Type‘ in der ‚PDN Connectivity Request‘ Nachricht und allen RRC Nachrichten darunter auf ‚emergency‘ gesetzt. Somit ist es möglich, diesem Bearer eine höhere Priorität im Netzwerk zu geben. Nach Aufbau des Bearers und einer SIP Emergency Registration sendet das Endgerät dann ein SIP ‚Emergency Invite‘ an das Netzwerk. Falls der Teilnehmer zuvor schon registriert war, kann das Netzwerk seine Teilnehmeridentität verifizieren. Falls der Teilnehmer noch nicht registriert war, z. B. weil das Endgerät aufgrund fehlender Netzwerkabdeckung ein fremdes LTE Netzwerk wählen musste, kann die Identität des Teilnehmers nicht überprüft werden. In den meisten Ländern ist es jedoch trotzdem erlaubt, den Notruf aufzubauen.

In der SIP ‚Invite‘ Nachricht geben zwei Parameter im Header dem Netzwerk weitere Informationen über die Art des Notrufs. Der allgemeine Fall sieht wie folgt aus:

```
INVITE urn:service:sos SIP/2.0
To: "112" <urn:service:sos>
```

Wurden weitere Notrufnummern für bestimmte Notrufzentralen definiert, z. B. 909 für die Feuerwehr, würde der Universal Resource Name (URN) wie folgt erweitert werden:

```
To: "909" <urn:service:sos.fire>
```

Zusätzlich sendet das Endgerät einen P-ANI (P-Access Network Identifier) Header mit, der die Cell-ID der Zelle enthält, in der sich der Teilnehmer gerade befindet. Der P-ANI Header ist nicht notrufspezifisch, sondern wird bei jedem VoLTE Gesprächsaufbau mitgesendet.

In manchen Ländern, wie z. B. den USA, müssen bei Aufbau eines Notrufs weitere Informationen über den Aufenthaltsort mitgegeben werden. In VoLTE wird dies über ‚Secure User Plane Location‘ (SUPL) Signalisierung während des Notrufaufbaus gemacht. Auf diese Weise teilt das Endgerät dem Notrufzentrum seine aktuellen GPS Koordinaten mit.

3.3 VoLTE Roaming

Während VoLTE heute in vielen Ländern und allen Teilen der Welt verbreitet ist, wird für die Sprachtelefonie im Roaming Fall aufgrund von fehlenden VoLTE Roaming Vereinbarungen und noch nicht durchgeföhrten Netzwerkupgrades hauptsächlich noch GSM und UMTS verwendet. Wenn sich heute ein VoLTE Endgerät in einem ausländischen Netz (VPLMN, Visited Public Land Mobile Network) registriert und dort kein VoLTE Roaming mit dem Heimatnetzwerk des Teilnehmers angeboten wird, deaktiviert es seine VoLTE Funktion und verhält sich wie ein ‚altes‘ Endgerät, das den Sprachdienst nur über CSFB (Circuit Switched FallBack) in GSM und UMTS Netzwerken unterstützt.

Bei erscheinen dieser Auflage ist jedoch erfreulicherweise zu beobachten, dass Netzbetreiber vermehrt VoLTE Roaming anbieten und viele Endgeräte dies nun auch unterstützen. Grundsätzlich gibt es zwei VoLTE Roaming Implementierungen: Die erste Lösung, die unter dem Namen ‚Local Breakout VoLTE Roaming‘ spezifiziert wurde, benötigt ein komplexes Zusammenspiel (Interworking) mit dem ausländischen Netz. Als Konsequenz daraus wurde etwas später dann eine sehr viel einfache Roamingvariante unter dem Namen ‚S8-Home Routing‘ standardisiert, die heute auch von den meisten Netzbetreibern, die VoLTE Roaming aktiviert haben, verwendet wird.

Um diese zwei Konzepte zu verstehen, ist es hilfreich, zunächst einen Blick auf das heutige Roamingsystem für IP Daten und leitungsvermittelte Gespräche zu werfen:

Wenn sich ein Endgerät im Ausland befindet und sich in ein VPLMN für den Internetzugang einbucht, kontaktiert die MME (Mobility Management Entity, vgl. Kap. 1) im Ausland die Home Subscriber Server (HSS) Datenbank im Heimatnetzwerk des

Teilnehmers, um Daten für die Authentifizierung und für die Aktivierung der Verschlüsselung zu bekommen. Nachdem ein Teilnehmer authentifiziert ist, kontaktiert die MME im Ausland dann das PDN-Gateway (P-GW) im Heimatnetzwerk des Teilnehmers und fordert für den Internetzugang den Aufbau eines Packet Data Bearers an. Der P-GW im Heimatnetzwerk teilt dem Endgerät dann eine IP Adresse zu und gibt diese an den MME im ausländischen Netzwerk zurück. Außerdem wird eine Verbindung zum ausländischen Serving-Gateway (S-GW) hergestellt, über das dann die IP Datenpakete des Nutzers geleitet werden. Das bedeutet, dass alle Datenpakete von und zum Endgerät und dem Internet immer zwischen dem ausländischen Netzwerk und dem P-GW im Heimatnetz ausgetauscht werden und erst von dort aus von und zum Internet gelangen. Dieses Konzept wird als Home Routing bezeichnet, da alle Datenpakete des Teilnehmers aus dem Ausland immer zunächst ins Heimatnetz weitergeleitet werden. Die Vorteile dieses Ansatzes sind, dass im Endgerät keine Konfigurationsänderungen notwendig sind und dass alle Dienste, die der Heimatnetzbetreiber seinen Kunden bietet, auch im Roaming Fall erreicht werden können. Nachteil dieses Ansatzes ist jedoch, dass lange Verzögerungszeiten (Delay) auftreten können, wenn der Nutzer sich sehr weit von seinem Heimatnetzwerk entfernt befindet. Neben den eigentlichen Nutzdaten müssen über die High Speed IP Verbindungen zwischen den Netzbetreibern auch Session Management Nachrichten ausgetauscht werden. Das Netzwerk, das Mobilfunknetzbetreiber untereinander verbindet, wird IPX (IP-Exchange) genannt²¹.

Gänzlich unterschiedlich funktioniert das Roaming für traditionelle Mobilfunktelefonie mit Mobile Switching Centern (MSCs). Bucht sich ein Endgerät in einem ausländischen Netz ein, führt es eine Location Update Prozedur mit der GSM oder UMTS MSC im Ausland durch. Falls ein LTE Netzwerk verwendet wird, führt das Endgerät einen ‚Combined LTE Attach‘ aus und die MME im Ausland registriert den Nutzer dann auch bei der zugehörigen ausländischen MSC. Details hierzu sind in Kap. 1 zu finden. Wenn der Nutzer im Ausland dann ein Gespräch zu einer Telefonnummer aus diesem Land aufbaut, kann die ausländische MSC das Gespräch ohne weitere Interaktion mit dem Heimatnetzwerk des Benutzers aufbauen. Das ausländische Netzwerk erzeugt dann einen ‚Billing Record‘ für das Gespräch und sendet diesen dann zum Heimatnetzwerk für die spätere Abrechnung. Wenn das Endgerät sich in einem ausländischen LTE Netzwerk befindet, führt es für den Anruf zunächst es eine Circuit Switched Fallback Prozedur (CSFB) nach GSM oder UMTS durch und fährt dann mit der oben beschrieben Prozedur fort. Das bedeutet also, dass leitungsvermittelte Telefonie während des Roamings komplett im Roamingnetzwerk verarbeitet werden kann und die Gespräche nicht in das Heimatnetz geroutet werden.

3.3.1 Option 1: VoLTE Local Breakout

Die erste VoLTE Roaming Option ist ab 3GPP Release 11 spezifiziert und wird als ‚Local Breakout‘ bezeichnet, da diese sich an der Funktionsweise der leitungsvermittelnden Telefonie im Roaming Fall orientiert²². Wenn ein VoLTE Roaming fähiges Endgerät sich im

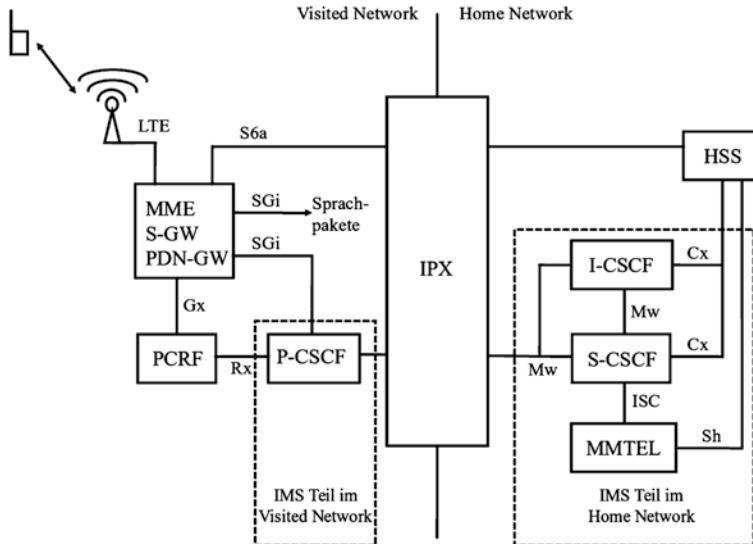


Abb. 3.14 VoLTE Local Breakout

ausländischen Netzwerk registriert und dieses Netz den VoLTE Local Breakout Mechanismus unterstützt baut, es einen IMS Default Bearer für die SIP Signalisierung auf. Das Netzwerk weiß, dass dieser Default Bearer nicht zum P-GW im Home Netzwerk aufgebaut werden soll, sondern zu einem lokalen P-GW. Somit hat das Endgerät Zugang zu einem P-CSCF im ausländischen Netzwerk. Dies ist wichtig, da so später ein Dedicated Bearer mit QoS Parametern für den Sprachdatenstrom aufgebaut werden kann. Die P-CSCF kann beim Verbindungsaufbau erkennen, dass das Endgerät zu einem ausländischen Teilnehmer gehört und leitet die SIP „Register“ Nachricht an einen I/S-CSCF im Heimatnetzwerk des Nutzers weiter. Zusätzlich zum normalen IMS SIP Registrierungsprozess wird noch eine „Transit and Routing Function“ (TRF) in den SIP Signalisierungspfad eingebunden. Die TRF wird benötigt, damit SIP Signalisierung und der Sprachdatenstrom für Gespräche aus dem Visited Network in das Zielnetzwerk weitergeleitet werden können. Die SIP Signalisierung wird dabei zunächst zur S-CSCF ins Heimatnetzwerk geschickt, von dort wieder zurück in das Visited Network und von dort aus dann zum Zielnetzwerk. Da die SIP Signalisierung und auch der Sprachdatenstrom aus Sicht des Zielnetzwerkes aus dem Visited Network kommen, ist es möglich, die gleiche Abrechnungsmethodik wie für leitungsvermittelte Gespräche zu verwenden (Abb. 3.14).

3.3.2 Option 2: VoLTE S8-Home Routing

Ein Nachteil der zuvor beschriebenen VoLTE Local Breakout Lösung ist, dass das ausländische Netzwerk eine IMS Infrastruktur benötigt und diese mit dem IMS Heimatnetzwerk des Roamers verbunden sein muss. Außerdem wird eine zusätzliche

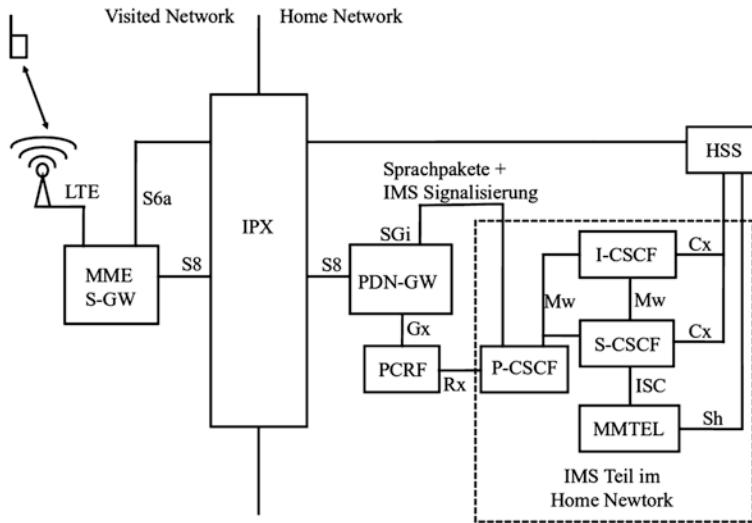


Abb. 3.15 VoLTE S8-Home Routing

Netzkomponente benötigt und beide IMS Netzwerke müssen VoLTE Roaming unterstützen. Da die Umsetzung dieses Systems in der Praxis sehr aufwendig ist, wurde in 3GPP eine weitere VoLTE Roaming Variante spezifiziert, die als VoLTE S8-Home Routing bezeichnet wird und in 3GPP TR 23.749 beschrieben ist²³. Der Name dieser Lösung stammt vom S8 Netzwerkinterface, das die MME im Ausland (Visited Network) mit dem P-GW im Heimatnetz eines Teilnehmers verbindet. Dies wird in Abb. 3.15 gezeigt. NTT DoCoMo in Japan war der erste Netzbetreiber, der diese Lösung eingeführt hat²⁴. Mittlerweile verwenden auch eine wachsende Anzahl Netzbetreiber in Europa und den USA diese Lösung.

Die Idee von VoLTE S8-Home Routing ist, keinen P-CSCF im Visited Network zu verwenden, sondern die komplette VoLTE Signalisierung und auch den Sprachpfad über die S8 Schnittstelle in das Heimatnetzwerk zu tunnen. Dies entspricht somit dem gleichen Prinzip wie die Nutzung des Internets über Mobilfunk im Roamingfall. Somit benötigt diese Implementierung keine zusätzlichen Funktionen im Visited Network und es wird auch keine Zusammenarbeit zwischen den zwei IMS Netzwerken im Visited- und Home Network benötigt. Theoretisch wäre es sogar möglich, VoLTE S8HR in einem Visited Network zu nutzen, das selber noch gar nicht VoLTE fähig ist.

Ein Nachteil dieser Lösung ist jedoch, dass eine VoLTE S8HR Lösung ohne weitere Zusätze keinen Dedicated Bearer im Visited Netzwerk für den Sprachpfad aufbauen kann. Die Sprachdatenpakete könnten dann somit nicht priorisiert werden. Ein fehlender Dedicated Bearer Setup würde auch bedeuten, dass die IMS VoLTE Applikation im Endgerät modifiziert werden müsste. In 3GPP wurden deshalb zwei Funktionalitäten spezifiziert, die für S8HR zu verwenden sind, also den Aufbau von Dedicated Bearern und

wie SRVCC nach GSM und UMTS während eines Gesprächs auch im Roamingfall verwendet werden kann. Dies wird wie folgt umgesetzt:

Wenn ein Endgerät ein VoLTE Gespräch in einem Visited Netzwerk aufbaut, kontaktiert die IMS im Home Network den P-GW für einen Dedicated Bearer mit einem QCI von 1 wie zuvor schon beschrieben. Der P-GW im Heimatnetz reicht diese Anforderung an die MME und den S-GW im Visited Network weiter und von dort wird der Request an das Radionetzwerk weitergeleitet. Damit dies funktioniert, muss der Betreiber des Visited Network dem Home Network des Teilnehmers erlauben, einen Dedicated Bearer über das IPX Roaming Exchange Netzwerk aufzubauen. Außerdem muss das Visited Network generell den Einsatz von Dedicated Bearern in seinem LTE Network unterstützen.

Der Handover eines VoLTE Telefonats nach GSM oder UMTS in einem Visited Network mit Single Radio Voice Call Continuity (SRVCC) wird wie folgt ausgeführt: Wie oben für SRVCC im Heimatnetz beschrieben, beginnt die Prozedur, nachdem der eNodeB bemerkt, dass die LTE Radiobedingungen schlechter werden und eine GSM oder UMTS Zelle verfügbar ist, in der das Telefonat weitergeführt werden kann. Der eNodeB weiß, dass sich das Endgerät in einem VoLTE Call befindet, da ein Dedicated Bearer mit QCI=1 aufgebaut ist und fordert deshalb von der MME einen Handover an. Die MME im Visited Network kontaktiert dann das lokale Mobile Switching Center (MSC), das wie auch eine MSC im Heimnetzwerk die Sv Schnittstelle unterstützen muss, damit es den Handover Request der MME empfangen kann. Teil der Handover Anfrage ist die Session Transfer Number für den SRVCC (STN-SR), die die MME vom HSS im Heimnetz des Teilnehmers während der Attach Prozedur erhalten hat. Wie die in Kap. 5 für GSM beschriebene Mobile Station Roaming Number (MSRN) hat diese die Aufgabe, den Teilnehmer temporär zu identifizieren. In diesem Kontext wird die STN-SR wie eine Telefonnummer verwendet, um damit eine Sprachverbindung zu einer anderen leitungsvermittelnden Komponente im Netzwerk aufzubauen. Dies ist z. B. eine andere MSC, oder in diesem Fall die IMS und ein Circuit Switched Media Gateway im Heimatnetzwerk. Im Heimatnetzwerk erkennen die IMS und das Media Gateway anhand der Telefonnummer in der SS7 IAM Nachricht (Initial Address Message, vgl. Kap. 5), die die STN-SR enthält, dass der ankommende Anruf Teil einer SRVCC Prozedur für einen Teilnehmer ist, der durch die STN-SR identifiziert wird. Die IMS im Heimnetzwerk leitet dann den Sprachpfad zum Media Gateway und somit zur MSC im Visited Network um. Weitere Details zu dieser Prozedur sind in 3GPP TS 23.216, Abschn. 7.2.2 zu finden²⁵.

3.4 Voice und 5G NR

In der Praxis gibt es heute 5G Non-Standalone (NSA) und 5G Standalone (SA) Netzwerke. In 5G Non-Standalone wird der IMS Sprachdienst über den LTE Anker einer 4G/5G Verbindung und über das LTE Kernnetz verwendet. Somit sind keine Änderungen

am System notwendig. In 5G NR Standalone Netzwerken kann der IMS Sprachdienst mit nur wenigen Änderungen weiterverwendet werden, auch wenn hier das 5G Zugangsnetz und das 5G Kernnetz verwendet werden. Da das IMS System grundsätzlich vom Zugangsnetzwerk und Kernnetz unabhängig ist, sind somit nur neue Schnittstellen für die Aktivierung der Quality of Service Mechanismen im 5G SA Kern- und Zugangsnetz notwendig.

Unterstützt ein Endgerät im 5G SA Mode sowie das 5G Netzwerk die notwendigen Quality of Service Mechanismen, wird sowohl die Signalisierung, als auch die IMS Sprachverbindung über das 5G Kern- und Zugangsnetzwerk geleitet. Aus dem ursprünglichen Voice over LTE (VoLTE) wird dann Voice over NR (VoNR). Dieses Voice Profil wird von der GSMA in NG.114 beschrieben²⁶. Wie auch die 3GPP Spezifikationen ist dieses Dokument im Internet kostenfrei erhältlich. Da die 5G Abdeckung in den nächsten Jahren jedoch noch geringer als die LTE Abdeckung in vielen Ländern sein wird, werden Mechanismen benötigt, während eines Gesprächs von 5G SA nach LTE zu wechseln.

Fehlt dem 5G SA Netzwerk oder dem Endgerät die Unterstützung der 5G Quality of Service Mechanismen, wird zwar die IMS Signalisierung über 5G SA geleitet, für Gespräche muss dann jedoch in das LTE Netzwerk gewechselt werden. Dies wird als EPS-Fallback bezeichnet. EPS steht dabei für Evolved Packet System und ist ein anderer Name für LTE. Der CS-Fallback Mechanismus, wie er noch bei LTE anzutreffen ist, wurde in 3GPP Release 15 für 5G nicht spezifiziert. Daraus folgt, dass 5G SA-fähige Endgeräte mit Telefoniefunktion auch zwingend das IMS System unterstützen müssen.

3.4.1 IMS Signalisierung über 5G SA

Möchte ein Endgerät den IMS Sprachdienst im 5G SA Netzwerk nutzen, setzt es nach dem Einschalten oder nach dem Verlassen des Flugmodus in der 5G SA Registration Request Nachricht (vgl. Kap. 2) zwei Parameter:

- **UE's Usage Setting:** Voice Centric. Endgeräte, die keinen Sprachdienst nutzen wollen, z. B. Tablets, setzen diesen Parameter stattdessen auf Data Centric.
- **5G Mobility Management (P5GMM) Capabilities:** S1=True (1). Damit teilt das Endgerät dem Netzwerk mit, dass es LTE unterstützt und somit auch EPS Fallback. Die Bezeichnung S1 stammt von Namen des Interfaces zwischen dem LTE eNB und den LTE Kernnetz.

Im weiteren Verlauf fragt das Netzwerk dann mit einer UE Capability Inquiry Nachricht, welche 5G Eigenschaften und Frequenzbänder das Endgerät unterstützt. Dieses antwortet dann mit einer UE Capability Information Nachricht. Wird neben EPS-Fallback auch Voice over NR unterstützt, fügt das Endgerät hier den VoiceOverNR Parameter ein und setzt diesen auf „Supported“.

Im Gegenzug teilt das Netzwerk dem Endgerät in der Registration Accept Nachricht mit, welche Funktionen für den IMS Voice Service im Netz unterstützt werden:

- **IMSVoPS3GPP:** Ist dieses Flag auf „True“ gesetzt, wird EPS-Fallback oder VoNR vom Netzwerk unterstützt. Ist dieses Flag auf „False“ gesetzt, wechseln „Voice Centric“ Geräte sofort nach LTE, da dann keine IMS Registrierung im Netzwerk möglich ist. Öffentliche 5G SA Netzwerke haben deshalb dieses Flag typischerweise gesetzt. Für andere Einsatzfälle, z. B. Campus Netzwerke, macht es jedoch Sinn, dieses Flag auf „False“ zu setzen, falls IMS kein Teil des Netzes ist. Das Campus Netzwerk kann dann jedoch nicht mehr mit Voice Centric Geräten verwendet werden. Ist das Flag gesetzt, werden Voice Centric Geräte dann auch sofort eine IMS Registrierung durchführen. Ist diese nicht erfolgreich, wird ebenfalls zu LTE gewechselt.
- **EMC:** IMS Emergency Call Support. Ist das Flag gesetzt, kann das 5G SA Netzwerk für Notrufe verwendet werden.
- **EMF:** IMS Emergency Call Fallback Support: Ist dieses Flag gesetzt, werden Notrufe über LTE unterstützt. Sind weder EMC noch EMF gesetzt, muss das Endgerät für Notrufe nach einem GSM oder UMTS Netzwerk suchen.
- **IWKN26:** Dieser Parameter gibt an, ob das N26 Interface zwischen dem 5G SA Kernnetz und dem 4G LTE Kernnetz vorhanden ist. Ist der Parameter auf 0 gesetzt, ist N26 vorhanden und ein schneller Wechsel zwischen 5G und 4G ist somit möglich. Ist der Parameter auf 1 gesetzt, ist das N26 Interface nicht vorhanden. Hier wird sehr untypisch also eine negative Logik verwendet.

Nachdem das Endgerät im 5G SA Netz registriert ist, wird dann neben der PDU Session für den Internetzugang eine weitere PDU Session mit dem APN/DNN „ims“ aufgebaut. Dies funktioniert in gleicher Weise wie schon bei LTE. Wichtig ist hier, dass die PDU Session Establishment Accept Nachricht eine Mapping Information der PDU Session auf einen LTE Bearer beinhaltet. Nur so ist es möglich, den Bearer später nach LTE zu übertragen. Fehlt diese Information, muss das Endgerät einen kompletten LTE Initial Attach durchführen und die PDU Sessions gehen verloren.

Sobald die IMS PDU Session aufgebaut ist, startet dann die IMS SIP Registrierung. Diese ist bis auf die Verwendung einer 5G Cell ID statt einer 4G Cell ID in der SIP REGISTER Nachricht identisch. Von einer 5G NR Zelle in TDD Band n78 sieht diese wie folgt aus:

P-Access-Network-Info:

3GPP-NR-TDD;utran-cell-id-3gpp=2620900AB313DD5E0047

3.4.2 5G NR EPS Fallback

Erste 5G SA Netze und Endgeräte unterstützen in der Praxis keine Quality of Service Mechanismen für die Sprachdatenpakete und schicken ein Endgerät während des Gesprächsaufbaus in das LTE Kern- und Zugangsnetzwerk. Nach dem Ende des Gesprächs kann das Endgerät dann wieder zu 5G SA zurückkehren. Im Standard wurden dafür zwei Varianten spezifiziert. Mit der etwas aufwendigeren Variante werden beim Gesprächsaufbau alle Bearer mit einem aktiv vom Netzwerk gesteuerten Handover von 5G nach LTE übergeben. Viele Netzbetreiber werden jedoch zunächst eher die etwas weniger aufwendigere EPS Fallback Variante verwenden, die das Endgerät mit einer RRC Release mit Redirect Nachricht nach LTE schickt. Während dies einfach zu implementieren ist, verzögert sich dadurch jedoch der Gesprächsaufbau. Die Signalisierung dieser Variante ist in 3GPP TS 23.502 in Abschn. 4.13.6.1 spezifiziert²⁷ und geschieht wie folgt:

Für ein abgehendes Gespräch sendet das Endgerät eine SIP INVITE Nachricht. Das IMS System sucht daraufhin den Zielteilnehmer und kontaktiert auch das 5G Kernnetz, um für die Sprachdatenpakete eine höhere Quality of Service anzufordern. Während im LTE Netz die PCRF dafür zuständig war, wird im 5G SA Kernnetz die PCF Funktion verwendet. Um am IMS System möglichst wenig ändern zu müssen, hat diese in Richtung des IMS Systems die bereits von LTE bekannte Schnittstelle. Über die PCF wird dann versucht, die notwendigen Quality of Service Parameter im Radionetzwerk, also dem gNB, für den Teilnehmer zu konfigurieren. Dies geschieht mit einer PDU Session Resource Modify Request Nachricht. Ist dies im gNB noch nicht implementiert, startet dieser daraufhin sofort die EPS Fallback Prozedur für den Teilnehmer. Optional konfiguriert der gNB dann zunächst das UE, inter-RAT LTE Messungen auf der Luftschnittstelle durchzuführen. Nach Antwort des Endgeräts oder einem Timeout schickt der gNB dann eine RRC Release Nachricht mit einer Redirect to LTE Information zum Endgerät. In Richtung des Kernnetz wird die Resource Modify Nachricht mit der Information abgelehnt, dass eine Mobility Prozedur nach LTE durchgeführt wird. Dies veranlasst das 5G Kernnetz dann, die Prozedur nicht abzubrechen, sondern zunächst zu warten.

Das Endgerät wechselt nach dem RRC Release nach LTE und führt dort einen LTE Tracking Area Update (TAU) durch. Über die dort enthaltene bisherige MME ID, die aus der AMF ID gebildet wurde, fordert die MME dann über die N26 Schnittstelle die Context Informationen von der AMF an. Außerdem werden die Internet- und IMS PDU Sessions zu einem LTE Serving-Gateway transferiert. Nachdem dies durchgeführt wurde, sendet die MME eine TAU Accept Nachricht. Außerdem baut das Netzwerk nun sofort den LTE Dedicated Bearer auf, der für die Sprachpakete verwendet wird. Ab diesem Zeitpunkt wird dann der IMS Gesprächsaufbau, wie schon von LTE gewohnt, weitergeführt. Wenn man sich die SIP Signalisierung während der gesamten Prozedur anschaut, wird man keinen Unterschied zum Aufbau eines VoLTE Calls sehen, der

Wechsel von 5G nach 4G LTE ist also auf IP Ebene völlig transparent. Das Endgerät kann jedoch während des Gesprächs eine SIP Re-Register Nachricht schicken, um das IMS System vom Wechsel der Radiotechnologie zu informieren.

Nachdem das Gespräch beendet wurde, kann das LTE Netzwerk eine RRC Connection Release Nachricht mit Redirect nach 5G SA senden. Dies ist jedoch nur optional. Wird dies nicht gemacht, verwendet das Endgerät zunächst weiter LTE, und, sofern vorhanden, 5G NSA. Erst nach dem Wechsel in den RRC-Idle Zustand wechselt das Endgerät dann wieder zur 5G SA Zelle in einem anderen Frequenzband. Auch dies kann das Netzwerk verhindern, in dem es keine entsprechenden inter-RAT Informationen in den LTE SIBs sendet. Somit wechselt dann das Endgerät nur dann nach 5G SA zurück, wenn es vom Netzwerk im RRC-Connected State entsprechende Anweisungen bekommt.

3.4.3 5G Voice over NR (VoNR)

Unterstützen sowohl das Endgerät als auch das Netzwerk die nötigen 5G SA Quality of Service Mechanismen, ist kein EPS-Fallback nach LTE beim Gesprächsaufbau nötig. Die 5G Basisstation (gNB) führt dann die Anweisungen in der PDU Session Resource Modify Request Nachricht aus und schickt dem Endgerät daraufhin eine RRC PDU Session Modification Command Nachricht. Diese Nachricht enthält unter anderem eine Paketfilterkonfiguration, in der beschrieben ist, von und zu welchen IP Adressen eine andere Quality of Service für die Sprachdatenpakete verwendet wird. Außerdem wird dem Endgerät in dieser Nachricht mitgeteilt, auf welche maximale Geschwindigkeit der Sprachdatenstrom limitiert ist.

Unterstützt das 5G SA Netzwerk VoNR, muss es auch in der Lage sein, ein IMS Gespräch am Rand der 5G Abdeckung nach LTE zu übergeben. Dies geschieht mit einem inter-RAT 5G nach LTE Handover, der in Kap. 2 beschrieben wurde und im Detail in 3GPP TS 23.502, Abschn. 4.11.1.2²⁸ beschrieben ist.

3.5 Voice over Wifi (VoWifi)

Wie schon am Anfang dieses Kapitels besprochen, wurde der IMS Service für LTE (VoLTE) so entwickelt, dass er möglichst unabhängig vom LTE Kernnetz und besonders vom LTE Zugangsnetzwerk (Radio Access) ist. Von der Schnittstelle für die Anfrage von Quality of Service für den Sprachpfad ist dies auch tatsächlich der Fall. Um den Dienst nicht nur über Mobilfunk nutzen zu können, spezifizierte 3GPP auch eine Möglichkeit, den VoLTE Dienst über ‚untrusted non-3GPP access‘, also über das Internet, zu verwenden. Da Smartphones und andere Endgeräte typischerweise Wi-Fi als weitere Schnittstelle zum Internet verwendet, wird diese Erweiterung als ‚Voice over Wifi‘ oder ‚VoWifi‘ bezeichnet. Nachfolgend wird VoWifi, wie in GSMA IR.51 definiert,

beschrieben²⁹. Diese VoWifi Variante wird von den meisten Netzbetreibern, die diesen Dienst anbieten, heute verwendet. Auf der Endgeräteseite ist VoWifi voll in die Telefoniefunktionalität integriert, es gibt also nur eine Applikation für die Sprachtelefonie und der Nutzer muss keine zusätzliche Software installieren. Abgesehen von einer Indikation in der Statusleiste ist es für den Nutzer transparent, ob für ein Telefonat Wifi statt LTE genutzt wird.

3.5.1 VoWifi Netzwerkarchitektur

Abb. 3.16 zeigt, wie das LTE und IMS Netzwerk für VoWifi erweitert werden muss. Die einzige zusätzliche Komponente ist das „evolved Packet Data Gateway“ (ePDG), das in 3GPP TS 23.402 spezifiziert wurde³⁰. Das ePDG ist das Gateway zwischen dem Internet auf der einen Seite und dem Mobilfunknetzwerk und der IMS Plattform eines Netzbetreibers auf der anderen Seite. Auf der Seite des Internets verhält sich das ePDG wie ein Virtual Private Network (VPN) Gateway. Auf der Mobilfunkseite verhält sich das ePDG wie eine LTE MME (Mobility Management Entity) und ein S-GW (Serving-Gateway), die in Kap. 1 besprochen wurden.

Ein Endgerät verbindet sich mit dem Mobilfunknetzwerk und dem IMS System über Wifi durch den Aufbau eines VPN Tunnels zum ePDG. Hierzu wird das IPsec (IP Security) Protokoll verwendet, das auch bei anderen VPN Servern Anwendung findet, z. B. bei der Anbindung von Home Office Mitarbeitern. In Abschn. 8.2 in 3GPP TS 33.402³¹ ist spezifiziert, welche IPsec Konfiguration für VoWifi verwendet wird. Wie bei LTE und VoLTE wird für die Authentifizierung und die Erzeugung der Session Keys während des ePDG Verbindungsaufbaus der geheime Schlüssel Ki und ein Algorithmus

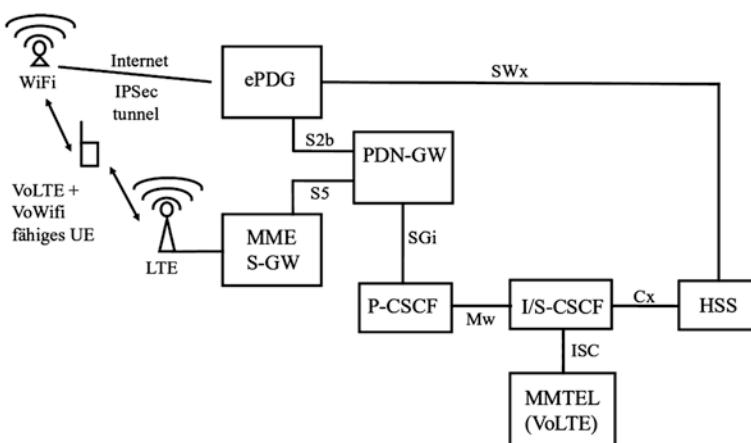


Abb. 3.16 VoWifi Netzwerkarchitektur und der ePDG

auf der SIM Karte und im HSS (Home Subscriber Server) auf der Netzwerkseite verwendet (vgl. Kap. 1). Während des Verbindungsaufbaus authentifizieren sich das Endgerät und der ePDG gegenseitig.

Abb. 3.17 zeigt einen typischen VoWiFi IPSec Verbindungsaufbau. Im ersten Schritt ermittelt das Endgerät die IP Adresse des ePDG seines Mobilfunknetzwerkes. Dazu erzeugt das Endgerät einen Fully Qualified Domain Name (FQDN) und sendet diesen für die Adressauflösung zu einem DNS Server im Internet. Der Aufbau des FQDNs für ePDGs ist standardisiert und enthält den Mobile Country Code (MCC) und den Mobile Network Code (MNC) des Heimatnetzwerkes. Der MCC und der MNC sind auch Teil der International Mobile Subscriber Identity (IMSI) des Teilnehmers, der auf der SIM-Karte gespeichert ist. Die generische FQDN eines ePDG sieht wie folgt aus:

epdg.epc.mncXXX.mccXXX.pub.3gppnetwork.org

Nachdem der DNS Server die Anfrage des Endgeräts beantwortet hat, sendet dieses dann das erste IPSec Paket an den ePDG. Dieses Paket enthält zunächst die unterstützten Authentifizierungs- und Verschlüsselungsalgorithmen. Der ePDG sucht sich dann eine auch von ihm unterstützte Kombination aus und sendet eine Antwort zurück. Um keine privaten Daten wie z. B. die IMSI während der Authentifizierungsprozedur unverschlüsselt zu übertragen, wird während dieser Phase ein nicht-authentifizierter temporärer Tunnel aufgebaut, der verschlüsselt ist. Weitere Details finden sich in RFC 5282³².

Nach Aufbau des temporären Tunnels überträgt das Endgerät seine Identität (IMSI) und informiert den ePDG, welche Netzwerkparameter es benötigt, um eine Verbindung durch den finalen IPSec Tunnel mit dem Kernnetzwerk und dem IMS System aufzubauen zu können. Dies sind:

- Eine IPv4 Adresse und/oder ein IPv6 Prefix für sich selber
- Die IPv4 Adresse und/oder IPv6 Adresse eines DNS Servers im Netzwerk
- Die IP Adresse(n) der P-CSCFs

Nachdem der ePDG die Anfrage des Endgeräts erhalten hat, fragt es dann beim Home Subscriber Server (HSS) nach den Authentifizierungsinformationen für den Teilnehmer und führt die gegenseitige Authentifizierung durch. Diese Prozedur stellt auch sicher, dass keine Man-In-The-Middle Attacke möglich ist. Dies kann z. B. durch die Verwendung des EAP-AKA Protokolls in Kombination mit einer Public/Private Certificate Authentifizierung wie in 3GPP TS 33.402, Kap. 8.2 beschrieben, gewährleistet werden. In der Praxis kann jedoch auch eine EAP-only Prozedur ohne Zertifikat verwendet werden, wie sie in RFC 5998 beschrieben ist³³. Während der Authentifizierungsprozedur wird der auch bei UMTS und LTE verwendete Authentifizierungsalgorithmus auf der SIM-Karte ausgeführt, wo sich auch der geheime Schlüssel Ki befindet. Das Ergebnis wird an das Endgerät zurückgegeben. Das Endgerät erzeugt damit dann eine Antwort für

das Netzwerk, prüft die Authentizität des Netzwerks und erstellt die Schlüssel für die Verschlüsselung.

Nachdem der IPSec Tunnel aufgebaut wurde, führt das Endgerät wie am Anfang des Kapitels gezeigt, eine normale IMS VoLTE Registrierung durch. Abb. 3.17 zeigt unter anderem das im IPSec Tunnel verschlüsselt übertragene IMS SIP ‚Register‘ Paket (Paket 1093 mit einer Länge von 1374 Bytes). Am ePDG angekommen wird das Paket entschlüsselt und dann weiter in das Kernnetz des Netzbetreibers gesendet. An dieser Stelle sei angemerkt, dass die IMS Registrierung, wie schon zuvor beschrieben, ebenfalls den Aufbau eines IPSec Tunnels vorsieht, und zwar zwischen Endgerät und der P-CSCF. Das bedeutet, dass es einen IMS IPSec Tunnel gibt, der im inneren des IPSec Tunnels transportiert wird, der zwischen dem Endgerät und dem ePDG aufgebaut wurde.

3.5.2 VoWifi Handover

Eine wichtige Funktionalität im Zusammenhang mit VoWifi ist die Übergabe eines laufenden Gesprächs zwischen VoWifi und VoLTE in beiden Richtungen. Dies ist möglich, da der ePDG sich wie ein MME/Serving-Gateway im LTE Kernnetz verhält und ein Übergang zwischen Wifi und LTE ähnlich wie ein Inter-MME/Inter-Serving Gateway Handover ausgeführt werden kann.

Von außen betrachtet bedeutet ein Handover von LTE nach Wifi, dass die IP Adresse des bestehenden LTE IMS Default Bearer für den IPSec Tunnel wiederverwendet wird, der während des ePDG Verbindungsaufbaus erzeugt wird. Der Signalisierungsaustausch für diesen Handover zwischen Endgerät und ePDG entspricht bis auf einen Unterschied dem zuvor besprochenen initialen ePDG IPSec Tunnel Aufbau. Der einzige Unterschied ist, dass das Endgerät keine neue IP Adresse für sich anfordert, sondern den ePDG

1017	14:09:06.460583	192.168.65.30	192.168.65.1	DNS	102	Standard query 0x31fe A epdg.epc.mnc001.mcc26
1020	14:09:06.492569	192.168.65.1	192.168.65.30	DNS	486	Standard query response 0x31fe A epdg.epc.mnc001.mcc26
1048	14:09:12.169353	192.168.65.30	109.237.187.230	ISAKMP	402	IKE_SA_INIT MID=00 Initiator Request
1049	14:09:12.194955	109.237.187.230	192.168.65.30	ISAKMP	94	IKE_SA_INIT MID=00 Responder Response
1050	14:09:12.208448	192.168.65.30	109.237.187.230	ISAKMP	426	IKE_SA_INIT MID=00 Initiator Request
1051	14:09:12.241551	109.237.187.230	192.168.65.30	ISAKMP	330	IKE_SA_INIT MID=00 Responder Response
1052	14:09:12.813614	192.168.65.30	109.237.187.230	ISAKMP	442	IKE_AUTH MID=01 Initiator Request
1053	14:09:12.885449	109.237.187.230	192.168.65.30	ISAKMP	186	IKE_AUTH MID=01 Responder Response
1054	14:09:12.901804	192.168.65.30	109.237.187.230	ISAKMP	186	IKE_AUTH MID=02 Initiator Request
1057	14:09:13.230498	109.237.187.230	192.168.65.30	ISAKMP	218	IKE_AUTH MID=02 Responder Response
1058	14:09:13.525740	192.168.65.30	109.237.187.230	ISAKMP	154	IKE_AUTH MID=03 Initiator Request
1061	14:09:14.114988	109.237.187.230	192.168.65.30	ISAKMP	122	IKE_AUTH MID=03 Responder Response
1062	14:09:14.228138	192.168.65.30	109.237.187.230	ISAKMP	138	IKE_AUTH MID=04 Initiator Request
1063	14:09:14.256081	109.237.187.230	192.168.65.30	ISAKMP	458	IKE_AUTH MID=04 Responder Response
1069	14:09:16.649997	109.237.187.230	192.168.65.30	ESP	174	ESP (SPI=0x391b27bc)
1082	14:09:19.655215	109.237.187.230	192.168.65.30	ESP	174	ESP (SPI=0x391b27bc)
1087	14:09:21.942410	192.168.65.30	109.237.187.230	ESP	174	ESP (SPI=0x17e15b28)
1090	14:09:22.136385	192.168.65.30	109.237.187.230	ESP	158	ESP (SPI=0x17e15b28)
1091	14:09:22.166072	109.237.187.230	192.168.65.30	ESP	158	ESP (SPI=0x391b27bc)
1092	14:09:22.180853	192.168.65.30	109.237.187.230	ESP	142	ESP (SPI=0x17e15b28)
1093	14:09:22.180868	192.168.65.30	109.237.187.230	ESP	1374	ESP (SPI=0x391b27bc)
1094	14:09:22.211408	109.237.187.230	192.168.65.30	ESP	142	ESP (SPI=0x391b27bc)

Abb. 3.17 Aufbau eines ePDG VPN Tunnels

informiert, dass es bereits eine IP Adresse für einen Default Bearer gibt, die nun wieder verwendet werden soll. Um dies zu erreichen, kontaktiert der PDN-Gateway die bisher zuständige MME und den Serving-Gateway (S-GW), informiert diese über den Wechsel und ändert dann das Ziel seiner Route für den GTP Kernnetztunnel des Teilnehmers vom S-GW zum ePDG. Trotz der vielen Aktionen die dafür notwendig sind, darf die Prozedur nur wenige hundert Millisekunden dauern, um die Unterbrechung auf dem Sprachkanal so kurz und unauffällig wie möglich zu halten.

Ein Handover eines VoWifi Gesprächs nach LTE wird durchgeführt, wenn der Nutzer mit seinem Endgerät den Abdeckungsbereich des Wifi Netzwerkes verlässt. Falls dann ein LTE Netzwerk in Reichweite ist, ist das Endgerät dort schon angemeldet (Attached) und ein Default Bearer für den Internetzugang ist ebenfalls schon vorhanden. Dies ist notwendig, da ein im LTE Netzwerk angemeldetes Endgerät mindestens einen Default Bearer aufgebaut haben muss. Der Bearer wird zu dieser Zeit nicht verwendet, da das Endgerät über Wifi mit dem Internet verbunden ist. Um das VoWifi Gespräch in das LTE Netzwerk zu übergeben, sendet das Endgerät einen PDN Connectivity Request. Der ‚Request Type‘ Parameter wird dabei nicht auf ‚Initial Request‘ sondern auf ‚Handover‘ gesetzt. Die MME startet dann den Transfer des IMS Default Bearer zu sich. Als Teil dieses Prozesses ändert der PDN-GW dabei seine Routing Tabelle, d. h. er tauscht den ePDG als Quelle und Ziel von Paketen für diese Verbindung mit dem MME/S-GW.

Der IMS Default Bearer wird nicht nur zwischen LTE und Wifi während eines Gespräches übergeben, sondern auch im Idle Zustand wenn gerade kein Telefonat geführt wird. Der Grund dafür ist, dass aus logischer Sicht ein Handover nicht für das Telefongespräch gemacht wird, sondern für den IMS Default Bearer, also für die IP Adresse. Für IMS und VoLTE ist der Transfer von einem Radionetzwerk zu einem anderen transparent, da das Endgerät ja seine IP Adresse behält und die IP Pakete im Kernnetz des Netzbetreibers nur an unterschiedliche Router weitergeleitet werden. Auch wenn dies aus technischer Sicht nicht nötig ist, wird die IMS und der VoLTE TAS trotzdem über den Zugangswechsel durch eine IMS Reregistration Prozedur informiert, in der der P-Access-Network-Identifier Parameter die Technologie des neuen Radionetzwerkes identifiziert. Diese Registrierung hat keinen Einfluss auf ein eventuell gerade stattfindendes Gespräch und das Gespräch würde auch weiterlaufen, wenn die Re-Registration nicht durchgeführt werden würde.

3.5.3 Wifi-Preferred und Cellular-Preferred

In der Praxis gibt es heute drei Betriebsmodi für VoWifi. Im „Cellular-Preferred“ Modus wird ein 2G, 3G oder LTE Netzwerk für Telefonie bevorzugt, auch wenn sich das Endgerät in einem Wifi Netzwerk befindet. Das Wifi Netzwerk wird dann nur für den Internetzugang verwendet. VoWifi wird in diesem Modus nur verwendet, wenn kein Mobilfunknetzwerk verfügbar ist, also z. B. in Kellern, in ländlichen Gegenden oder wenn das Endgerät im Flugmodus ist und Wifi aktiviert wurde. Ein Vorteil dieses

Modus ist, dass die Sprachqualität vom Netzwerk sichergestellt werden kann, solange das Endgerät ein Mobilfunknetzwerk empfangen kann. In Wifi Netzwerken kann dies üblicherweise nicht gewährleistet werden, da die verschlüsselten Sprachpakete anderen Paketen nicht bevorzugt werden. Dies ist vor allem dann der Fall, wenn das Wifi Netzwerk an einem DSL oder Kabelanschluss mit sehr limitiertem Uplinkkanal verwendet wird. Die Sprachqualität und die Verzögerungszeit werden schnell schlechter, sobald andere Endgeräte im Netzwerk ebenfalls Daten im Uplink übertragen. Im Downlink kann die Sprachqualität auch negativ beeinflusst werden, wenn z. B. Video Streaming den Großteil der verfügbaren Bandbreite benötigt.

Ein Nachteil des Cellular-Preferred Modus ist, dass ein Handover eines Gesprächs nur zwischen LTE und Wifi möglich ist. Ein Telefongespräch, das im 2G oder 3G Netzwerk aufgebaut wurde und somit leitungsvermittelt ist, kann bei Problemen mit der Netzabdeckung nicht nach VoWifi übergeben werden, da hier ja eine paketvermittelnde und IP basierte Übertragung des Gesprächs benötigt wird. Statt einem Handover bricht das Gespräch in einem solchen Szenario ab. Auch ein Gespräch, das über VoLTE im LTE Netzwerk aufgebaut wurde, dann aber über den SRVCC Mechanismus in ein 2G oder 3G Netzwerk übergeben wurde, kann ebenfalls nicht nach VoWifi übergeben werden. In der Praxis kommt es deshalb auf den VoWifi Handover Algorithmus des Endgeräts an, ob der Handover von LTE nach Wifi eines Gesprächs früher vom Endgerät eingeleitet wird, als der Handover von LTE nach 2G oder 3G, der vom Netzwerk gesteuert wird.

Im VoWifi „Wifi-Preferred“ Modus verbindet sich das Endgerät mit dem ePDG, sobald eine Wifi Verbindung vorhanden ist. Telefonate werden dann immer über Wifi aufgebaut, auch wenn ein Mobilfunknetzwerk empfangen werden kann. Dies vermeidet auf der einen Seite Probleme mit 2G/3G Gesprächsabbrüchen, auf der anderen Seite können aber Sprachqualitäts- und Laufzeitprobleme auftreten, wenn die Anbindung des Wifi Netzwerkes durch andere Endgeräte überlastet wird.

Und schließlich gibt es einen „IMS-preferred“ Modus. In dieser Konfiguration präferiert das Endgerät LTE und Wi-Fi Netzwerke vor 2G oder 3G Netzen. Auf diese Weise wird versucht, möglichst wenige Handover für laufende Gespräche nach 2G oder 3G durchzuführen, da es aus diesen Netzen während eines laufenden Gesprächs nicht möglich ist, nach LTE oder Wi-Fi zurückzukehren.

Welcher der drei VoWifi Modi verwendet wird, hängt von der Präferenz des Netzbetreibers ab und ob das Endgerät dem Nutzer die Möglichkeit gibt, den Modus auch selber zu wechseln.

3.5.4 SMS, MMS und Supplementary Services über Wifi

Zusätzlich zum IMS Sprachdienst wird die VoWifi Lösung auch verwendet, um weitere Netzbetreiberdienste über die IPSec Verbindung zum ePDG zu tunneln. Nachdem ein Gerät im VoLTE Netzwerk registriert ist, entweder über LTE oder über Wifi, können SMS Nachrichten über SIP gesendet und empfangen werden. SMS Nachrichten

werden dann nicht mehr über die 2G, 3G oder LTE Signalisierungskanäle übertragen. Da das Zugangsnetzwerk für das IMS System transparent ist, sind somit für die SMS Übertragung über VoWifi weder auf der Netzwerkseite noch auf der Endgeräteseite Erweiterungen nötig.

Auch Supplementary Services wie z. B. das Ändern der Einstellungen für die Gesprächsweiterleitung sind IP basiert und werden typischerweise über den LTE Default Bearer für den Internetzugang oder einen anderen vom Netzbetreiber definierten LTE Default Bearer übertragen. Da VoWifi zunächst nur den IMS Default Bearer über den IPSec Tunnel überträgt, ist es über diesen also nicht möglich, den Service im Netzwerk zu kontaktieren. Es muss deshalb also z. B. zum Ändern der Einstellungen für die Gesprächsweiterleitung temporär ein zweiter IPSec Tunnel zum ePDG für den für diesen Dienst genutzten Default Bearer aufgebaut werden. Nachdem die Einstellungen geändert wurden, wird der Tunnel dann wieder abgebaut. Dieses Verfahren wird auch verwendet, um MMS Nachrichten über (Vo)Wifi zu übertragen.

3.5.5 VoWifi Roaming

Im Prinzip ist VoWifi ein IP basierter Dienst und ist somit zunächst einmal aus technischer Sicht nicht an das Land eines Netzbetreibers gebunden. Trotzdem gibt es Unterschiede zwischen der Nutzung von VoWifi im Heimatland und im Ausland.

Die einzige technische Limitation heute ist, dass VoLTE nur im Heimnetzwerk funktioniert, da die meisten Netzbetreiber VoLTE Roaming heute noch nicht unterstützen. Das bedeutet, dass es nicht möglich ist, ein Handover im Ausland eines VoWifi Telefonats nach LTE durchzuführen. Wenn sich also ein Endgerät nicht sicher ist, ob VoLTE Roaming am aktuellen Aufenthaltsort unterstützt wird, darf es deshalb keine PDN Connectivity Request Nachricht vom Typ ‚Handover‘ schicken. Daraus folgt, dass im Ausland ein VoWifi Telefonat an der Grenze der Wifi Hotspot Abdeckung abbricht.

Eine nicht-technische VoWifi Limitierung im Roamingfall ist, dass manche Netzbetreiber keine Verbindung zum ePDG über das Internet zulassen, wenn sie feststellen, dass sich der Teilnehmer nicht im Heimatland befindet. Dies kann er z. B. ermitteln, indem er die IP Adresse des Teilnehmers während des Verbindungsaufbaus analysiert oder im HLR überprüft, ob sich der Teilnehmer zuvor in einem ausländischen Mobilfunknetzwerk eingebucht hat. Ob es klug ist, den VoWifi Dienst im Ausland zu blockieren, ist fraglich.

Für Telekommunikationsnetzbetreiber gibt es in manchen Ländern Gesetze, die vorschreiben, dass zu allen Gesprächen, die in diesem Land aufgebaut werden, Zugang durch sogenannte ‚Bedarfsträger‘ (Polizeibehörden, etc.) zu ermöglichen ist. Da dies nicht möglich ist, wenn der Teilnehmer im Ausland einen IPSec Tunnel zu einem ePDG im Heimnetzwerk in einem anderen Land aufbaut, hat 3GPP in TS 23.402, Abschn. 4.5.4, mehrere Möglichkeiten spezifiziert, wie Endgeräte erkennen können, ob sie sich mit einem ePDG im Roaming Land verbinden sollen oder ob sie sich zum ePDG

im Heimatland verbinden sollen. Gibt es im Ausland einen ePDG, der Verbindungen von VoWiFi Roamern zulässt, muss dieser auch mit dem Heimnetz des Nutzers verbunden sein. In der Praxis wird sich zeigen, ob Endgeräte wie heute immer einen Tunnel zum ePDG im Heimatland aufbauen oder ob lokale ePDGs in Zukunft tatsächlich verwendet und von Endgeräten auch genutzt werden müssen.

3.6 VoLTE und Festnetz IMS – Ein Vergleich

Aus historischer Sicht ist es interessant, dass SIP als Teil des IP Multimedia Subsystem (IMS) im Mobilfunk zunächst ein relativ einfaches Protokoll war, das gar nicht für den Mobilfunk gedacht war. Nachdem die Entwicklung in 3GPP für die mobile Welt schon sehr fortgeschritten war, entschlossen sich Netzbetreiber und Hersteller, für die nächste Generation der Festnetztelefonie ebenfalls die 3GPP IMS Spezifikationen zu verwenden. Wenn man den SIP Gesprächsaufbau eines Festnetz IMS Endgerätes wie in Abb. 3.18 betrachtet, sieht dies zunächst einem VoLTE Gesprächsaufbau im Mobilfunknetz sehr ähnlich. Während die SIP ‚Invite‘, ‚100 Trying‘, ‚183 Session Progress‘ und die ‚200 OK‘ Nachrichten auch in VoLTE verwendet werden, gibt es auch Unterschiede und Nachrichten, die nicht verwendet werden. Große Unterschiede sind z. B., dass es im Festnetz keinen IPSec Tunnel gibt, es werden nur ‚well known‘ UDP Ports verwendet, TCP wird für SIP gar nicht verwendet und es gibt keinen Precondition Mechanismus, keine Bandbreitenaushandlung, keine Aushandlung von Early-Media und keinen SRVCC Mechanismus.

Eine SIP Nachricht, die bei VoLTE nicht verwendet wird, ist die SIP ‚407 Proxy Authentication Required‘ Nachricht. Diese wird im Festnetz nach der SIP ‚Invite‘ Nachricht gesendet und wird benötigt, da kein IPSec verwendet wird, um eine authentifizierte und gesicherte Verbindung zwischen SIP User Agent und dem Festnetz IMS System herzustellen. Somit muss das Endgerät vor jedem Gesprächsaufbau authentifiziert werden.

Weiterhin verwenden IMS Festnetzsysteme andere Sprachcodecs als in der Mobilfunkwelt, die im folgenden Ausschnitt aus dem SDP (Session Description Protocol) Teil einer SIP Nachricht gezeigt werden:

Filter: ((sip.Status-Code sip.Method) && !tcp.analysis.retr.)							Expression...	Clear	Apply	Save	IPv6 Prefix Filter
No.	Time	Source	Destination	Protocol	Dst Prt	Src Prt	Length	Info			
330	16:11:48.374488	192.168.2.1	111.0.11.111	SIP/SDP	5060	5060	1258	Request: INVITE sip:040428990@tel.xyzabcde.de			
331	16:11:48.572888	111.0.11.111	192.168.2.1	SIP	5060	5060	542	Status: 407 Proxy Authentication Required 279932023			
332	16:11:48.575482	192.168.2.1	111.0.11.111	SIP	5060	5060	427	Request: ACK sip:040428990@tel.xyzabcde.de			
334	16:11:48.584495	192.168.2.1	111.0.11.111	SIP/SDP	5060	5060	54	Request: INVITE sip:040428990@tel.xyzabcde.de			
336	16:11:48.815714	111.0.11.111	192.168.2.1	SIP	5060	5060	338	Status: 100 Trying			
349	16:11:50.01274	111.0.11.111	192.168.2.1	SIP/SDP	5060	5060	959	Status: 183 Session Progress			
1070	16:11:56.795104	111.0.11.111	192.168.2.1	SIP/SDP	5060	5060	1114	Status: 200 OK			
1073	16:11:56.810811	192.168.2.1	111.0.11.111	SIP	5060	5060	597	Request: ACK sip:sgc_c@111.0.11.111;transport=udp			
2846	16:12:13.974866	192.168.2.1	111.0.11.111	SIP	5060	5060	1151	Request: BYE sip:sgc_c@111.0.11.111;transport=udp			
2853	16:12:14.049285	111.0.11.111	192.168.2.1	SIP	5060	5060	498	Status: 200 OK			

Abb. 3.18 Gesprächsaufbau in einem Festnetz IMS System

```
m=audio 7078 RTP/AVP 9 8 0 2 102 100 99 101 97 120 121
a=sendrecv
a=rtpmap:2 G726-32/8000
a=rtpmap:102 G726-32/8000
a=rtpmap:100 G726-40/8000
a=rtpmap:99 G726-24/8000
a=rtpmap:101 telephone-event/8000
a=fmtp:101 0-15
a=rtpmap:97 iLBC/8000
a=fmtp:97 mode=30
a=rtpmap:120 PCMA/16000
a=rtpmap:121 PCMU/16000
a=rtcp:7079
```

Ein interessanter Codec, der von vielen Festnetzgeräten unterstützt wird, ist G.722 für die Wideband Sprachübertragung im Festnetz. Mit diesem wird ein Sprachkanal mit einer Datenrate von 64 kbit/s übertragen. Wie auch im Mobilfunknetzwerk bietet dieser Codec eine sehr viel besseren Sprachqualität, falls die Gegenstelle diesen Codec ebenfalls unterstützt. Manche Mobilfunkbetreiber nutzen Media Gateways, um zwischen G.722 (64 kbit/s) im Festnetz und G.722.2 (WB-AMR mit 12,65–23,85 kbit/s) im Mobilfunknetzwerk zu übersetzen und ermöglichen somit Gespräche mit Wideband Sprachübertragung über die Netzwerkgrenze hinweg. G.722 findet sich nicht in den zuvor gezeigten Codec Details, da G.722 mit der Standard RTP Profilnummer 9 identifiziert wird³⁴ und somit nur durch seine Nummer in der ersten Zeile (m=) der SDP Nachricht in der Codec Liste enthalten ist.

3.7 Fragen und Aufgaben

1. Welches sind die wichtigsten IMS Komponenten und deren Aufgaben?
2. Wie wird sichergestellt, dass SIP Nachrichten nur von einem authentisierten Gerät gesendet werden können?
3. Was sind ‚Preconditions‘?
4. Was sind ‚Asserted Identities‘?
5. Warum wird bei VoLTE ‚Header Compression‘ verwendet?
6. Wie werden Gesprächsweiterleitungen in VoLTE konfiguriert?
7. Wie werden im VoLTE System Notrufe gemacht?
8. Was ist Single Radio Voice Call Continuity (SRVCC)?
9. Welches sind die wichtigsten Schritte bei der Übergabe eines Gesprächs von VoLTE nach VoWifi?
10. Was sind die Unterschiede zwischen VoWifi Cellular-Preferred und Wifi-Preferred?

Lösungen sind auf der Website zum Buch unter <http://www.cm-networks.de> zu finden.

Anmerkungen

1. The GSM Association, IR.92 IMS Profile for Voice and SMS version 10.0, <http://www.gsma.com/newsroom/all-documents/ir-92-ims-profile-for-voice-and-sms/>, Accessed in 2017.
2. Homepage des Asterisk Projekts, <https://www.asterisk.org>.
3. J. Rosenberg et al., SIP: Session Initiation Protocol, IETF RFC 3261.
4. G. Combs, Wireshark, <http://www.wireshark.org>.
5. H. Schulzrinne, RTP: A Transport Protocol for Real-Time Applications, IETF RFC 3550.
6. M. Handley, V. Jacobson and C. Perkins, SDP: Session Description Protocol, IETF RFC 4566.
7. V. Fajardo et al., Diameter Base Protocol, IETF RFC 6733.
8. 3GPP, IP Multimedia Core Network Subsystem (IMS) Multimedia Telephony Service and supplementary services; Stage 1, TS 22.173.
9. The GSM Association, IR.92 IMS Profile for Voice and SMS version 10.0, <http://www.gsma.com/newsroom/all-documents/ir-92-ims-profile-for-voice-and-sms/>, Accessed in 2017.
10. 3GPP, Signalling flows for the session setup in the IP Multimedia core network Subsystem (IMS) based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); Stage 3, TS 24.930, Chapter 5.3.1.
11. G. Camarillo et al., Integration of Resource Management and Session Initiation Protocol (SIP), RFC 3312.
12. M. Handley et al., SDP: Session Description Protocol, RFC 4566.
13. 3GPP, IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction, TS 26.114.
14. G. Camarillo et al., Early Media and Ringing Tone Generation in the Session Initiation Protocol (SIP), RFC 3960.
15. 3GPP, Common Basic Communication procedures using IP Multimedia (IM) Core Network (CN) subsystem; Protocol specification, TS 24.628.
16. 3GPP, 3G security; Access security for IP-based services, TS 33.203.
17. H. Schulzrinne et al., RTP Payload for DTMF Digits, Telephony Tones, and Telephony Signals, RFC 4733.
18. 3GPP, Support of Short Message Service (SMS) over generic 3GPP Internet Protocol (IP) access; Stage 2, TS 23.204.
19. 3GPP, IP Multimedia Subsystem (IMS) Service Continuity; Stage 2, TS 23.237.
20. 3GPP, LTE; Architectural Requirements, TS 23.221.
21. Wikipedia, IP Exchange, https://en.wikipedia.org/wiki/IP_exchange.
22. I. Tanaka, VoLTE Roaming and Interconnection Standard Technology, NTT Docomo Technical Journal, Volume 15 No. 2, 2013, https://www.nttdocomo.ne.jp/english/binary/pdf/corporate/technology/rd/technical_journal/bn/vol15_2/vol15_2_037en.pdf.
23. 3GPP, Study on S. 8 Home Routing Architecture for VoLTE, TR 23.749.

24. M. Sauter, Docomo Doesn't Want to Wait and Launches S8HR VoLTE Roaming on Its Own, <https://blog.wirelessmoves.com/2015/10/docomo-doesnt-want-to-wait-and-launches-s8hr-volte-roaming-on-its-own.html>.
25. 3GPP, Single Radio Voice Call Continuity (SRVCC); Stage 2, TS 23.216.
26. GSMA, IMS Profile for Voice, Video and Messaging over 5GS.
27. 3GPP, Procedures for the 5G System (5GS), TS 23.502, Kap. 4.13.6.1 EPS fallback for IMS voice.
28. 3GPP, Procedures for the 5G System (5GS), TS 23.502, Kap. 4.11.1.2 Handover Procedures.
29. The GSM Association, IMS Profile for Voice, Video and SMS over Wi-Fi – GSMA, IR.51.
30. 3GPP, Architecture enhancements for non-3GPP accesses, TS 23.402.
31. 3GPP, 3GPP System Architecture Evolution (SAE); Security aspects of non-3GPP accesses, TS 33.402.
32. D. Black et al., Using Authenticated Encryption Algorithms with the Encrypted Payload of the Internet Key Exchange version 2 (IKEv2) Protocol, IETF RFC 5282.
33. P. Eronen et al., An Extension for EAP-Only Authentication in IKEv2, RFC 5998.
34. H. Schulzrinne et al., RTP Profile for Audio and Video Conferences with Minimal Control, RFC 3551 Chapter 6.



Wireless LAN IEEE 802.11

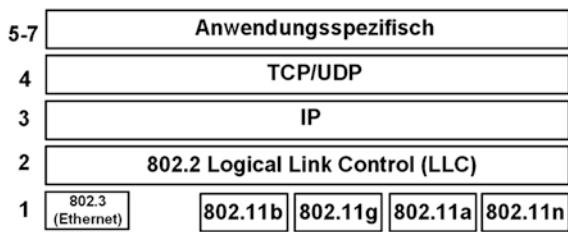
4

Mitte der neunziger Jahre fristete eine neue Technologie namens Wireless LAN noch ein Schattendasein. Dies änderte sich sehr jedoch sehr rasch, nachdem die benötigte Hardware deutlich billiger wurde. Wireless LAN wurde so schnell das optimale Medium, um Computer und später auch Smartphones und Tablets drahtlos untereinander und mit dem Internet zu verbinden. Kap. 4 dieses Buches beschäftigt sich ausführlich mit diesem System, das vom IEEE (Institute of Electrical and Electronics Engineers) unter der Bezeichnung 802.11 standardisiert wurde¹. Der erste Teil des Kapitels beschreibt zunächst die technischen Grundlagen, die sich über die Zeit nur wenig geändert haben. Im zweiten Teil werden dann die evolutionären Schritte besprochen, die über die Jahre für deutliche Geschwindigkeitssteigerungen sorgten. Außerdem werden Sicherheits- und Verschlüsselungsverfahren und optionale Funktionen, die in der Praxis jedoch durchaus verwendet werden, beschrieben.

4.1 Wireless LAN Überblick

Wireless LAN (Local Area Network) trägt seinen Namen zu Recht, denn es basiert im Wesentlichen auf LAN Standards, die ursprünglich vom IEEE für die drahtgebundene Vernetzung von Computern in den 802.X Standards beschrieben sind². Diese LAN Standards werden im täglichen Sprachgebrauch auch oft als „Ethernet“ bezeichnet. Die drahtlose Variante, also das Wireless LAN (WLAN), wurde in den 802.11 Standards spezifiziert. Wie in Abb. 4.1 zu sehen ist, dient WLAN heute hauptsächlich dazu, auf Schicht 3 des ISO Modells IP-Pakete zu transportieren. Schicht 2, der Data Link Layer, wurde mit wenigen Änderungen aus der drahtgebundenen „Ethernetwelt“ übernommen. Um der drahtlosen Natur des Netzwerkes Rechnung zu tragen, wurden zusätzlich für Layer 2 einige Management-Operationen definiert, die in Abschn. 4.2 beschrieben

Abb. 4.1 WLAN Protokollstack



werden. Lediglich Schicht 1, der Physical Layer, wurde komplett neu entwickelt, da bei WLAN kein Kabel, sondern Funkwellen für die Übertragung der Datenpakete verwendet werden.

4.2 Geschwindigkeiten und Standards

Seit Bestehen der 802.11 Standards gab es zahlreiche Weiterentwicklungen bei der Funkübertragung. Aus diesem Grund gibt es mehrere Physical Layer, die in den Spezifikationen abgekürzt PHY genannt werden. Neue PHYs und andere Erweiterungen sind weiterhin Teil der 802.11 Spezifikationen und erhalten jeweils eine eigene Buchstabenkombination hinter der allgemeinen Bezeichnung 802.11. Die folgende Tabelle gibt einen Überblick über aktuelle spezifizierte PHYs und deren Namen.

Standard	Wi-Fi Alliance Bezeichnung	Frequenzband	Theoretische Maximalgeschwindigkeit
802.11b ³		2,4 GHz (2,401–2,483 GHz)	1–11 Mbit/s
802.11g ⁴		2,4 GHz (2,401–2,483 GHz)	6–54 Mbit/s
802.11a ⁵		5 GHz (5,150–5,350 GHz und 5,470–5,725 GHz)	6–54 Mbit/s
802.11n	Wi-Fi 4	2,4 GHz (wie oben) 5 GHz (wie oben)	6–600 Mbit/s
802.11ac	Wi-Fi 5	5 GHz (wie oben)	bis zu 6,93 Gbit/s
802.11ax	Wi-Fi 6	2,4 GHz und 5 GHz (+6 GHz mit Wi-Fi 6E)	bis zu 9,5 Gbit/s

Der große Durchbruch für WLAN erfolgte mit dem 802.11b Standard, mit dem Datenraten von 1–11 MBit/s möglich sind. Die Übertragungsrate richtete sich dabei hauptsächlich nach der Entfernung zwischen Sender und Empfänger, sowie nach der Anzahl der Hindernisse wie Wände oder Decken. 11 MBit/s waren dabei in Gebäuden nur über kurze Entferungen in Größenordnungen von 10–20 m möglich. Die Redundanz in den Datenpaketen wurde je nach Übertragungsqualität automatisch angepasst und reduzierte so die Geschwindigkeit bei sehr schlechten Bedingungen auf bis zu 1 MBit/s. Der

802.11b Standard sendete im 2,4 GHz ISM (Industrial, Scientific and Medical)-Band, das in den meisten Ländern lizenzenfrei verwendet werden darf. Wichtigste Bedingung für die Verwendung dieses Bandes ist die Beschränkung der maximalen Sendeleistung auf 100 mW. Das ISM-Band ist ein öffentliches Frequenzband, neben WLAN senden hier auch noch andere Funksysteme wie z. B. Bluetooth.

Im 802.11g Standard wurde ein im Vergleich zum 802.11b Standard weit komplexerer PHY spezifiziert, der Datenraten je nach Qualität des Übertragungsmediums von bis zu 54 MBit/s erlaubt. Auch dieser Standard sendete auf dem 2,4 GHz ISM-Band und wurde so gestaltet, dass das Verfahren rückwärtskompatibel zu 802.11b war. Somit wurde sichergestellt, dass 802.11b-Geräte auch mit 802.11g-Geräten kommunizieren konnten.

Zusätzlich zum 2,4 GHz ISM-Band wurde auch im 5 GHz Frequenzbereich ein Band für WLAN freigegeben, für das zunächst der 802.11a Standard spezifiziert wurde. Wie beim 802.11g Standard waren auch hier Datenraten von 6–54 Mbit/s möglich. 802.11a-Endgeräte wurden jedoch aufgrund der nötigen Rückwärtskompatibilität zu 802.11b/g nie besonders erfolgreich, da die Unterstützung von zwei Frequenzbereichen damals deutlich höhere Kosten verursachte. In der Praxis standen dem jedoch früher keine nennenswerten Vorteile gegenüber.

Aufgrund der steigenden Datenraten bei lokalen Netzwerken und auch bei Internetzugängen per Kabel oder ADSL wurde bald klar, dass auch bei Wireless LANs weitere Geschwindigkeitssteigerungen notwendig waren. Nach einigen Jahren Standardisierungsarbeit einigten sich schließlich die beteiligten Firmen auf ein gemeinsames neues Verfahren, das im IEEE 802.11n Standard definiert ist. Durch doppelte Kanalbreiten und zahlreichen weiteren Neuerungen, die im Laufe dieses Kapitels näher beschrieben werden, erreicht dieser Standard theoretische Spitzengeschwindigkeiten von bis zu 600 Mbit/s auf Layer 1. In der Praxis können mit diesem Standard typische Datenraten auf Layer 3 (IP) von 70 bis 150 Mbit/s unter guten Übertragungsbedingungen erreicht werden. Außerdem unterstützt der Standard sowohl das 2,4 GHz Band als auch das 5 GHz Band. Dies wurde notwendig, da das 2,4 GHz Band bereits sehr stark genutzt wurde und besonders in Städten viele Netze gleichzeitig einen Kanal nutzen. Das 5 GHz Band wird dagegen aktuell noch wenig verwendet und bietet deshalb in vielen Situationen besser Übertragungsbedingungen, die für Anwendungen mit großem Bandbreitenbedarf wie z. B. Video Streaming notwendig sind.

Der nächste Schritt in der WLAN Evolution war 802.11ac. Dieser wird oft auch als Wi-Fi 5 bezeichnet, ein Begriff, der durch die Wi-Fi Alliance eingeführt wurde. Durch Kanalbandbreiten von 80 MHz im 5 GHz Band, und optional auch 160 MHz, verbesserter Modulation und weiteren Verfahren zur Geschwindigkeitssteigerung, ist eine theoretische Spitzendatenrate von 6,9 Gbit/s möglich. In der Praxis sind die erreichbaren Geschwindigkeiten jedoch wesentlich geringer, gehen jedoch noch immer weit über die Möglichkeiten von 802.11n Endgeräten hinaus. 802.11ac kompatible Geräte erreichen heute unter idealen Bedingungen und auf sehr kurzen Distanzen Übertragungsgeschwindigkeiten von bis zu 700 Mbit/s auf Layer 3 (IP). Dies ist in der Praxis

jedoch eher eine Ausnahme, unter normalen Bedingungen sind jedoch noch immer Geschwindigkeiten von 200 bis 400 Mbit/s möglich.

Während 802.11ac Endgeräte heute große Verbreitung haben, gibt es auch zunehmend Geräte, die auch den aktuell neuesten WLAN Standard, 802.11ax, unterstützen. Dieser wird von der Wi-Fi Alliance auch als Wi-Fi 6 bezeichnet. Hier können zum ersten Mal in größerem Umfang Kanäle im 5 GHz Band mit 160 MHz Bandbreite verwendet werden, was somit die Datenrate gegenüber den bei 802.11ac üblichen 80 MHz Kanälen weiter steigert. Außerdem führt Wi-Fi 6 auch neue Verfahren für die gleichzeitige Datenübertragung mit mehreren Endgeräten ein. Dadurch steigert sich zwar die Spitzendatenrate eines einzelnen Endgeräts nicht, ein Kanal kann dadurch jedoch an Orten mit vielen Geräten, wie z. B. in Büros, besser genutzt werden.

Zusätzlich zu den Frequenzbereichen bei 2,4 GHz und 5 GHz wurde von der Bundesnetzagentur und anderen nationalen Regulierungsbehörden in 2021 auch ein neuer, etwa 500 MHz breiter Bereich im 6 GHz Band für die Nutzung von WLAN und anderen Funktechnologien freigegeben. Dies ist wichtig, da heute auch das 5 GHz Band zunehmend intensiver genutzt wird, und sich dies natürlich negativ auf die Datenraten in Netzwerken vor allem in dicht besiedelten Gebieten, auswirkt. WLAN Geräte, die dieses Frequenzband unterstützen, werden auch als Wi-Fi 6E Geräte vermarktet, da für die Unterstützung des 6 GHz Frequenzbandes der 802.11ax Standard entsprechend erweitert wurde.

Weitere 802.11 Standards, die in der nachfolgenden Tabelle aufgelistet sind, spezifizieren diverse zusätzliche optionale Wireless LAN-Funktionalitäten:

Standard	Beschreibung
802.11e ⁶	Wichtigste neue Funktionalität des Standards sind Methoden, um für eine Übertragung eine bestimmte Quality of Service (QoS) zu gewährleisten. Damit ist es möglich, Bandbreite und schnellen Medienzugriff für Echtzeitanwendungen wie z. B. Voice over IP (VoIP) auch in stark ausgelasteten Netzen zu gewährleisten. Außerdem spezifiziert der Standard das Direct Link Protocol (DLP), mit dem zwei WLAN-Endgeräte auch direkt unter Umgehung des Access Points Daten austauschen können. Dies steigert die Übertragungsgeschwindigkeit zwischen zwei drahtlosen Endgeräten wesentlich
802.1f ⁷	Spezifikation für den Datenaustausch zwischen Access Points. Mehr dazu in Abschn. 4.3.1 über Extended Service Sets (ESS)
802.11h ⁸	Ergänzung für Standards im 5 GHz-Bereich für Leistungsregelung und dynamische Frequenzwahl. In Europa sind ab einer gewissen Sendeleistung nur 802.11a-Systeme zugelassen, die sich an diese Erweiterungen halten
802.11i ⁹	Standardisiert erweiterte Authentifizierungs- und Verschlüsselungsalgorithmen für WLAN. Wichtiger Bestandteil von 802.11i ist 802.1x. Mehr hierzu in Abschn. 4.7 zum Thema WLAN-Sicherheit
802.11w	Einführung von Management Frame Sicherungsmechanismen gegen De-Authentication/Disassociation Angriffe. Neue, zu 802.11ac kompatible Geräte, müssen diese Erweiterung unterstützen

Standard	Beschreibung
802.11k	Network Assisted Roaming: In WLAN Netzwerken mit mehreren Access Points ermöglicht diese Erweiterung, Informationen über benachbarte APs des gleichen Netzwerkes den Endgeräten mitzuteilen. Endgeräte können diese Information dann verwenden, um z. B. bei einem schlechten Signalpegel zu einem anderen AP zu wechseln
802.11v	Network Assisted Roaming: Falls vom AP und Endgerät unterstützt, kann ein Endgerät aufgefordert werden, zu einem anderen AP zu wechseln. Dies kann für Funktionen wie Load Balancing verwendet werden, oder auch, um ein Endgerät zu einem anderen AP zu schicken, falls der Empfangspegel für das Endgerät zu niedrig wird

Da es heute eine große Anzahl an Firmen gibt, die WLAN Endgeräte herstellen, ist die Interoperabilität zwischen unterschiedlichen Endgeräten von entscheidender Bedeutung. Um dies sicherzustellen, wurde 1999 die schon erwähnte Wi-Fi Alliance gegründet. Heute sind hunderte Firmen Mitglied der Wi-Fi Alliance und nutzen das Wi-Fi Alliance Zertifizierungsprogramm, um ihre Produkte zu testen und das ‚Wi-Fi certified‘ Marketing Logo für ihr Produkt zu bekommen. Da dieses Label heute auf den meisten Verpackungen von WLAN fähigen Produkten zu finden ist, wird 802.11 WLAN auch oft als Wi-Fi bezeichnet.

4.3 WLAN-Konfigurationen: Von Ad-hoc bis Wireless Bridging

Alle Stationen, die auf dem gleichen Übertragungskanal Daten austauschen, werden im 802.11 Standard unter dem Begriff Basic Service Set (BSS) zusammengefasst. Die Definition des BSS umfasst auch den geografischen Bereich, in dem sich die Teilnehmer des BSS aufhalten können. Ein BSS kann in folgenden unterschiedlichen Modi betrieben werden:

4.3.1 Ad-hoc, BSS, ESS und Wireless Bridging

Im Ad-hoc Mode, auch Independent BSS (IBSS) genannt, kommunizieren zwei oder mehr WLAN-Endgeräte direkt miteinander. Jede Station ist gleichberechtigt, und Daten werden direkt von Endgerät zu Endgerät gesendet. Der Ad-hoc Mode entspricht also im Wesentlichen einem drahtgebundenen Ethernet, in dem ebenfalls alle Stationen gleichberechtigt sind und Datenpakete ebenfalls direkt zwischen zwei Teilnehmern ausgetauscht werden. Die gesendeten Daten werden zwar auch von allen anderen Teilnehmern des Netzwerks empfangen, von diesen aber ignoriert, weil die Zieladresse des Pakets nicht mit ihrer eigenen Adresse übereinstimmt. Alle Teilnehmer des Ad-hoc-Netzes müssen sich zu Beginn auf die Werte für einige Parameter einigen und diese dann

in ihren Endgeräten entsprechend konfigurieren. Wichtigster Parameter ist die Service Set ID (SSID), die als Namen für das Netzwerk dient. Weiterhin müssen alle Teilnehmer des Netzwerkes die gleiche Kanalnummer einstellen und auch der Verschlüsselungskey muss bei allen Teilnehmern gleich konfiguriert werden. Zwar kann das Ad-hoc-Netzwerk auch ohne Verschlüsselung betrieben werden, aus Sicherheitsgründen ist hiervon jedoch abzuraten. Schließlich müssen sich die Teilnehmer noch auf die zu verwendenden IP-Adressen einigen und auch diese entsprechend in ihren Endgeräten eintragen. Die komplizierte Konfiguration ist einer der Gründe, warum der Ad-hoc-Modus in der Praxis selten verwendet wird.

Eine der Hauptanwendungen eines WLAN-Netzwerkes ist der Zugang zu einem Firmen- oder Heimnetzwerk, sowie dem Internet. Für diesen Zweck eignet sich der Infrastructure BSS Mode des WLAN Standards am besten. Im Unterschied zum Ad-hoc Mode gibt es hier einen sogenannten Access Point, der eine zentrale Rolle übernimmt.

Der Access Point bildet, wie in Abb. 4.2 gezeigt, den Übergang zwischen dem drahtlosen und drahtgebundenen Netzwerk für alle Endgeräte im BSS. Außerdem kommunizieren Endgeräte in einem Infrastructure BSS nicht direkt miteinander, sondern immer über den Access Point. Möchte Endgerät A an Endgerät B ein Datenpaket schicken, sendet es dies zunächst an den Access Point. Der Access Point analysiert die Zieladresse und stellt das Paket danach an Teilnehmer B zu. Auf diese Weise ist es möglich, Endgeräte im drahtlosen und im drahtgebundenen Netzwerk zu erreichen, ohne dass Teilnehmer wissen müssen, um welche Sorte Endgerät es sich handelt. Der zweite Vorteil dieses Verfahrens liegt darin, dass auch drahtlose Endgeräte über den Access Point miteinander kommunizieren können, die sich für eine direkte Kommunikation zu weit auseinander befinden. Dies kann z. B. der Fall sein, wenn sich, wie in Abb. 4.2 gezeigt, der Access Point zwischen Endgerät A und B befindet. Die Sendeleistung jedes Endgeräts reicht zwar aus, den Access Point zu erreichen, nicht jedoch das jeweils andere Gerät. Großer Nachteil dieses Verfahrens ist jedoch, dass bei Kommunikation zwischen zwei drahtlosen Teilnehmern das Datenpaket zweimal über die Luftschnittstelle übertragen wird und somit die maximale Bandbreite des BSS halbiert wird. Aus diesem Grund wurde als Teil des 802.11e Standards das optionale Direct Link Protocol (DLP)

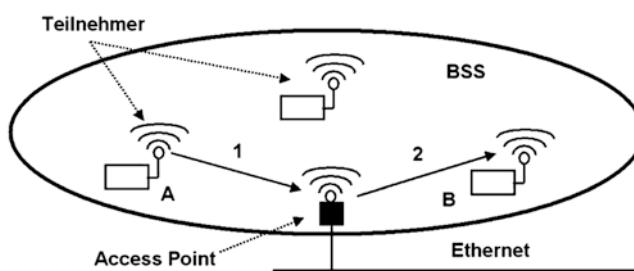


Abb. 4.2 Infrastructure BSS

eingeführt, das eine direkte Kommunikation zwischen zwei Endgeräten erlaubt. In der Praxis findet diese Erweiterung jedoch nur wenig Anwendung. Weitere Informationen hierzu in Abschn. 4.8.

Oft ist ein WLAN Access Point mit weiteren Funktionen ausgestattet:

- 1 Gbit/s-Anschlüsse und High-End APs auch mit 2,5 Gbit/s Anschlüsse für drahtgebundene Ethernet-Endgeräte mit Layer 2 Switching-Funktionalität.
- Oft dient ein WLAN Access Point im Heimbereich auch gleichzeitig als IP Router zum Internet und kann per Ethernet mit einem DSL oder Kabelmodem verbunden werden.
- Eine ebenfalls übliche Variante ist, dass der WLAN Access Point Teil eines DSL-oder Kabelrouters ist. Dies ist sehr praktisch, da weniger Geräte verkabelt werden müssen und nur noch ein Netzteil für die Stromversorgung benötigt wird. Ein solcher voll integrierter Access Point ist in Abb. 4.3 gezeigt.
- Um Endgeräte automatisch für das Netzwerk zu konfigurieren, ist üblicherweise auch ein DHCP (Dynamic Host Configuration Protocol) Server in einem Access Point integriert¹⁰. Dieser über gibt allen drahtlosen und drahtgebundenen Endgeräten die benötigten Netzwerkeinstellungen wie individuelle IP-Adressen, sowie die Adresse des DNS Servers für die Namensauflösung und die IP-Adresse des Internet Gateways.

Da ein WLAN Access Point (AP) aufgrund seiner geringen Sendeleistung nur eine begrenzte Reichweite hat, sind in manchen Fällen mehrere APs notwendig, um ein bestimmtes Gebiet zu versorgen. Ändert ein mobiler Teilnehmer seinen Aufenthaltsort und kann dadurch von einem anderen AP besser versorgt werden, meldet sich die Netzwerkkarte automatisch beim neuen AP an. Eine solche Anordnung wird Extended Service Set (ESS) genannt und ist in Abb. 4.4 dargestellt. Meldet sich ein Endgerät an

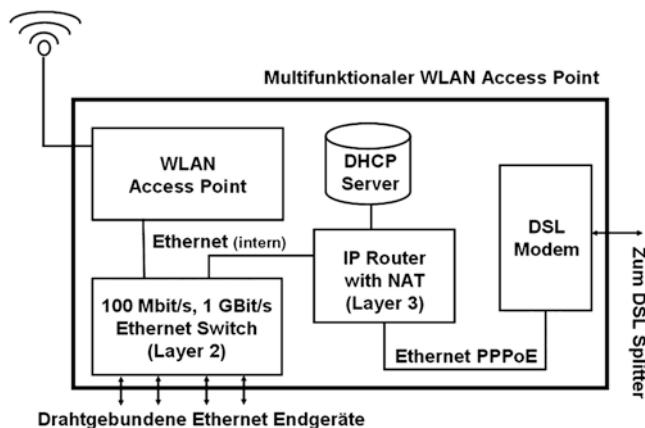


Abb. 4.3 Access Point, IP-Router und DSL-Modem in einem Gerät

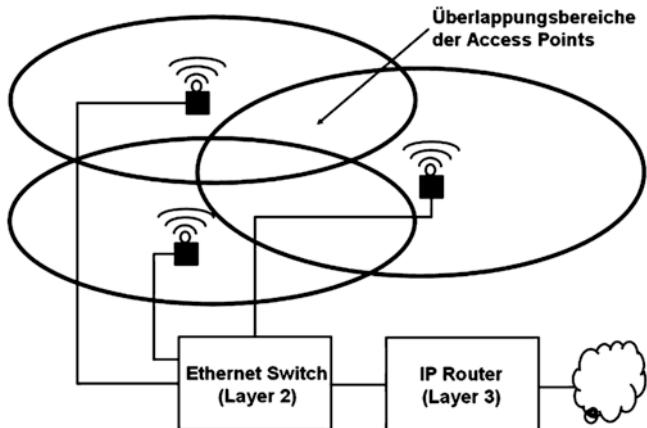


Abb. 4.4 Extended Service Set (ESS) mit 3 Access Points

einem anderen Access Point des ESS an, tauschen der neue und bisherige AP über die Ethernet Verbindung, die in den WLAN Standards auch Distribution System genannt wird, Teilnehmerinformationen aus. Zukünftig werden dann Pakete, die über das Distribution System für den Teilnehmer eingehen, über den neuen AP an den Teilnehmer zugestellt, der alte Access Point ignoriert fortan diese Pakete. Für höhere Schichten des Protokollstacks ist der Wechsel des Access Points in einem ESS nicht sichtbar, die IP-Adresse kann deshalb beibehalten werden.

Folgende Bedingungen müssen erfüllt sein, um den reibungslosen Übergang eines Teilnehmers zu einem anderen Access Point (AP) in einem ESS zu gewährleisten:

- Alle APs eines ESS müssen sich im gleichen IP-Subnetz befinden, es dürfen also keine IP Router zwischen den APs liegen. Ethernetswitches, die auf ISO Layer 2 arbeiten, sind jedoch erlaubt. Dies limitiert das Ausbreitungsgebiet eines ESS beträchtlich, da IP-Subnetze oft nicht sehr groß sind (z. B. ein Gebäude oder ein Stockwerk).
- Alle APs müssen die gleiche BSS Service ID (SSID) besitzen. Mehr zur SSID in Abschn. 4.3.2.
- APs müssen auf unterschiedlichen Frequenzen senden und sich bei der Frequenzwahl an ein Muster halten, das in Abb. 4.5 gezeigt wird.
- Viele APs verwenden für den Austausch der Teilnehmerinformationen bei einem AP-Wechsel ein proprietäres Protokoll. Aus diesem Grund sollten alle APs eines ESS vom gleichen Hersteller stammen. Um ein ESS mit APs unterschiedlicher Hersteller zu ermöglichen, wurde vom IEEE Anfang 2003 der Standard 802.11f (Recommended Practice for Multi-Vendor Access Point Interoperability) verabschiedet, der aber nicht verpflichtend für Hersteller ist.

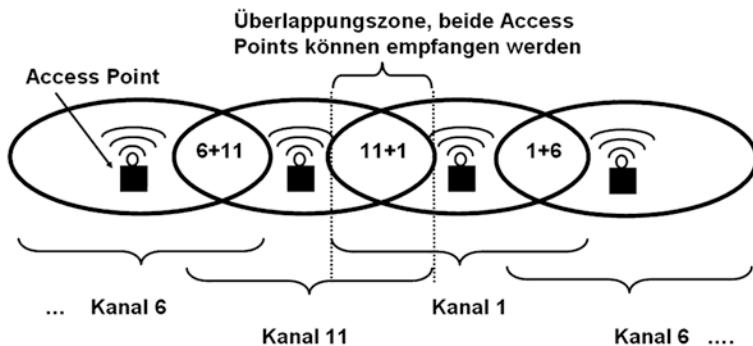


Abb. 4.5 Überlappende Abdeckung von Access Points

- Zwischen den Abdeckungsbereichen der einzelnen Access Points muss es eine Überlappung geben, damit Endgeräte auch in den Randgebieten die Netzabdeckung nicht verlieren. Da die APs mit unterschiedlichen Frequenzen senden, stellt diese Überlappung aber kein Problem dar.

Eine weitere WLAN-Variante ist das Wireless Bridging. In dieser Betriebsart wird das drahtgebundene Ethernet Distribution System zwischen zwei oder mehr APs eines ESS durch eine Funkstrecke ersetzt. In der Praxis gibt es heute eine Vielzahl von Wireless Hotspots, die auch als Wireless Bridge betrieben werden können.

4.3.2 SSID und Frequenzwahl

Bei Inbetriebnahme eines Access Points gibt es zwei grundsätzliche Parameter, die individuell vergeben werden müssen:

Der erste Parameter ist die Basic Service Set ID, kurz SSID genannt. Die SSID wird vom Access Point über Beacon Frames, die in Abschn. 4.4 besprochen werden, in regelmäßigen Abständen über die Luftschnittstelle bekannt gegeben (Broadcast). Das Wort „Frame“ wird bei WLAN synonym zu „Paket“ verwendet. Die SSID identifiziert ein Netzwerk eindeutig und ermöglicht es, mehrere unterschiedliche Access Points, die Zugriff auf unterschiedliche Netzwerke gewähren, am gleichen Ort zu betreiben. Eine Konfiguration von unabhängigen APs sollte nicht mit einem ESS verwechselt werden, das für alle APs die gleiche SSID verwendet. Üblicherweise wird für die SSID ein Textstring gewählt, da dieser bei der Konfiguration der Endgeräte später in einer Dialogbox dem User zur Auswahl des Access Points angeboten wird.

Zweiter wichtiger Parameter, der bei Vorhandensein mehrerer APs sorgfältig gewählt werden sollte, ist die Sendefrequenz. Das ISM-Band im 2,4 GHz-Bereich von 2,410 bis 2,483 MHz ist je nach Land in bis zu 13 Kanäle von jeweils 5 MHz Bandbreite

unterteilt. Da ein WLAN-Kanal eine Bandbreite von 25 MHz benötigt, sollten unterschiedliche WLAN-Netze, die sich überlappen oder die gleiche Fläche abdecken, mindestens 5 ISM Kanäle Abstand zueinander halten. Wie in Abb. 4.5 dargestellt, können auf diese Weise 3 unabhängige BSS oder ein ESS mit sich überlappenden Grenzen von 3 Access Points betrieben werden. Bei 3 unabhängigen BSS ist diese Überlappung nicht unbedingt gewünscht, lässt sich in der Praxis jedoch oft nicht vermeiden. Bei 3 Access Points, die zusammen ein ESS bilden, ist diese Überlappung jedoch notwendig, um einen nahtlosen Wechsel von einem AP zum anderen zu ermöglichen. Um mindestens 5 Kanäle Abstand einzuhalten, müssen in den Access Points jeweils Kanal 1, 6 und 11 eingestellt werden.

Da die Kanäle 12 und 13 nur in Europa zugelassen sind, wird bei der Installation der meisten WLAN Geräte das Land abgefragt. Manche Produkte sparen sich jedoch diese Abfrage und blockieren Kanal 12 und 13 permanent. Steht deshalb beim Aufbau eines Access Points nicht fest, mit welchen Endgeräten später auf das Netzwerk zugegriffen wird, sollten Kanal 12 und 13 nicht verwendet werden.

802.11a, 11n, 11ac und 11ax Systeme senden im 5 GHz-Bereich in Europa von 5,150–5,350 GHz und von 5,470–5,725 GHz. Zusammen sind dies 455 MHz, in denen 18 unabhängige 802.11a WLAN-Netzwerke Platz finden. Gegenüber den 3 unabhängigen Netzen im 2,4 GHz Band ist dies eine enorme Verbesserung. Mit 802.11n wurde die Kanalbreite dann auf 40 MHz verdoppelt, es sind also noch 9 unabhängige Netze möglich. 802.11ac verwendet typischerweise 80 MHz Kanäle, 802.11ax ermöglicht sogar eine Bündelung von 160 MHz in der Praxis. Dies schränkt die Anzahl der unabhängigen Netze im 5 GHz Frequenzbereich natürlich wieder stark ein. Da für diesen Frequenzbereich eine automatische Frequenzwahl vorgeschrieben ist, suchen sich Access Points automatisch einen freien Kanal.

Auf Endgeräteseite ist die Grundkonfiguration des Wireless LANs für ein BSS und ESS einfacher. Das Endgerät sucht bei der Konfiguration alle Frequenzen nach vorhandenen Access Points ab und zeigt dann die gefundenen SSIDs an. Der Benutzer hat daraufhin die Möglichkeit, eine SSID auszuwählen. Dies ist in Abb. 4.6 gezeigt. Der Sendekanal hingegen muss nicht ausgewählt werden, da das Endgerät beim Einschalten immer alle Kanäle nach einem Access Point mit der ausgewählten SSID durchsucht. Werden mehrere Access Points auf unterschiedlichen Frequenzen mit der gleichen SSID gefunden, handelt es sich um ein ESS. Das Endgerät wählt dann den Kanal, auf dem die Beacon Frames (vgl. nächster Abschnitt) am besten empfangen werden.

Neben SSID und Sendekanal ist die Konfiguration der Verschlüsselung für Heim- und Firmennetzwerke ebenfalls sehr wichtig. Manche Produkte haben diese noch immer standardmäßig bei Auslieferung deaktiviert. Dies stellt ein großes Sicherheitsrisiko dar, da Funkwellen nicht an der Wohnungs- oder Bürotür hält machen. Mehr zu diesem wichtigen Thema in Abschn. 4.7.



Abb. 4.6 Endgerätekonfiguration für ein BSS oder ESS

4.4 Management-Operationen

Im drahtgebundenen Ethernet genügt es, ein Endgerät mit einem Kabel am nächsten Hub oder Switch anzuschließen, um dem Endgerät Zugriff auf das Netzwerk zu gewähren. Ein solches physisches „Einstecken“ ist bei einem WLAN-Endgerät nicht möglich. Zusätzlich verfügt ein WLAN-Endgerät über Funktionen wie automatisches Roaming zu anderen Access Points eines ESS, oder die Verschlüsselung der Datenpakete auf Layer 2, die mit dem Netzwerk koordiniert werden müssen. Aus diesem Grund definiert der 802.11 Standard eine Reihe von Management-Operationen und Nachrichten auf Layer 2, sowie zusätzliche Informationen im MAC Header von Datenpaketen, die im drahtgebundenen Ethernet nicht notwendig sind.

In einem BSS nimmt der Access Point (AP) eine zentrale Rolle ein und stellt gleichzeitig den Übergang zum drahtgebundenen Ethernet her. Alle Datenpakete im WLAN werden immer an den AP geschickt, der dann die Weiterleitung an mobile und drahtgebundene Endgeräte übernimmt. Damit ein WLAN-Endgerät beim Einschalten einen aktiven AP erkennen kann, sendet dieser in regelmäßigen Abständen (typisch sind 100 ms) Beacon Frames aus. Wie in Abb. 4.7 auszugsweise gezeigt, enthalten Beacon Frames neben der SSID des Access Points noch eine Menge weiterer Informationen, die einem Endgerät Aufschluss über Funktionen und Optionen des Access Points liefern.

Jedes Bit des 2 Byte langen Capability Information Element (Capability IE) gibt Auskunft über eine bestimmte Eigenschaft. So ist zum Beispiel in Abb. 4.7 zu sehen, dass der Access Point keine Verschlüsselung aktiviert hat (Privacy Disabled). Für umfangreichere Informationen wie z. B. die unterstützten Übertragungsraten, die mehr als ein Bit benötigen, werden eigene Information Elements (IE) im Beacon Frame verwendet. Jedes Information Element hat seine eigene ID wie z. B. 0 für das IE, das die SSID enthält, oder 1 für das IE „Supported (Data-) Rates“. Da IEs unter Umständen variable Längen haben (z. B. das SSID IE), folgt auf die ID eine Längenangabe. Somit ist es für das Endgerät möglich, optionale und evtl. unbekannte IEs, die Information für neuere Geräte enthalten, bei der Dekodierung der Nachricht zu überspringen.

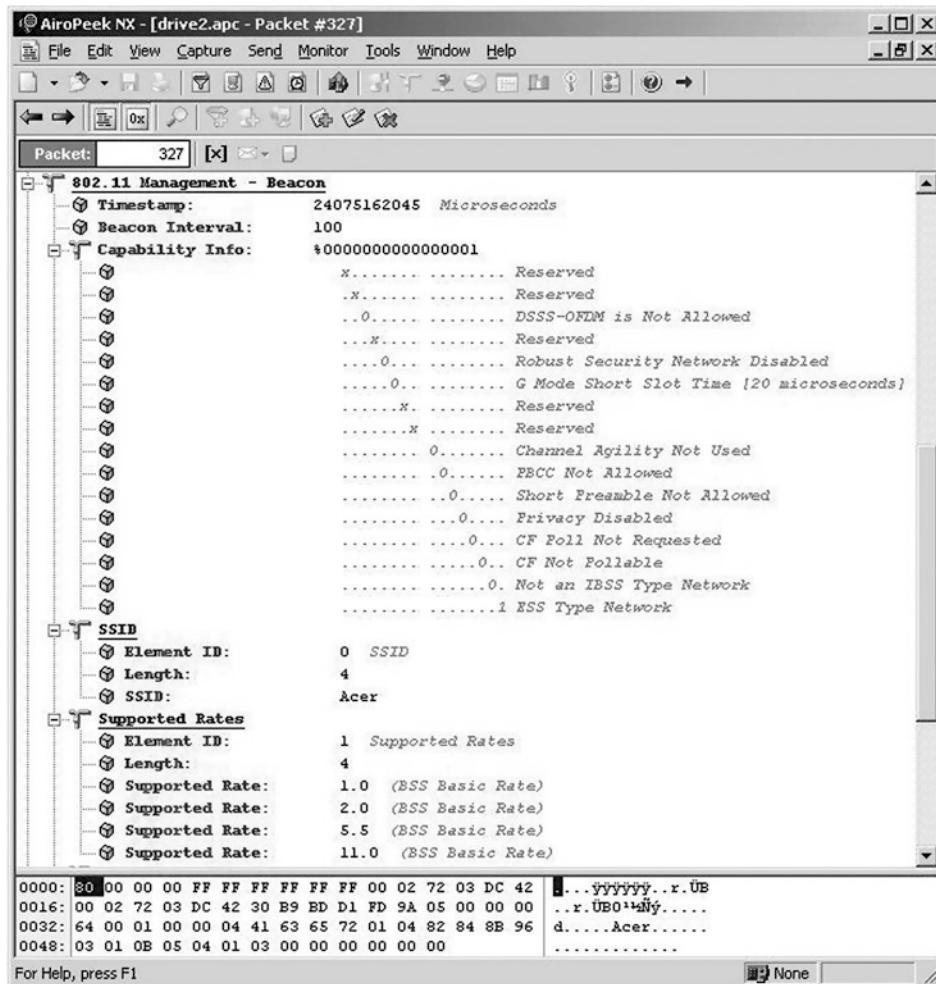


Abb. 4.7 Ausschnitt aus einem Beacon Frame

Ein Endgerät hat die Möglichkeit, bei der Netzsuche entweder nur passiv alle Kanäle nach Beacon Frames zu durchsuchen, oder aktiv mit Probe Request Frames einen Access Point zu suchen. In der Praxis verwenden die meisten Endgeräte beide Methoden. Empfängt ein Access Point einen Probe Request Frame, antwortet er mit einem Probe Response Frame, der die gleichen Informationen wie ein Beacon Frame enthält.

Nachdem ein Endgerät einen geeigneten Access Point gefunden hat, folgt im nächsten Schritt die sogenannte Authentifizierung. Der Standard definiert dazu zwei Verfahren:

Die Open System Authentication, die heute üblicherweise verwendet wird, trägt ihren Namen zu Unrecht, denn bei diesem Verfahren findet keine Authentifizierung des Endgerätes im eigentlichen Sinne statt. Das Endgerät sendet in diesem Schritt einen Authentication Frame mit einer Authentifizierungsanforderung an den Access Point (Authentication Request), der als Authentifizierungsalgorithmus „Open System“ fordert. Weitere Authentifizierungsinformationen sind nicht nötig. Lässt der Access Point eine solche „Authentifizierung“ zu, antwortet er mit einem positiven Statuscode, und das Endgerät ist „authentifiziert“.

Die Shared Key Authentication ist das zweite Authentifizierungsverfahren, wird aber heute nicht mehr verwendet, da die eigentliche Authentifizierung erst während der WPA Authentifizierungs- und Verschlüsselungsphase gemacht wird, die erst später folgt.

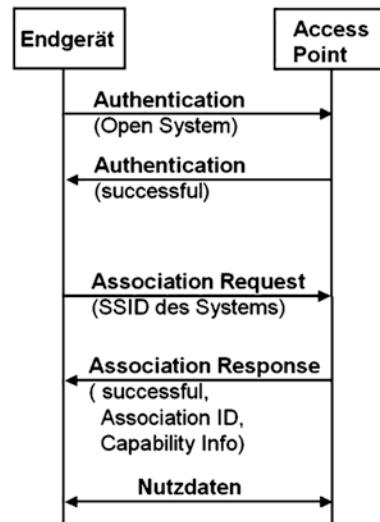
Im nächsten Schritt sendet ein Endgerät einen Association Request (Zuordnungsanforderung) an den Access Point. Der Access Point antwortet daraufhin mit einer positiven Association Response-Nachricht, in der die wichtigsten Systeminformationen wie das Capability Information Element noch einmal wiederholt werden. Außerdem wird dem Endgerät eine Association ID übergeben, die später für den Power Save Mode benötigt wird. Eine Trennung zwischen Authentication und Association wurde eingeführt, um einem Endgerät den schnellen Wechsel zwischen Access Points des gleichen ESS zu ermöglichen.

Abb. 4.8 zeigt die zwei für die Verbindungsaufnahme mit dem Netzwerk nötigen Prozeduren Authentication und Association. Acknowledgement Frames, die in Abschn. 4.5 eingeführt werden, wurden zur besseren Übersicht weggelassen.

Nach erfolgreicher Association des Endgeräts mit einem Access Point können bei offenen Netzwerken sofort Nutzdatenpakete übertragen werden. Bei verschlüsselten Netzwerken muss zuvor noch die WPA oder WPA-2 Authentifizierungs- und Schlüsselaustauschfunktion durchlaufen werden, die in Abschn. 4.7 beschrieben wird.

Befindet sich ein Endgerät in einem ESS mit mehreren Access Points (vgl. Abb. 4.4), kann es jederzeit zu einem anderen Access Point mit besserem Empfang für den aktuellen Aufenthaltsort wechseln. Dies wird als Wi-Fi Roaming bezeichnet. Um andere APs eines ESS zu finden, kann ein Endgerät, wenn keine Daten übertragen werden, alle Kanäle nach Beacon Frames von anderen APs absuchen. Da alle APs des gleichen ESS die gleiche SSID in den Beacon Frames verwenden, kann ein Endgerät sehr einfach zwischen APs des eigenen Netzwerkes von APs anderer Netzwerke unterscheiden. Außerdem wurden weitere Verbesserungen spezifiziert, die diese Prozeduren beschleunigen und in der Praxis auch eingesetzt werden. Wenn AP und Endgerät die

Abb. 4.8 Authentication und Association (ohne Acknowledgment Frames)



802.11k Neighbor Reporting Erweiterung unterstützen, kann der AP die Endgeräte über andere APs im ESS informieren. Dies wird jedoch bisher nur von neuen und teuren Endgeräten unterstützt. Mit der 802.11v BSS Transition Erweiterung kann ein AP die Endgeräte auffordern, zu einem anderen AP des gleichen ESS zu wechseln. Dies ermöglicht eine Lastverteilung im Netzwerk oder den schnellen Transfer eines Endgeräts zu einem anderen AP bei schwächer werdendem Signalpegel. 802.11v wird von den meisten Endgeräten unterstützt, die in den letzten Jahren verkauft wurden. Im Unterschied zu Mobilfunknetzwerken können Endgeräte in Wi-Fi Netzwerken jedoch selbstständig entscheiden, wann sie zu einem anderen AP des gleichen Netzwerks wechseln möchten. Die dazugehörige Prozedur wird Reassociation genannt und ist in Abb. 4.9 dargestellt. Um zu

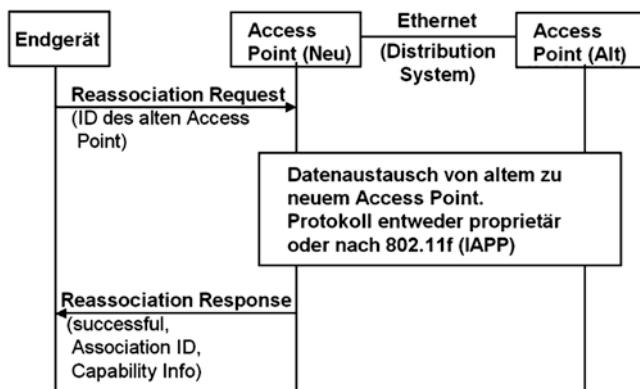


Abb. 4.9 Reassociation (ohne Acknowledgement Frames)

einem neuen Access Point zu wechseln, ändert das Endgerät die Sende- und Empfangsfrequenz und sendet auf der neuen Frequenz einen Reassociation Request Frame. Dieser entspricht im Wesentlichen einem Association Request Frame mit der Ausnahme, dass zusätzlich noch die ID des vorherigen Access Points übergeben wird. Der neue Access Point sucht daraufhin über das drahtgebundene Ethernet (Distribution System) mit der übergebenen ID den bisherigen Access Point des Teilnehmers und informiert diesen über den Wechsel. Der bisherige Access Point sendet dem neuen Access Point dann eventuell zwischengepufferte Datenpakete des Endgeräts und löscht dessen Hardwareadresse und Association ID dann aus seiner Teilnehmerliste. Zukünftig eingehende Datenpakete über das drahtgebundene Distribution System, die immer von allen APs eines ESS empfangen werden, werden fortan nur vom neuen AP zum Teilnehmer übertragen und vom bisherigen AP ignoriert. Abgeschlossen wird die Reassociation-Prozedur durch Senden einer positiven Reassociation Response-Nachricht an das Endgerät.

Anfangs war nur die Signalisierung zwischen Endgerät und neuem Access Point der Reassociation-Prozedur standardisiert. Für die drahtgebundene Kommunikation zwischen den Access Points gab es lange Zeit keinen Standard, sodass diese Prozedur von den Herstellern mit proprietären Protokollen gelöst wurde. Mit Verabschiedung der 802.11f-Empfehlung und des Inter Access Point Protocol (IAPP) können nun auch Access Points unterschiedlicher Hersteller miteinander kooperieren. Es bleibt jedoch den Herstellern überlassen, ob sie dieses Protokoll in ihren Access Points implementieren.

Um die Laufzeit batteriebetriebener Geräte zu erhöhen, gibt es in den 802.11 Standards auch einen Stromsparmodus (Power-Saving Mode, PS). Dieser bremst die Datenübertragung in bestimmten Situationen etwas, reduziert aber die Leistungsaufnahme wesentlich.

Ist der Sendepuffer eines Endgeräts leer und wurden seit einiger Zeit auch keine Daten vom Access Point empfangen, kann ein Endgerät den PS-Mode aktivieren. Dazu sendet das Endgerät einen leeren Frame an den Access Point, in dessen MAC-Header das PS Bit gesetzt ist. Der Access Point puffert danach alle für das Endgerät eingehenden Frames, und das Endgerät kann somit die Stromzufuhr zu seinem Sender und Empfänger abschalten. Die Zeit zwischen letztem Datenpaket und dem Einschalten des PS-Modus kann vom Endgerätehersteller selbst bestimmt werden. In der Praxis werden hier Werte z. B. von batteriebetriebenen Geräten wie Mobiltelefonen mit Wi-Fi-Funktionalität von 0,5 s gewählt.

Möchte ein Endgerät wieder Daten senden, schaltet es seine Sende- und Empfangsstufe wieder ein und sendet einen leeren Frame mit deaktiviertem PS Bit. Danach können die neuen Frames mit Nutzdaten sofort gesendet werden (Abb. 4.10).

Bei den meisten Anwendungen auf mobilen Endgeräten, wie z. B. dem Webbrowsern, treffen nur in Ausnahmefällen nach dem Einschalten des PS-Modus weitere Daten ein. Damit diese nicht verloren gehen, werden die Frames im Access Point zwischengespeichert. Aus diesem Grund muss das Endgerät auch im PS-Modus periodisch seinen Empfänger aktivieren, um diese Pakete gegebenenfalls abholen zu können. Um ein Endgerät über gepufferte Frames zu informieren, gibt es in Beacon Frames das Traffic

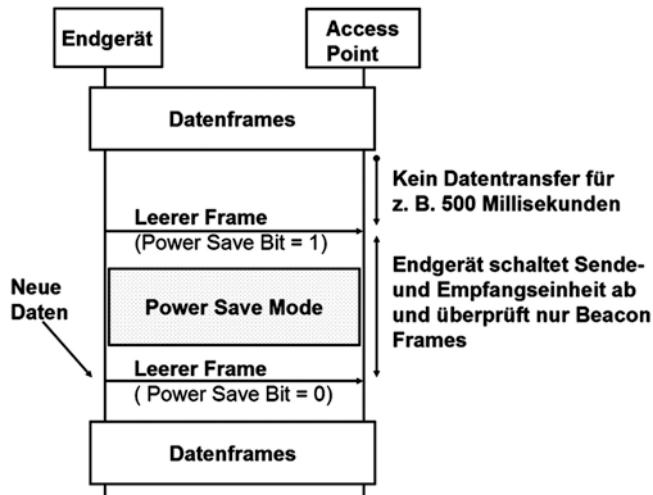


Abb. 4.10 Ein- und Ausschalten des Stromsparmodus (ohne Acknowledge Frames)

Indication Map (TIM) Information Element. Für jedes Endgerät ist in der TIM ein Bit vorhanden, das anzeigt, ob gepufferte Daten vorliegen. Das Endgerät identifiziert sein Bit in der TIM über seine Association ID (AID), die ihm bei der Association-Prozedur übergeben wurde. Über die AID können bis zu 2007 Endgeräte angesprochen werden, die TIM ist also maximal 2007 Bits lang. Um die Beacon Frames möglichst klein zu halten, wird mithilfe eines Offsets und einer Längenangabe nur ein Teil der TIM im Beacon Frame übertragen. Dies ist auch sinnvoll, da meist nur wenige Endgeräte an einem Access Point gleichzeitig betrieben werden.

Damit ein Endgerät nicht für jeden Beacon Frame seinen Empfänger einschalten muss, übergibt das Endgerät bei der Association-Prozedur ein Listen-Intervall an den Access Point, das vorgibt, in welchen Abständen die Beacon Frames überprüft werden. Akzeptiert der Access Point dieses Intervall, muss er eingehende Daten mindestens für diesen Zeitraum puffern. In der Praxis wird für das Listen-Intervall zum Beispiel ein Wert von 3 verwendet. Dies bedeutet, dass das Endgerät nur jeden dritten Beacon Frame empfängt und somit seinen Empfänger für 300 ms abschalten kann. Ist das TIM Bit für das Endgerät nicht gesetzt, kann es nach Empfang des Beacon Frames seinen Empfänger wieder für die nächsten 300 ms deaktivieren.

Ist das TIM Bit für ein Endgerät gesetzt, aktiviert es neben seinem Empfänger auch seine Sendeeinheit und ruft die gepufferten Datenpakete über PS-Poll Frames beim Access Point ab. Als Antwort auf einen PS-Poll Frame erhält das Endgerät dann einen gepufferten Frame. Ist im MAC-Header des Frames das More Bit gesetzt, sind noch weitere Frames im Access Point gepuffert, die dann jeweils durch einen weiteren PS-Poll Frame angefordert werden müssen.

Auch Broadcast und Multicast Frames, die an mehrere oder alle Endgeräte gerichtet sind, müssen für Endgeräte im Power-Save Mode gepuffert werden. Statt jedoch diese Frames für jedes Endgerät einzeln zu puffern, gibt stattdessen das erste Bit in der TIM an (AID 0), ob Broadcastdaten gepuffert wurden. Diese Frames werden dann automatisch nach einem Beacon Frame gesendet, der statt einer TIM periodisch eine Delivery TIM (DTIM) enthält. In welchen Abständen statt der TIM eine DTIM gesendet wird, wird über eine Periode und einen Count Down-Zähler in der TIM den Endgeräten mitgeteilt.

4.5 Die MAC-Schicht

Das Medium Access Control Protocol (MAC, Layer 2) hat bei WLAN ähnlich wie im drahtgebundenen Ethernet unter anderem folgende Aufgaben:

- Es regelt den Zugriff der Endgeräte auf das Übertragungsmedium.
- Jedem Datenpaket wird ein MAC Header vorangestellt, der unter anderem die Adresse des Senders und Empfängers (MAC-Adressen) enthält.

4.5.1 Zugriffssteuerung auf das Übertragungsmedium

Aufgrund der höheren Fehleranfälligkeit der Datenübertragung über die Luftschnittstelle werden bei WLAN alle Datenpakete von der Gegenstelle nach korrektem Empfang durch ein Acknowledgement (ACK) Frame bestätigt. Dies ist ein großer Unterschied zum drahtgebundenen Ethernet, in dem Pakete nicht bestätigt werden. In allen bisherigen Abbildungen dieses Kapitels wurden diese Frames zur Übersichtlichkeit weggelassen. Abb. 4.11 zeigt den Austausch von Frames zwischen Access Point und einem Endgerät zum ersten Mal mit ACK Frames. Jeder Frame, der entweder Nutzdaten oder Management-Daten (Authentication, Association, etc.) enthält, muss von der Gegenseite durch ein ACK Frame bestätigt werden. Erst danach darf der nächste Nutzdatenframe vom gleichen oder einem anderen Endgerät gesendet werden. Bleibt der ACK Frame aus, muss das Datenpaket wiederholt werden.

Durch einen sehr kurzen Sendeabstand zwischen Datenframe und ACK Frame, der Short Interframe Space (SIFS) genannt wird, ist sichergestellt, dass kein anderes Endgerät einen Frame dazwischen senden kann. Für normale Frames wird deshalb ein längerer Sendeabstand zum letzten Paket eingehalten, der DCF Interframe Space (Distributed Coordination Function Interframe Space, abgekürzt DIFS) genannt wird. Somit kann der ACK Frame auf jeden Fall gesendet werden, bevor eine andere Station den Kanal für einen normalen Frame verwenden darf. Mehr zum Thema DCF im nächsten Abschnitt.

Optional gibt es für ein Endgerät die Möglichkeit, die Luftschnittstelle für die Übertragung eines Frames im Vorhinein zu reservieren. Dies ist in Fällen sinnvoll, in denen

Abb. 4.11 Bestätigung (Acknowledgement) für jeden Frame

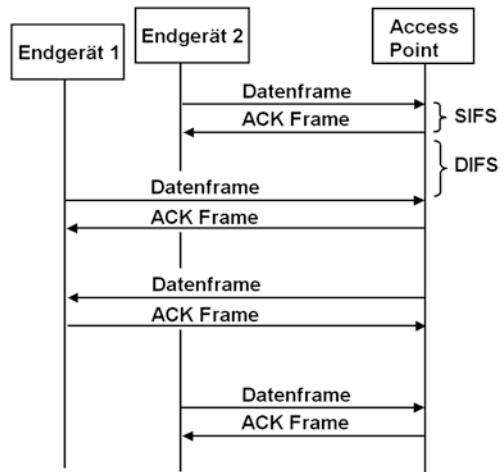
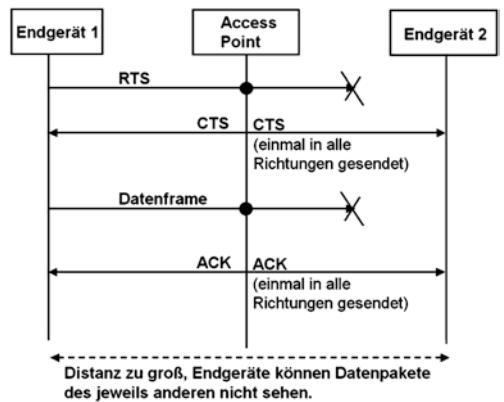


Abb. 4.12 Reservierung der Luftschnittstelle mit RTS/CTS



Teilnehmer eines BSS zu weit voneinander entfernt sind, um die Datenpakete des jeweils anderen zu sehen. In diesen Fällen kann es passieren, dass beide Stationen gleichzeitig einen Frame an den Access Point senden und sich die Frames am Access Point gegenseitig stören. Dieses Szenario wird auch „Hidden Station“-Problem genannt. Um dieses Problem zu umgehen, sendet ein Endgerät wie in Abb. 4.12 gezeigt vor dem Datenframe zuerst einen kurzen RTS (Ready to Send) Frame an den Access Point. Der Access Point antwortet daraufhin mit einem kurzen CTS (Clear to Send) Frame, und die Luftschnittstelle ist für den Teilnehmer reserviert. Während der RTS Frame vom zweiten Endgerät aufgrund des zu großen Abstands nicht gesehen wird, sieht es aber auf jeden Fall den CTS Frame, da dieser vom näheren Access Point gesendet wird. Damit das zweite Endgerät weiß wie lange es nicht senden darf, enthalten RTS und CTS Frames die Information, wie lange die Luftschnittstelle reserviert ist. Abgeschlossen wird die Übertragung des Frames wieder durch ein ACK Frame. Ob ein Endgerät ein Frame mit oder

ohne RTS/CTS-Sequenz überträgt, ist in den meisten Endgeräten in Abhängigkeit der Framegröße konfigurierbar. Dies ist sinnvoll, da der zusätzliche Zeitaufwand des RTS/CTS-Mechanismus nur bei großen Paketen sinnvoll ist. Meist ist diese Option jedoch per Default deaktiviert und muss manuell konfiguriert werden.

Bei Wireless LAN gibt es keine zentrale Steuerung, welcher Teilnehmer zu welchem Zeitpunkt auf das Übertragungsmedium (Luftschnittstelle) zugreifen darf. Jeder Teilnehmer trifft für sich die Entscheidung, wann ein anstehendes Datenpaket übertragen wird. Da aber möglichst keine Kollisionen mit anderen Teilnehmern auftreten sollen, koordinieren sich die Teilnehmer mit einem Verfahren, das Distributed Coordination Function (DCF) genannt wird. Dieser Ansatz unterscheidet sich grundlegend vom zentral gesteuerten Medienzugriff aller anderen Systeme, die in diesem Buch vorgestellt werden. Diese haben alle eine verwaltende Instanz, die genau vorgibt, welcher Teilnehmer zu welchem Zeitpunkt und für wie lange senden darf. Vorteil des DCF-Verfahrens ist die einfache Implementierung in Endgeräten. Großer Nachteil des Verfahrens ist jedoch, dass keine Bandbreite reserviert oder garantiert werden kann. Besonders für Echtzeitanwendungen wie Sprach- oder Videotelefonie ist dies ein Problem, wenn das Medium von anderen Teilnehmern stark ausgelastet wird. Aus diesem Grund wurde im 802.11e-Standard für Geräte und Anwendungen, die eine hohe Anforderung bezüglich konstanter Bandbreite und Medienzugriffszeit haben, eine DCF-Erweiterung spezifiziert, die in Abschn. 4.8 beschrieben wird.

Wichtigster Teil der DCF ist das Medienzugriffsverfahren, das Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) genannt wird. CSMA/CA ist CSMA/CD (CSMA/Collision Detect) sehr ähnlich, das im drahtgebundenen Ethernet verwendet wird, bietet aber einige zusätzliche Möglichkeiten, Kollisionen zu vermeiden.

Möchte ein Endgerät ein Datenpaket versenden, und es wird keine Aktivität auf der Luftschnittstelle festgestellt, kann das Datenpaket ohne Verzögerung gesendet werden. Wird jedoch zu diesem Zeitpunkt gerade ein Datenpaket eines anderen Teilnehmers übertragen, muss das Endgerät zunächst warten, bis diese Übertragung abgeschlossen ist. Danach wartet das Endgerät noch das Ende der DIFS-Periode ab. Um zu vermeiden, dass mehrere sendebereite Endgeräte danach gleichzeitig ihre Pakete absenden, wird zusätzlich noch eine per Zufallsgenerator in jedem Endgerät ermittelte Backoff-Zeit gewartet. Da mit großer Wahrscheinlichkeit jeder Teilnehmer eine andere Backoff-Zeit ermittelt hat, sendet somit nur ein Endgerät. Alle anderen sendebereiten Endgeräte sehen das Datenpaket, brechen ihre Backoff-Wartezeit ab und starten ihre Zugriffsprozedur erneut von vorn. Sollten trotz dieser Prozedur einmal zwei Endgeräte gleichzeitig senden, stören sich die Pakete gegenseitig und der Acknowledgement Frame bleibt aus. Beide Stationen müssen dann erneut versuchen, ihr Datenpaket zu senden. Bei einem Übertragungsfehler vergrößert sich jedoch die Zeitspanne für die mögliche Backoff-Zeit für das Endgerät. Somit wird erreicht, dass bei hoher Auslastung die Anzahl der Kollisionen gering bleibt.

Die Backoff-Zeit wird in Slots zu $20 \mu\text{s}$ eingeteilt. Beim ersten Sendevorschuss gibt es 31 Slots, von denen einer per Zufallsgenerator ausgewählt wird. Schlägt die Übertragung

fehl, vergrößert sich das Fenster auf 63 Slots, danach auf 127 Slots usw., bis maximal 1023 Slots, was maximal 20 ms entspricht. Bei 802.11n wurde das erste Backoff-Fenster auf 15 Slots verkleinert, was 0,3 ms entspricht.

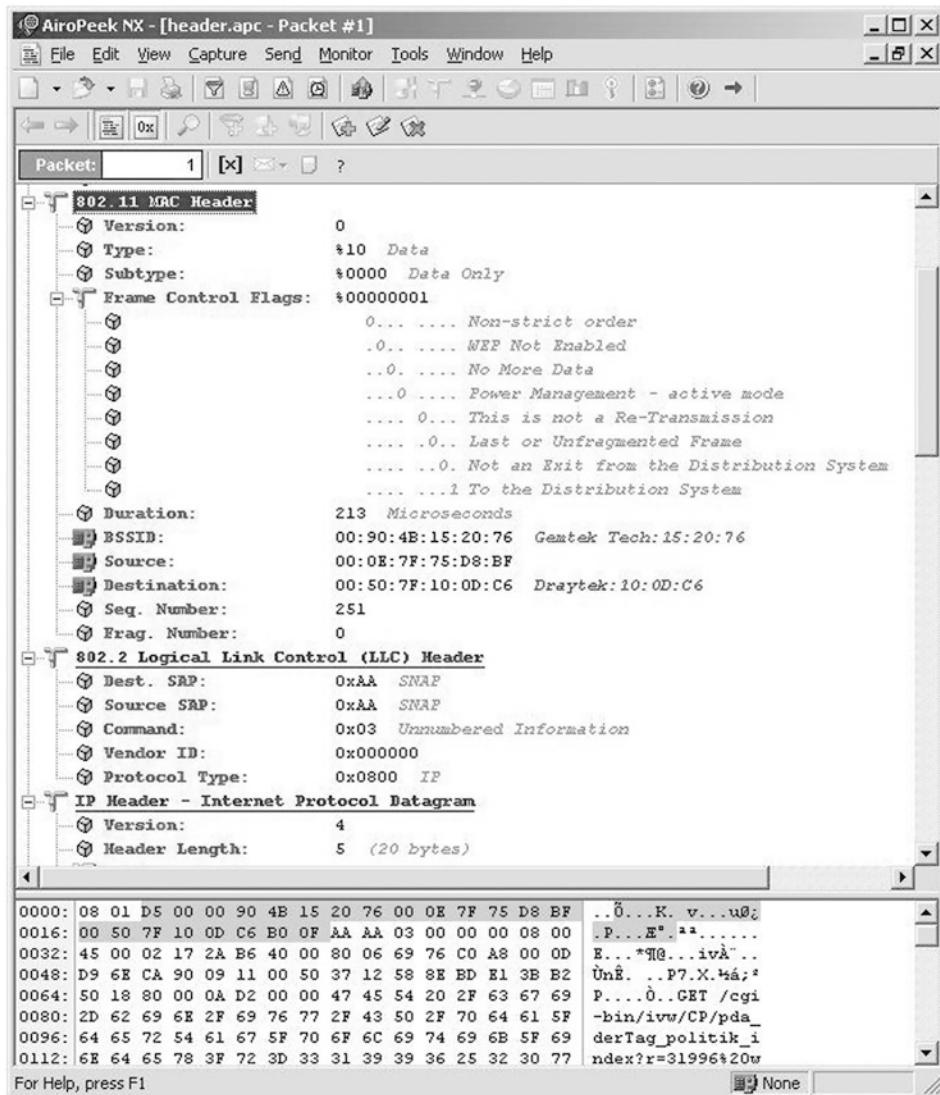
Zusätzlich zur Erkennung einer laufenden Datenübertragung und anschließender Backoff-Zeit enthält jedes Datenpaket auch eine Zeitspanne, wie lange die Übertragung des Datenpakets und anschließendem ACK Frame dauert. Diese Zeitspanne wird Network Allocation Vector (NAV) genannt. Diese zusätzliche Funktion ist vor allem dann sinnvoll, wenn wie in Abb. 4.12 gezeigt, die Luftschnittstelle mit RTS und CTS Frames reserviert wird. Das RTS Frame enthält die Information, wie viel Zeit für die Übertragung des CTS, des Datenpakets und das anschließende ACK Frame benötigt wird. Das anschließende CTS-Paket der Gegenseite enthält dann einen etwas kleineren NAV, der nur noch die Zeitspanne für das anschließende Datenpaket und den ACK Frame enthält.

4.5.2 Der MAC Header

Wichtigste Aufgabe des MAC Headers auf Layer 2 ist die Adressierung der Endgeräte im lokalen Netzwerk. Zu diesem Zweck enthält der MAC Header eines WLAN Frames in gleicher Weise wie ein Frame im drahtgebundenen Ethernet die 48 Bit langen MAC-Adressen von Sender (Source) und Empfänger (Destination). In einem Basic Service Set (BSS) wird ein Datenframe jedoch nicht direkt vom Sender zum Empfänger geschickt, sondern immer zuerst zum Access Point. Aus diesem Grund enthält der MAC Header eines Frames, wie in Abb. 4.13 gezeigt, nicht zwei, sondern drei MAC-Adressen. Die dritte MAC-Adresse ist dabei die Adresse des Access Points. Dieser empfängt das Paket und überprüft, ob die MAC-Adresse des Empfängers zu einem drahtlosen oder einem drahtgebundenen Endgerät gehört und leitet den Frame entsprechend weiter. Somit spielt es für das Endgerät keine Rolle, ob der Empfänger des Frames ein WLAN oder Ethernet-Endgerät ist.

Weitere wichtige Elemente im MAC Header sind der Frame Type und Subtype. Das Frame Type Element gibt an, ob es sich beim aktuellen Frame um einen Nutzdatenframe, Management Frame (z. B. Association Request) oder Control Frames (z. B. ACK) handelt. Je nach Frame Type enthält das Subtype Element dann weitere Informationen. Bei Management Frames gibt das Subtype-Feld an, um welche Management-Operation es sich handelt (z. B. Authentication, Association, Beacon, etc.).

In den Frame Control Flags werden in jedem Frame zusätzliche Informationen zwischen zwei Teilnehmern ausgetauscht. Dort ist unter anderem angegeben, ob die Nutzdaten des Frames verschlüsselt sind (WEP enabled Bit, veraltet und nicht mehr verwendet), ob das Endgerät in den Stromsparmodus wechseln wird (Power Management Bit) und ob der Frame für einen Access Point bestimmt ist (To Distribution System Bit).

**Abb. 4.13** MAC und LLC Header eines WLAN Frames

Bei Nutzdatenframes folgt auf den MAC Header, in gleicher Weise wie im drahtgebundenen Ethernet, der Logical Link Control Header (LLC Header, Layer 2). Dessen wichtigste Aufgabe ist es, das Layer 3-Protokoll zu identifizieren, das anschließend folgt.

4.6 Physical Layer und MAC-Erweiterungen

Auf Layer 1, dem Physical Layer, der auch als PHY bezeichnet wird, gibt es wie in Abschn. 4.2 gezeigt, unterschiedliche Varianten mit unterschiedlichen Geschwindigkeiten, die in den Standards IEEE 802.11b, g, a, n, ac und ax beschrieben sind.

4.6.1 IEEE 802.11b mit bis zu 11 Mbit/s

Der Durchbruch von WLAN im Consumer Segment kam Anfang der 2000er Jahre mit der Einführung von 802.11b kompatiblen Endgeräten, die eine maximale Geschwindigkeit von 11 Mbit/s boten. Auch wenn heute PHYs wie 802.11ac und 11ax den Markt dominieren und weit höhere Geschwindigkeiten ermöglichen, sind die grundsätzlichen Mechanismen für die Zugriffskontrolle und das Netzmanagement noch die Gleichen. Auch sind alle neueren PHYs noch rückwärtskompatibel zu 802.11b. Es können somit auch noch sehr alte 802.11b Geräte in aktuellen Netzwerken verwendet werden. In neuen APs ist es jedoch möglich, die 802.11b Unterstützung zu deaktivieren und somit den Datendurchsatz zu erhöhen. Um die Grundzüge der WLAN Technik und auch die Rückwärtskompatibilität besser zu verstehen, sowie die Kompromisse, die eingegangen werden müssen, um auch ältere Geräte zu unterstützen, gibt dieser Abschnitt eine kurze Einführung in den 802.11b PHY, auch wenn dieser in der Praxis heute keine große Bedeutung mehr hat.

Um ein Gefühl für ein 802.11 System im Vergleich zu anderen Technologien zu bekommen, sind folgende Daten hilfreich:

- Die maximale Leistung der Endgeräte und Wi-Fi Access Points ist auf 0,1 W begrenzt. Mobiltelefone hingegen dürfen mit 1 bis 2 W je nach Frequenzband senden. Auf Netzwerkseite können LTE und 5G NR Basisstationen mit bis 100–200 W pro Sektor und Frequenz senden.
- 22 MHz Bandbreite pro Kanal. Somit können im ISM-Band drei Netze am gleichen Ort überlappend betrieben werden. GSM nutzt 0,2 MHz (200 kHz) pro Kanal, während LTE eine Kanalbandbreite von bis zu 20 MHz nutzt. 5G NR verwendet sogar eine Kanalbandbreite von bis zu 100 MHz. Zusätzlich wird Carrier Aggregation eingesetzt, um die Kanalbandbreite weiter zu erhöhen.
- Framegröße: 4–4095 Bytes, IP Frames sind jedoch meist nicht größer als etwa 1500 Bytes. Interessant ist hier der Vergleich zu anderen Technologien: In einem GPRS-Paket, das aus 4 Bursts zu je 114 Bits besteht, können nur 456 Bits übertragen werden. Bei Coding Scheme 2 bleiben hier nach Abzug der Fehlerkorrekturbits nur 240 Bits, also 30 Bytes. Während ein IP-Paket über WLAN komplett in einem Paket übertragen wird, wird dieses in GPRS über mehrere Pakete aufgeteilt.

- Übertragungszeit eines großen Pakets: Dies ist zum einen von der Größe des Pakets und zum anderen von der Übertragungsraten abhängig. Wird ein großes Paket mit z. B. 1500 Bytes mit einer Übertragungsgeschwindigkeit von 1 Mbit/s übertragen, dauert die Übertragung 12 ms. Bei gutem Empfang und einer Übertragungsgeschwindigkeit von 11 Mbit/s dauert die Übertragung hingegen nur etwa 1,1 ms. Hinzu kommt noch die Übertragungszeit für den ACK Frame, sowie die Sendepause zwischen den Frames.
- Zeit zwischen Datenframe und ACK Frame (SIFS): 10 μ s, oder 0,01 ms.
- Tritt ein Übertragungsfehler auf, wird das im letzten Absatz beschriebene Backoff-Verfahren angewandt. Ein Backoff Slot, von denen es bei der ersten Wiederholung 63 gibt, hat eine Länge von 20 μ s oder 0,02 ms.
- Zu Beginn jedes Frames wird eine Präambel gesendet, die anderen Endgeräten die Übertragung ankündigt. Dies ist notwendig, damit sich alle anderen Endgeräte auf den Frame synchronisieren können. Die Präambel hat eine Länge von 144 μ s, oder 0,144 ms.

Die Präambel ist Teil des Physical Layer Convergence Protocol (PLCP) Header, der vor jedem Frame gesendet wird. Der PLCP Header enthält auch die Information, mit welcher Übertragungsraten der nachfolgende MAC Frame gesendet wird. Bei 802.11b kann ein MAC Frame mit 1 Mbit/s, 2, 5,5 und 11 Mbit/s gesendet werden. Diese Flexibilität ist nötig, da Endgeräte mit schlechtem Empfang mit geringer Geschwindigkeit senden können, um somit die Redundanz zu erhöhen. Üblicherweise entscheidet das Endgerät automatisch anhand der Übertragungsbedingungen, mit welcher Geschwindigkeit ein Frame gesendet werden soll. Beacon Frames hingegen werden beispielsweise von manchen Access Points immer mit 1 oder 2 Mbit/s übertragen. Auf diese Weise können auch weiter entfernte Geräte die Beacon Frames noch korrekt empfangen. Dies ist aber nicht vorgeschrieben, und so senden manche Access Points die Beacon Frames mit 11 Mbit/s. Dies erhöht zwar den Durchsatz des Netzwerkes geringfügig, weit entfernte Stationen werden aber Probleme haben, die Beacon Frames korrekt zu empfangen.

Für die Codierung der Daten eines Frames für die Übertragung mit 1 oder 2 Mbit/s wird das Direct Sequence Spread Spectrum-Verfahren (DSSS) verwendet. Ein Bit wird dabei nicht direkt übertragen, stattdessen werden 11 Chips übertragen. Für ein Bit mit dem Wert 1 wird die Chipsequenz „0,1,0,0,1,0,0,0,1,1,1“ übertragen, für ein Bit mit dem Wert 0 die Sequenz „1,0,1,1,0,1,1,1,0,0,0“. Diese Sequenzen werden auch Barker Code genannt. Da statt einem Wert nun 11 Werte pro Bit übertragen werden, erhöht sich die Redundanz ganz erheblich. Somit ist es möglich, auch bei einigen nicht korrekt empfangenen Chips das übertragene Bit dennoch korrekt zu erkennen.

Auch UMTS macht sich dieses Verfahrens, das „spreizen“ genannt wird, für die Erhöhung der Redundanz zunutze. Während bei WLAN jedoch nur ein Endgerät zu einer Zeit sendet (Time Division Multiple Access), ermöglicht das Spreizen bei UMTS zusätzlich die gleichzeitige Datenübertragung von mehreren Endgeräten (Code Division

Multiple Access). Bei UMTS werden jedoch keine festen Sequenzen, sondern variable orthogonale Codes verwendet.

Die Barker Chip-Sequenz wird anschließend mit dem Differential Binary Phase Shift Keying (DBPSK)-Verfahren mit einer Übertragungsgeschwindigkeit von 11 Mchips/s übertragen. Dies ergibt somit eine Bitrate von 1 Mbit/s. Beim DBPSK-Verfahren ändert sich bei jeder Übertragung eines Chips mit dem Wert 1 die Phasenlage des Sinussignals um 180 Grad. Bei einem Chip mit dem Wert 0 hingegen ändert sich die Phasenlage nicht.

Für eine Übertragungsgeschwindigkeit von 2 Mbit/s wird statt DBPSK das Differential Quadrature Phase Shift Keying (DQPSK)-Verfahren angewandt. Statt einem Chip pro Übertragungsschritt werden hier 2 Chips übertragen. Die 4 (quadrature) möglichen unterschiedlichen Werte (00, 01, 10 oder 11) der 2 Chips werden in diesem Verfahren mit 90 Grad Phasenwechseln pro Übertragungsschritt kodiert.

Um bei gleicher Bandbreitennutzung noch schnellere Datenraten zu ermöglichen, wurde mit dem 802.11b Standard das Complementary Code Keying (CCK)-Verfahren unter dem Namen High Rate DSSS (HR/DSSS) eingeführt. Statt ein Bit statisch in einer 11 Chip Barker-Sequenz zu kodieren, werden die Bits beim CCK-Verfahren wie folgt übertragen:

Um eine Datenrate von 11 Mbit/s zu erhalten, werden die Bits eines Frames wie in Abb. 4.14 gezeigt, zunächst in 8 Bit-Blöcke eingeteilt. Die ersten zwei Bits werden wie beim vorhergehenden Verfahren auch per DQPSK in eine Änderung der Phasenlage in 90 Grad Schritten übertragen.

Aus den restlichen 6 Bits wird danach ein 8 Chip-Codewort gebildet. Dieses 8 Chip-Codewort wird auch Symbol genannt. Da 6 Bit in einem 8 Bit Symbol kodiert werden, ist auch hier noch eine gewisse Redundanz enthalten. Das so erhaltene Symbol wird wiederum in 4 Teile zu 2 Bits unterteilt und danach in Phasenänderungen kodiert übertragen.

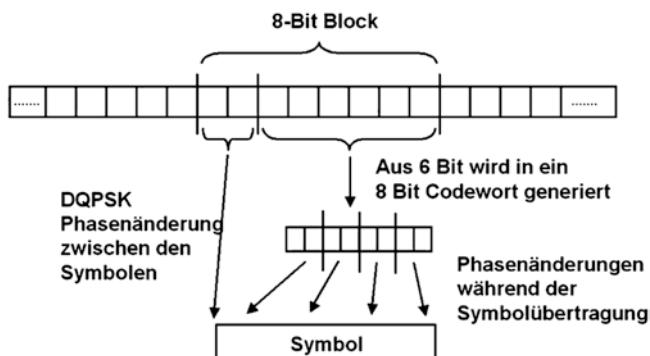


Abb. 4.14 Complementary Code Keying für 11 Mbit/s

Da die Taktgeschwindigkeit zum 1 bzw. 2 Mbit/s Verfahren nicht geändert wurde, können auf diese Weise 11 Mbit/s übertragen werden. Nachteil ist jedoch, dass in den übertragenen Informationen wesentlich weniger Redundanz vorhanden ist.

Damit auch Endgeräte mit schlechten Empfangsbedingungen keine Kollisionen erzeugen, muss zumindest der Beginn eines Frames korrekt empfangen werden können. Um dies zu gewährleisten, wird der PLCP Header immer mit einer Geschwindigkeit von 1 Mbit/s übertragen, auch wenn der anschließende MAC Frame mit 11 Mbit/s übertragen wird. Da auch die Übertragungszeit für den anschließenden MAC Frame im PLCP Header enthalten ist, weiß ein empfangendes Endgerät genau, wie lange das Medium besetzt ist, auch wenn die nachfolgenden „schnellen“ Bits nicht korrekt empfangen werden können.

Vergleicht man die tatsächliche Geschwindigkeit eines 11 Mbit/s Wireless LANs mit einem damals üblichen 10 Mbit/s drahtgebundenen Ethernet, ist ein deutlicher Unterschied sichtbar. Ein 10 Mbit/s Ethernet ermöglichte unter idealen Bedingungen einen maximalen Durchsatz von etwa 700–800 kB/s. Bei 802.11b betrug die maximale Geschwindigkeit beim Datenaustausch zwischen zwei mobilen Endgeräten hingegen „nur“ 300 kB/s. Dies lag an folgenden WLAN-Eigenschaften, die in diesem Abschnitt beschrieben wurden:

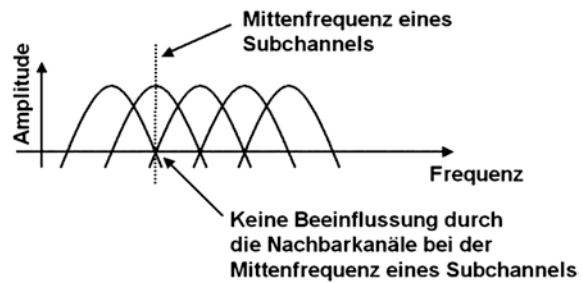
- Der PLCP Header jedes WLAN Frames wird mit 1 Mbit/s gesendet.
- Auf jeden Frame muss der Empfänger mit einem ACK Frame antworten. Auch dies kostet zusätzliche Zeit.
- Während bei Ethernet ein Frame direkt zum Empfänger geschickt wird, muss ein Frame in einem WLAN BSS zuerst an den Access Point geschickt werden. Dieser sendet das Paket dann an den Empfänger. Die Luftschnittstelle wird somit durch das Paket zweimal belegt. Die maximale Datenrate reduziert sich somit um die Hälfte.

4.6.2 IEEE 802.11g mit bis zu 54 Mbit/s

Um die Übertragungsgeschwindigkeit weiter zu erhöhen, wurde für die 802.11g-Standarderweiterung zum ersten Mal in den Wireless LAN Spezifikationen das Orthogonal Frequency Division Multiplexing (OFDM) Modulationsverfahren verwendet. Mit dieser Modulation waren bei etwa gleicher Bandbreitennutzung wie bei 802.11b Geschwindigkeiten von bis zu 54 MBit/s möglich. Im Standard wird dieser Physical Layer als Extended-Rate (ERP) PHY bezeichnet. Auch alle nachfolgenden Spezifikationserweiterungen wie 802.11n, 11ac und 11ax verwenden dieses Modulationsverfahren.

Das OFDM-Modulationsverfahren unterscheidet sich grundlegend von den in 802.11b verwendeten Techniken. Wie Abb. 4.15 vereinfacht zeigt, teilt OFDM den Übertragungskanal von etwa 20 MHz in 52 Unterkanäle (Sub-Channels) auf, über die unabhängig voneinander Daten übertragen werden können.

Abb. 4.15 Vereinfachte Darstellung des OFDM-Modulationsverfahrens



Die Unterkanäle werden als Orthogonal bezeichnet, weil die Amplituden der Nachbarkanäle an der Mittenfrequenz eines anderen Kanals genau Null sind. Somit haben sie keinen Einfluss auf die Amplitude eines anderen Kanals. Um Daten auf die Unterkanäle aufzumodulieren, wird beim OFDM-Verfahren keine Phasenverschiebung wie in den bisherigen Verfahren verwendet, stattdessen werden die Informationen über die Höhe der Amplitude kodiert. Je nach Empfangsqualität des Signals wird die Amplitude in eine unterschiedliche Anzahl von Stufen aufgeteilt.

Um das Signal zu demodulieren, wird im Empfänger für jeden Übertragungsschritt eine FFT (Fast Fourier Transformation)-Analyse durchgeführt. Mit diesem Verfahren ist es möglich, die Signalenergie (Amplitude) über das Frequenzband zu berechnen. Das Ergebnis einer FFT ist vereinfacht in Abb. 4.15 gezeigt. Das Frequenzband ist dabei auf der x-Achse aufgezeichnet (statt wie bei anderen Verfahren die Zeit), die Amplitude auf der y-Achse.

Die folgende Tabelle zeigt die bei 802.11g möglichen Geschwindigkeiten:

Geschwindigkeit (MBit/s)	Modulation und Coding	Kodierte Bits pro Kanal	Kodierte Bits in 48 Kanälen	Datenbits pro Schritt
6	BPSK, R=1/2	1	48	24
9	BPSK, R=3/4	1	48	36
12	QPSK, R=1/2	2	96	48
18	QPSK, R=3/4	2	96	72
24	16-QAM, R=1/2	4	192	96
36	16-QAM, R=3/4	4	192	144
48	64-QAM, R=2/3	6	288	192
54	64-QAM, R=3/4	6	288	216

Bei günstigen Übertragungsbedingungen kann z. B. das 64 Quadrature Amplitude Modulation (64QAM)-Verfahren in den Unterkanälen verwendet werden. Zusammen mit einem $\frac{3}{4}$ Convolutional Coder (3 Datenbits pro 4 übertragenen Bits) und einer Schrittgeschwindigkeit (Symbol Speed) von 250.000 Symbolen/s wird dadurch eine Geschwindigkeit von 54 Mbit/s erreicht (216 Bits pro Schritt * 250.000

Symbole/s = 54 Mbit/s). Der Convolutional Coder, auch Faltungskodierer genannt, dient zur Erhöhung der Redundanz und wird auch bei GSM und UMTS verwendet. 802.11g-Endgeräte und Access Points sind abwärtskompatibel zu langsameren 802.11b-Geräten. Das bedeutet, dass ein 802.11g Access Point auch 802.11b Endgeräte unterstützt, die mit maximal 11 Mbit/s senden können. Im umgekehrten Fall können auch 802.11g-Endgeräte mit 802.11b Access Points kommunizieren, wobei dann die Datenrate natürlich auf 11 Mbit/s begrenzt ist.

Da langsame 802.11b Endgeräte die für sie unbekannte OFDM Modulationsart nicht erkennen können, müssen 802.11g-Geräte Schutzmaßnahmen ergreifen, sobald sich ein älteres 802.11b Gerät am Netzwerk anmeldet. Während mindestens ein solches Gerät am Access Point angemeldet ist, informiert dieser über einen Parameter in den Beacon Frames alle Teilnehmer des Netzwerkes. 802.11g-Geräte senden dann vor dem eigentlichen Datenpaket ein Clear To Send (CTS) Paket. Dieses kann auch von 802.11b Endgeräten dekodiert werden und enthält die Zeitdauer, die die Luftschnittstelle danach belegt ist. Somit ist sichergestellt, dass 802.11b Endgeräte nicht gleichzeitig mit 802.11g-Geräten senden. Außerdem muss der PLCP Header jedes Frames mit 1 Mbit/s gesendet werden, um von allen Geräten korrekt erkannt zu werden. Zusammen bringt dies in der Praxis jedoch aufgrund des zusätzlichen Overheads einen Geschwindigkeitsverlust von bis zu 40 % mit sich. Aus diesem Grund kann in den meisten Access Points auch ein „G-Only“ Mode eingeschaltet werden, der diesen zusätzlichen Overhead vermeidet, ältere 802.11b-Geräte jedoch ausschließt.

Unter optimalen Übertragungsbedingungen sind in der Praxis Übertragungsgeschwindigkeiten von etwa 2500 kB pro Sekunde möglich. Kommunizieren zwei drahtlose Endgeräte miteinander, reduziert sich die maximale Geschwindigkeit auf etwa 1200 kB pro Sekunde, da alle Frames zuerst zum Access Point übertragen werden und erst von dort zum Empfänger weitergeschickt werden. Im Vergleich zu einem 802.11b-Netz mit 600 bzw. 300 kB/s zwischen zwei mobilen Endgeräten stellt der 802.11g Standard einen beachtlichen Fortschritt dar. Jedoch bleibt der Standard noch weit hinter einem 100 Mbit/s drahtgebundenen Ethernet zurück, das mit einer Datenrate von etwa 7000 kB/s immer noch etwa um den Faktor 3 schneller ist.

4.6.3 IEEE 802.11a mit bis zu 54 Mbit/s

Der 802.11a Standard ist im Wesentlichen mit dem zuvor beschriebenen 802.11g Standard identisch. Dieser Standard sendet jedoch nur im 5 GHz-Bereich und ist somit nicht mit 802.11b-Netzen kompatibel. Dies hat jedoch auch den Vorteil, dass die bei 802.11g verwendeten Verfahren für die Rückwärtskompatibilität hier nicht angewandt werden müssen und der PLCP Header statt mit 1 Mbit/s mit 6 Mbit/s gesendet werden kann. Reine 802.11a-Netze sind somit deutlich schneller als gemischte 802.11b/g-Netze und haben auch gegenüber reinen 802.11g-Netzen einen kleinen Geschwindigkeitsvorteil durch den schnelleren PLCP Header. In der Praxis sind 802.11a Netzwerke heute

nur noch sehr selten anzutreffen, da, wie nachfolgend beschrieben, im 5 GHz Band heute 802.11ac und 11ax kompatible Geräte verwendet werden.

4.6.4 IEEE 802.11n mit bis zu 600 Mbit/s

Wie in Abschn. 4.6.2 gezeigt, waren mit dem 802.11g Standard Übertragungsgeschwindigkeiten unter günstigen Bedingungen von 20–25 MBit/s auf Applikationsebene zu erreichen. Für ADSL oder Kabelanschlüsse war diese Geschwindigkeit ausreichend. Heutige ADSL2+, VDSL, Fiber to the Home (FTTH) und Kabelanschlüsse bieten jedoch weit höhere Geschwindigkeiten. Auch für die Anbindung von Endgeräten an zentrale Datei- oder Medienserver im Büro oder im Heimbereich, sowie für Anwendungen wie High Definition Video Streaming wird das Wireless LAN-Netzwerk schnell zum Nadelöhr. Aus diesen Gründen entschlossen sich eine große Anzahl von Firmen in der 802.11n Arbeitsgruppe den Standard weiterzuentwickeln. Hauptziel für viele Firmen war die Erhöhung der Datenrate. Weitere Ziele waren die Erhöhung der Reichweite und die Einführung von Quality of Service (QoS)-Mechanismen, um Applikationen wie Sprachtelefonie über IP (VoIP) oder Videostreaming auch in stark genutzten Drahtlosnetzwerken oder größeren Entfernungsmit guter Qualität zu ermöglichen. Aufgrund der großen Anzahl an Firmen, die sich an der Standardisierung beteiligten, wurde die 802.11n Erweiterung des Wireless LAN Standards sehr umfangreich und enthält zahlreiche optionale Funktionalitäten, die in der Praxis jedoch kaum genutzt wurden. Im Folgenden werden deshalb zunächst jene neuen Funktionen des High Throughput (HT) Physical Layers (PHY), sowie jene MAC Layer-Erweiterungen beschrieben, die im Standard fest vorgeschrieben sind, sowie jene Optionen, die auch im Consumer-Segment weite Verbreitung fanden.

Einfachstes Mittel, um die Geschwindigkeit zu steigern, war die Verbreiterung des Übertragungskanals. Zusätzlich zu 20 MHz-Kanälen erlaubt der Standard nun auch die Verwendung von 40 MHz-Kanälen. In der Praxis wurde dies schon von vielen Herstellern mit 802.11g proprietär implementiert, Endgeräte unterschiedlicher Hersteller waren jedoch nicht untereinander kompatibel.

Für die Datenübertragung werden bei 802.11n statt 52 OFDM-Subkanäle wie bei 802.11g nun 56 OFDM-Subkanäle in einem 20 MHz-Kanal verwendet. Die Bandbreite pro Subkanal ist bei beiden Varianten 312,5 kHz. Dies wurde erreicht, indem jeweils rechts und links im Frequenzband zwei weitere Subkanäle verwendet werden, die bei 802.11g noch nicht genutzt wurden. Die Anzahl der Pilotkanäle, die dem Empfänger das Ausmessen des Kanals ermöglichen und keine Nutzdaten übertragen, bleibt in beiden Varianten bei vier. In einem 40 MHz-Kanal werden insgesamt 114 Subkanäle verwendet, von denen 6 als Pilot verwendet werden.

	20 MHz non-HT (wie 802.11g)	20 MHz HT	40 MHz HT
Anzahl Carrier	48	52	108 (2 * 54)
Anzahl Pilots	4	4	6
Gesamte Anzahl an Carriern	52	56	114 (2 * 57)
Nicht benutzte Carrier in der Mitte	1	1	3

Der ursprüngliche Wireless LAN Standard verlangte nach jeder Übertragung eines Pakets eine Empfangsbestätigung der Gegenstelle durch ein Acknowledgement (ACK) Frame wie zuvor in Abb. 4.11 gezeigt. Dies ist bei einem unzuverlässigen Übertragungsmedium wichtig, um Übertragungsfehler schnell korrigieren zu können, hat jedoch den Nachteil, dass die Luftschnittstelle nicht sehr effizient genutzt wird. Erst mit 802.11e wurden effizientere Verfahren standardisiert, die in Abschn. 4.8 und Abb. 4.31 näher beschrieben werden. Um den Overhead weiter zu reduzieren, wurde im 802.11n Standard auf dem MAC Layer ein weiteres Verfahren eingeführt, um Pakete gebündelt übertragen zu können. Dieses Verfahren wird Frame Aggregation genannt. Statt jedes Paket einzeln zu übertragen und danach auf eine Bestätigung zu warten, kann der Sender jetzt Pakete auf dem MAC Layer bis zu einer Gesamtgröße von 65.535 Byte bündeln und gemeinsam übertragen. Der Empfänger bestätigt dann das gesamte Bündel mit nur einem ACK-Paket. Der Overhead wird dadurch vor allem dann stark minimiert, wenn ein Endgerät große Datenmengen überträgt und somit den Sendepuffer der Netzwerkkarte ständig gefüllt hält. Ein großer Nachteil ist jedoch, dass bei einem Übertragungsfehler das komplette Paket erneut übertragen werden muss (Abb. 4.16).

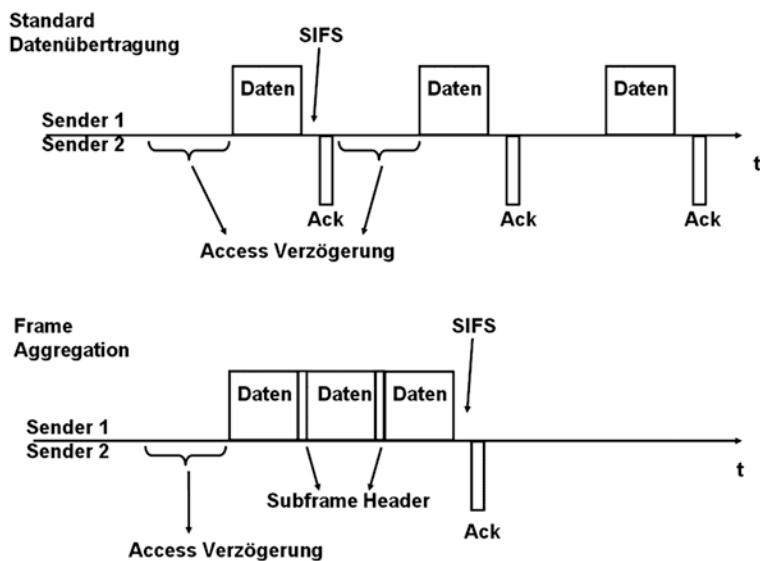


Abb. 4.16 Normale Datenübertragung im Vergleich mit Frame Aggregation

Ein weiterer Parameter für die Optimierung der Luftschnittstelle ist das OFDM Guard-Intervall. Dieses ist bei OFDM-Übertragungen notwendig, um die Interferenz zwischen aufeinander folgenden Symbolen abklingen zu lassen. In der Praxis zeigte sich, dass für die meisten Umgebungen ein Guard-Intervall von 400 ns pro Symbol statt bisher 800 ns ausreicht. Die Übertragungszeit eines OFDM Symbols verringert sich dadurch deutlich von 4 auf 3,6 µs, d. h. es können im gleichen Zeitraum mehr Symbole, also mehr Daten übertragen werden.

Eine weitere Möglichkeit, die Geschwindigkeit leicht zu steigern, ist die Anzahl der Fehlerkorrekturbits weiter zu senken. Die niedrigste Codierrate in 802.11g-Netzwerken ist 3/4, d. h. in 4 Bits sind drei Nutzdatenbits und ein Fehlerkorrekturbit enthalten. Bei 802.11n ist jetzt bei sehr guten Übertragungsbedingungen auch ein 5/6 Codierverfahren erlaubt, das für 5 Nutzdatenbits nur ein Fehlerkorrekturbit enthält.

Alle bisherigen Maßnahmen zusammen steigern die Geschwindigkeit um etwa das 2,5-fache verglichen mit 802.11g auf bis zu 150 Mbit/s. Wie bei früheren Standards auch, bleibt aufgrund der Acknowledgement Frames und anderen Eigenschaften der Luftschnittstelle für Applikationen in etwa die Hälfte dieser Geschwindigkeit übrig.

Wie am Anfang des Kapitels in Abb. 4.5 gezeigt, finden im 2,4 GHz ISM-Band nur drei unabhängige Netzwerke mit einer Bandbreite von 20 MHz Platz. Besonders in Städten teilen sich jedoch weit mehr Netze das ISM-Band. In einer solchen Situation schreibt der Standard vor, dass ein Access Point bei Empfang von Frames anderer Netzwerke in einem der zwei für den 40 MHz-Doppelkanal verwendeten Bänder sofort in den 20 MHz-Kanalmodus zurückschalten muss und erst 30 min nach dem letzten Auffinden eines Frames eines anderen Netzwerkes den breiteren Kanal wieder aktivieren darf. In der Praxis kann somit der 40 MHz-Kanalmodus im 2,4 GHz Band nur in den wenigsten Fällen verwendet werden. Zwar kann der Access Point in einem solchen Fall die Frequenz wechseln und dies den Endgeräten über Channel Switch Announcement Management Frames mitteilen, dies wird jedoch im überfüllten 2,4 GHz Band nur in den seltensten Fällen helfen. Manche Hersteller sind deshalb dazu übergegangen, diese Vorgabe zu ignorieren.

Der Standard ermöglicht auch die Verwendung des 5 GHz Bandes, in dem bis zu neun 40 MHz oder achtzehn 20 MHz Netzwerke Platz finden. Da dieser Frequenzbereich bisher nur selten genutzt wird, ist es hier meist ohne Probleme möglich, einen breiteren Kanal zu betreiben. In der Praxis bieten mittlerweile jedoch viele Access Point- und Endgerätehersteller 802.11n-Geräte für den 5 GHz-Bereich an. Nachteil des 5 GHz Bandes ist jedoch die kürzere Reichweite und die schlechtere Durchdringung von Wänden und anderen Hindernissen. In der Praxis führt dies dazu, dass im 5 GHz Band die Reichweite eines Access Points wesentlich geringer ist, als im 2,4 GHz Band.

Um die Geschwindigkeit und Reichweite weiter zu steigern, wurden im Standard sowohl für 20 MHz wie auch für 40 MHz-Kanäle diverse Multiple Input – Multiple

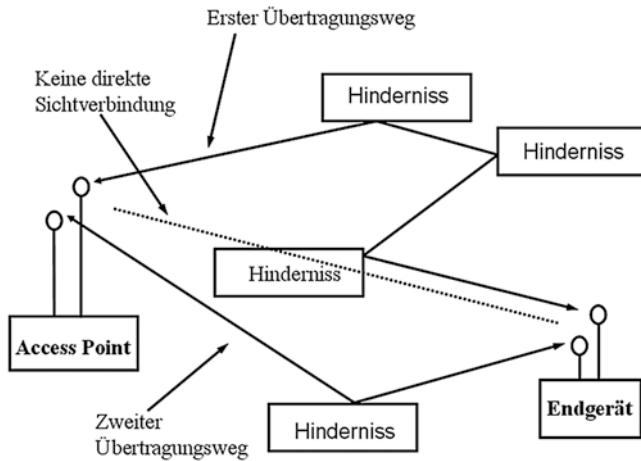


Abb. 4.17 2 × 2 MIMO

Output (MIMO)-Verfahren spezifiziert. MIMO nutzt den Umstand, dass bei der Funkübertragung zwischen einem Sender und einem Empfänger ein Signal an Objekten reflektiert und der Empfänger somit nicht nur ein Signal, sondern mehrere identische sieht, die jedoch aus unterschiedlichen Richtungen kommen. Bei MIMO Spatial Multiplexing haben sowohl Sender wie auch Empfänger mehrere Antennen und auch mehrere Sende- bzw. Empfangsstufen. Der Sender sendet nun auf jeder Antenne auf der gleichen Frequenz einen anderen Datenstrom, die dann am Empfänger wieder von getrennten Empfangsstufen empfangen werden. Dies ist in Abb. 4.17 gezeigt.

Im Standard sind bis zu 4 MIMO-Kanäle vorgesehen. Access Points müssen mindestens 2 MIMO-Kanäle unterstützen, andere 802.11n-Endgeräte wie z. B. Smartphones dürfen auch mit nur einen MIMO Pfad ausgestattet sein. Diese Regelung ist sinnvoll, da Access Points meistens nur wenige Restriktionen für Baugröße und Stromaufnahme haben. Kleine batteriebetriebene Geräte jedoch können mit nur einem MIMO-Zweig kleiner und stromsparender sein. Zudem werden solche Endgeräte in der Praxis kaum die höheren Geschwindigkeiten benötigen. Da Endgeräte während der Association-Prozedur dem Access Point ihre Fähigkeiten mitteilen können, kann dieser dann z. B. einen 20 MHz-Kanal ohne MIMO für ein VoIP-Telefon verwenden und für das nächste Paket an ein anderes Endgerät einen 40 MHz-Kanal mit zwei oder mehr MIMO-Zweigen.

Insgesamt gibt es aufgrund der zahlreichen Variablen wie Anzahl der MIMO-Kanäle, langer oder kurzer Guard Time, Modulation und Kodierung nun 77 mögliche Kombinationen, die zu unterschiedlichen Übertragungsgeschwindigkeiten führen. Die nachfolgende Tabelle zeigt exemplarisch einige Möglichkeiten.

	20 MHz, kein MIMO	20 MHz, 2 MIMO Streams	40 MHz 2 MIMO Streams
802.11b	1, 2, 5, 5, 11 Mbit/s		
802.11g	1, 2, 6, 9, 12, 18, 24, 36, 48, 54 Mbit/s		
802.11n, GI 800ns	6,5, 13, 19,5, 26, 39, 52, 58,5, 65 Mbit/s	13, 26, 39, 52, 78, 104, 117, 130 Mbit/s	27, 54, 81, 108, 162, 216, 243, 270 Mbit/s
802.11n, GI 400ns	7,2, 14,4, 21,7, 28,9, 43,3, 57,8, 65, 72,2 Mbit/s	14,4, 28,9, 43,3, 57,8, 86,7, 115,6, 130, 144,4 Mbit/s	30, 60, 90, 120, 180, 240, 270, 300 Mbit/s

Die Tabelle zeigt auch anschaulich den Einfluss der Kanalbündelung und des kürzeren Guard-Intervalls (GI). Durch die Kanalbündelung wird die Geschwindigkeit etwas mehr als verdoppelt, da zwischen den zwei Kanälen keine ungenutzten Subkanäle liegen und weniger Pilotkanäle verwendet werden. Der Einfluss des kürzeren Guard-Intervalls zeigt sich vor allem bei einem 40 MHz-Kanal mit zwei MIMO Streams. Durch das kürzere Guard-Intervall kann die maximale Geschwindigkeit von 270 Mbit/s auf 300 Mbit/s gesteigert werden.

Zusammen mit den zuvor beschriebenen Verfahren, ergibt sich mit 2×2 MIMO (2 Senderantennen, 2 Empfängerantennen) eine maximale Geschwindigkeitssteigerung gegenüber 802.11g von Faktor 5 auf etwa 300 Mbit/s auf der Luftschnittstelle. In einem 4×4 MIMO-System, das 4 Antennen sowohl beim Sender als auch beim Empfänger benötigt, ist eine theoretische Maximalgeschwindigkeit von bis zu 600 Mbit/s möglich.

In der Praxis erreichen 2×2 MIMO-Systeme auf dem Applikationslayer eine maximale Geschwindigkeit zwischen 80 und 110 Mbit/s. Dies kann aber nur unter günstigen Bedingungen, also auf kurzer Distanz von wenigen Metern, keine dicken Mauern zwischen den Geräten und im Greenfield Mode erreicht werden. Außerdem muss der Access Point unbedingt Gigabit Ethernet Ports unterstützen, um Datenraten über 100 Mbit/s auch tatsächlich weiterleiten zu können. Unter weniger optimalen Bedingungen wählen die Endgeräte automatisch statt einer 64-QAM Modulation eine robustere Modulation (16-QAM, QPSK oder BPSK) und statt einer 5/6 Fehlerkorrektur-Kodierung nur 3/4, 2/3 oder 1/2.

Eine weitere wichtige Eigenschaft von 802.11n zertifizierten Endgeräten ist die vorgeschriebene Implementierung der in 802.11e spezifizierten Quality of Service (QoS)-Erweiterungen für die Luftschnittstelle. Mit dieser Erweiterung ist es möglich, dass Applikationen wie Voice over IP bevorzugt behandelt werden. Somit können Telefoniekopakete oder Daten von anderen Applikationen, die eine konstante Bandbreite benötigen, auch in Perioden mit hoher Netzwerklast (Streaming oder Übertragung von großen Dateien) deterministisch und zur richtigen Zeit übertragen werden. Mehr zu diesem Thema in Abschn. 4.8.

Beacon Frames von 802.11n Access Points enthalten eine Anzahl neuer Parameter. Der erste nennt sich „HT Capabilities“ (Element ID 45) und beschreibt, welche High

Throughput-Funktionen der Access Point unterstützt. Die folgende Liste gibt einen Überblick über die wichtigsten Funktionen:

- Unterstützung des 40 MHz-Modus (ja/nein).
- Anzahl der gleichzeitig unterstützten MIMO Streams und mögliche Modulations- und Kodiermodi (MCS).
- Unterstützung der auf 400 ns verkürzten Guard Time.
- Ob der optionale MCS Feedback-Modus unterstützt wird. Mit diesem kann der Empfänger dem Sender eine Rückmeldung über die zu verwendende Modulation geben und somit die Datenrate optimal an die Übertragungsbedingungen anpassen.
- STBC Diversity Support (siehe unten).
- Power Save Multipoll Support (PSMP), eine verbesserte Stromsparoption.
- Zahlreiche Parameter für das optionale MIMO Beamforming (siehe unten).
- Zahlreiche Parameter für die optionale Unterstützung diverser dynamischer Antennenauswahlverfahren. (siehe unten).

Der zweite neue Parameter in Beacon Frames ist der ‚HT Information‘ Parameter (Element ID 61). In diesem teilt der Access Point den Endgeräten mit, welche HT-Funktionalitäten aktuell verwendet werden dürfen und welche nicht. In der nachfolgenden Liste sind die wichtigsten Informationen zusammengefasst.

- Ob aktuell ein 40 MHz-Kanal verwendet werden darf oder ob Übertragungen auf den 20 MHz-Primärkanal limitiert sind.
- Operating Mode: Greenfield, HT-Mixed, Non-Member Protection Mode (Endgeräte, die mit anderen Access Points kommunizieren, senden im gleichen Band).
- Ob es Endgeräte im Netzwerk gibt, die nicht Greenfield Mode kompatibel sind.
- Overlapping BSS Protection: Entdeckt der Access Point Beacon Frames von anderen Access Points im gleichen Frequenzband, die nicht HT-fähig sind oder im Mixed Mode arbeiten, kann mit diesem Bit Endgeräten signalisiert werden, ebenfalls den HT-mixed Mode zu aktivieren. Benachbarte Access Points, die dieses Bit sehen, selber jedoch keine nicht-HT-Endgeräte beobachten können, müssen keine Sicherungsmaßnahmen treffen. Auf diese Weise wird erreicht, dass HT-Netzwerke auf nicht kompatible Netzwerke in der Nähe Rücksicht nehmen, sich dies aber nicht über deren Grenzen hinaus fortsetzt.
- Secondary Beacon: Gibt an, ob dieses Beacon-Paket im primären 20 MHz-Kanal eines 40 MHz-Kanals gesendet wurde, oder im zweiten 20 MHz-Kanal.

Zusätzlich zu den Beacon Frames werden die HT Capability und HT Information Parameter von Access Points auch in Association-, Reassociation- und Probe Response Frames gesendet. Endgeräte erhalten somit auch während der Anmeldung und beim Wechsel des Access Points noch einmal zusätzlich alle unterstützten Parameter und die aktuelle Konfiguration.

Außer HT Parameter müssen 802.11n kompatible Access Points auch Informationen für das in der 802.11e-Erweiterung spezifizierte Quality of Service Handling in den Beacon Frames übertragen. Weitere Details hierzu in Abschn. 4.8.

Damit der Access Point auch über die Fähigkeiten jedes einzelnen Endgerätes im Netzwerk Bescheid weiß, sendet auch ein Endgerät während der Association-Prozedur seine „HT Information“ an den Access Point. Somit ist es dann möglich, dass der Access Point für ein Endgerät Daten in einem 40 MHz-Kanal mit kurzen Guard-Intervall und zwei MIMO Streams überträgt, während Daten für ein Gerät mit weniger Fähigkeiten automatisch im 20 MHz-Kanal, mit 800 ns Guard-Intervall und ohne MIMO geschickt werden.

Aufgrund der nötigen Rückwärtskompatibilität zu 802.11b, g und a, sowie den vielen Optionen der 802.11n Erweiterung muss ein Endgerät vor der Übertragung eines Datenpaketes aus zahlreichen Optionen wählen. Wird ein Datenpaket an ein 802.11b Endgerät geschickt, kommt die HR/DSSS Modulation zum Einsatz. In Abhängigkeit des Übertragungskanals muss dann noch eine entsprechende Coderate gewählt werden. Für 802.11g-Endgeräte wird für die Übertragung eine OFDM Modulation mit weniger Subkanälen (Non-HT Format) als für 802.11n Endgeräte verwendet, sowie ein 802.11g PLCP Header. Bei Übertragungen zwischen zwei 802.11n-Geräten kommt ebenfalls die OFDM Modulation zum Einsatz, der PLCP Header ist jedoch kürzer und enthält HT-spezifische Informationen (HT Greenfield Mode). Sind im Netzwerk 802.11n und 802.11g-Geräte angemeldet (HT-Mixed Mode), wird, wie in Abb. 4.18 gezeigt, ein entsprechend rückwärtskompatibler PLCP Header gesendet. Dieser kann auch von

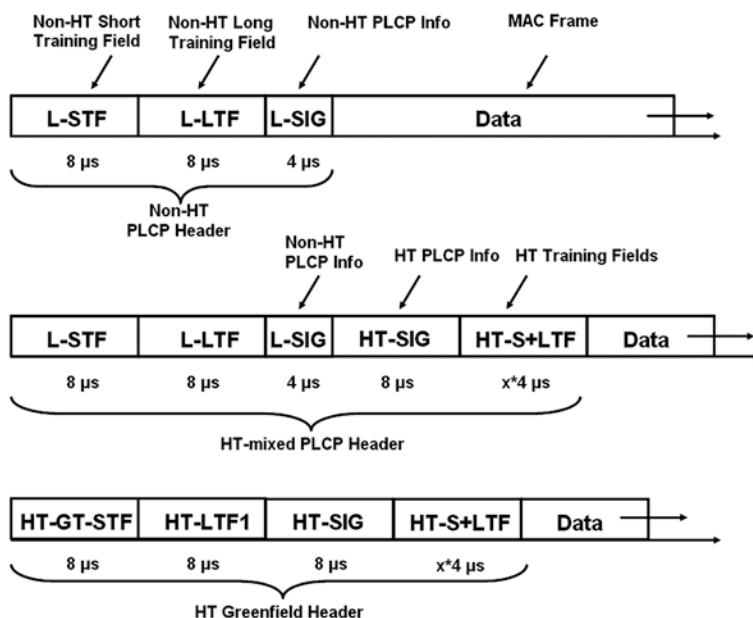


Abb. 4.18 PLCP Headervarianten

802.11g-Endgeräten dekodiert werden, umfasst jedoch einige Bytes mehr. Außerdem werden weniger OFDM-Subkanäle verwendet. Falls auch 802.11b Endgeräte vorhanden sind, muss zudem noch ein CTS-Paket in HR/DSSS Modulation der eigentlichen Übertragung vorangehen. Des Weiteren muss ein 802.11n Endgerät wissen, welche 802.11n-Funktionalitäten die Gegenstelle unterstützt. Dies ist notwendig, um die OFDM-Modulation entsprechend zu steuern (z. B. kurzes Guard-Intervall), die Wahl zwischen einem 20 oder 40 MHz-Kanal zu treffen, sowie die Anzahl der MIMO-Kanäle und die Codierrate in Abhängigkeit der Kanalqualität zu bestimmen.

Selbst diese umfangreiche Liste berücksichtigt noch nicht zahlreiche weitere optionale 802.11n-Funktionen, die nachfolgend beschrieben werden. Da viele dieser Funktionen optional sind, bleibt es in der Praxis dem Endgerät überlassen, welche davon implementiert und auch verwendet werden.

Für batteriebetriebene Endgeräte ist es sehr wichtig, dass der Wireless LAN Chip in Zeiten, in denen keine Daten übertragen werden, nur minimale Energie benötigt. Für diesen Zweck gibt es den in Abschn. 4.4 vorgestellten Power Save (PS) Mode, der heute auch von vielen Endgeräten verwendet wird. Dieser Power Save Mode kann aber nicht aktiviert werden, wenn Multimedia-Anwendungen wie Voice over IP z. B. alle 20 ms ein kleines Datenpaket von wenigen Mikrosekunden übertragen und dann für den Rest des Intervalls keine Daten übertragen. Auch wenn keine Daten übertragen werden, benötigt der WLAN Chip trotzdem Energie, da der Funkkanal weiterhin abgehört werden muss. Für solche Anwendungen wurde im 802.11n Standard optional ein zusätzlicher Stromsparmechanismus eingeführt, der Power Save Multi Poll (PSMP) genannt wird. Bei diesem Verfahren beantragt ein Endgerät beim Access Point, periodisch Datenpakete einer bestimmten Größe senden und empfangen zu dürfen. Der Access Point setzt daraufhin ein PSMP-Fenster auf und teilt dem Endgerät mit, zu welchen Zeiten dieses Fenster genutzt werden kann. Das Endgerät schaltet seinen Transceiver dann nur während dieses Fensters ein und empfängt seine Datenpakete. Nach dem Downlink Fenster folgt automatisch ein Uplink Fenster, in dem ein Endgerät ohne vorherige Reservierung des Mediums seine Daten schicken kann. Während der restlichen Zeit kann das Endgerät dann seinen Transceiver komplett abschalten und somit die Batterielaufzeit erhöhen.

Datenpakete in beide Richtungen enthalten im PSMP Modus nicht nur Nutzdaten, sondern auch Acknowledgement Informationen für die jeweils zuletzt empfangenen Datenpakete. Während eines PSMP-Fensters kann ein Endgerät mehrere Datenpakete senden bzw. empfangen. Werden diese einzeln verschickt, muss zwischen den Datenpaketen eine SIFS Pause eingelegt werden oder optional eine kürzere Sendepause, die RIFS (Reduced Inter Frame Space) genannt wird. Datenpakete können auch mit dem weiter oben beschriebenen Frame Aggregation-Verfahren in einem Physical Frame gebündelt werden.

Wie in Abb. 4.19 gezeigt, kann ein PSMP-Fenster auch von mehreren Endgeräten geteilt werden. Ein PSMP Frame am Anfang des Intervalls enthält Informationen für alle Endgeräte, zu welchen Zeiten jedes einzelne Endgerät im PSMP-Fenster Daten

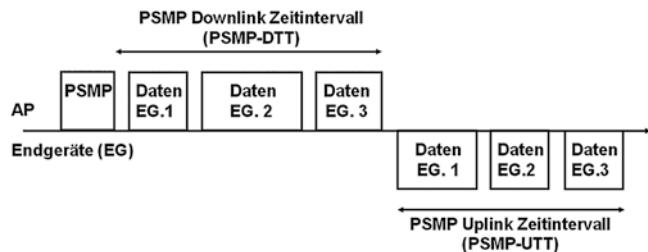


Abb. 4.19 Ein Power Save Multi Poll-Fenster (PSMP), in dem mehrere Endgeräte senden und empfangen

empfangen und senden darf. Der Standard gibt vor, dass PSMP-Fenster alle 5 bis 40 ms eingelegt werden sollen, mit einer Granularität von 5 ms. Für Voice over IP ist z. B. ein Intervall von 20 ms interessant, da Sprachcodecs üblicherweise Sprachinformationen über diesen Zeitraum komprimieren und dann in einem kleinen Paket übertragen.

Die PSMP-Fenster und die für jedes Endgerät vorhandenen Übertragungszeiten sind für eine kontinuierliche und gleich bleibende Übertragung gedacht und so optimiert, dass bei konstanter Nutzung möglichst wenig Bandbreite ungenutzt bleibt. Nun kann es jedoch sein, dass ein Endgerät kurzzeitig mehr Bandbreite benötigt oder ein Paket aufgrund eines Übertragungsfehlers erneut übertragen werden muss. Dies kann dann nicht im normalen Zeitfenster geschehen, da für solche zusätzlichen Übertragungen kein Platz vorhanden ist. Für zusätzliche Uplink-Kapazität kann das Endgerät deshalb dem Access Point über ein Flag im MAC Header mitteilen, dass zusätzliche Bandbreite benötigt wird. Der Access Point hat dann die Möglichkeit, an das nächste PSMP-Fenster ein weiteres PSMP-Fenster direkt anzuhängen und teilt dies im PSMP Frame, das jedem PSMP-Fenster voransteht, den Endgeräten entsprechend mit. Tritt ein Übertragungsfehler in Uplink Richtung auf, signalisiert dies der Access Point dem Endgerät durch ein negatives Acknowledgement im nächsten PSMP Downlink-Abschnitt und fügt ebenfalls ein PSMP-Fenster an.

Eine weitere Stromsparfunktion wurde im 802.11n Standard für MIMO Spatial Multiplexing (SM)-fähige Endgeräte eingeführt. Auch wenn keine Daten übertragen werden, müssen diese im Standardmodus ständig mehrere Empfänger bereithalten, da der Access Point ihnen ja zu jeder Zeit ein Paket schicken kann. Um die Stromaufnahme für batteriebetriebene Geräte zu reduzieren, wurden zwei optionale MIMO SM Power Save Modi spezifiziert. Im statischen Modus signalisiert ein Endgerät einem Access Point über eine „SM Power Save Management Action Frame“-Nachricht, wenn es den SM Power Save Modus an- oder abschaltet. Zusätzlich gibt es auch SM Power Save Bits im HT Capabilities Parameter, die ein Endgerät während der Association-Prozedur verwenden kann, um dem Access Point mitzuteilen, dass es aktuell nur Single Stream-Übertragungen zulässt. Des Weiteren gibt es auch einen dynamischen SM Power Save Modus. Hier schaltet das Endgerät alle zusätzlichen Empfänger ab und arbeitet im Single

Stream-Modus. Das Endgerät aktiviert seinen MIMO Modus wieder automatisch, sobald der Access Point das Endgerät mit einem Paket wie z. B. eine RTS/CTS-Sequenz im Single Stream-Modus adressiert. Alle nachfolgenden Frames schickt der Access Point ohne weitere Vereinbarung dann mit mehreren MIMO Streams.

MIMO Spatial Multiplexing steigert zwar die Datenrate, nicht jedoch die Reichweite eines Netzwerkes. Deshalb gibt es im Standard optional weitere Möglichkeiten, die zusätzlichen Sende- und Empfangseinheiten (Transceiver) statt für erhöhten Durchsatz für eine bessere Reichweite zu nutzen.

Eines dieser Verfahren ist das MIMO Beamforming. Hier wird über alle Transceiver der gleiche Datenstrom gesendet. Durch geschickte Kombination der Sendeleistung und zeitlichen Versatz der Datenströme kann jedoch eine Richtwirkung erzeugt werden. Somit wird die gesamte Übertragungsleistung nicht gleichmäßig im Raum verteilt, sondern gezielt in der Umgebung des Empfängers konzentriert. Damit Beamforming funktioniert, benötigt der Sender Rückmeldungen vom Empfänger, um den Strahl (Beam) in die richtige Richtung zu dirigieren. Somit müssen sowohl der Sender als auch der Empfänger MIMO Beamforming unterstützen.

Statt Beamforming kann die Reichweite eines Netzwerkes auch mit einem Verfahren gesteigert werden, das ein mathematisches Verfahren namens Space Time Block Code (STBC) nutzt. Unterstützen Sender und Empfänger diesen Modus, wird z. B. in einer 2×2 MIMO Konfiguration auch hier ein einzelner Datenstrom getrennt über zwei Pfade übertragen. STBC kodiert jedoch den Datenstrom für jeden Transmitter unterschiedlich und in einer Weise, dass diese zueinander orthogonal sind. Auf der Empfängerseite erhöht dies den Signal- zu Rauschabstand, was wiederum dabei hilft, die Signalstärke und damit den Durchsatz bei weiter entfernten Endgeräten zu steigern.

Unterstützt eine Gegenstelle keine der optionalen MIMO-Funktionalitäten, gibt es für einen Empfänger noch andere optionale Möglichkeiten, die Signalqualität zu steigern. Hat das Endgerät mehrere Antennen, kann es untersuchen, mit welcher Antenne Daten am besten empfangen werden, und verwendet dann diese. In der Praxis kann dies durchaus eine deutliche Signalverbesserung für weiter entfernte Endgeräte bedeuten. Dies lässt sich anschaulich bei Endgeräten mit nur einer Antenne und schlechten Empfangsbedingungen nachvollziehen. Hier reicht bei schlechten Empfangsbedingungen oft schon das manuelle Versetzen der Antenne um wenige Zentimeter, um den Empfang zu verbessern. Diese Funktionalität ist nicht 802.11n-spezifisch sondern wird auch schon bei 802.11g Access Points eingesetzt, die mehrere Antennen haben. Ein etwas aufwendigeres Verfahren ist das Maximum Ratio Combining (MRC). Hier untersucht der Empfänger den eingehenden Datenstrom auf mehreren Receivern und kombiniert die zwei getrennt empfangenen Signale, um so den Signal- zu Rauschabstand zu verbessern.

Für Endgeräte, die sich näher am Access Point befinden, ist natürlich das zuerst beschriebene MIMO Spatial Multiplexing das Mittel der Wahl, um die Übertragungsgeschwindigkeit zu steigern. Die Transceiver werden dann genutzt, um mehrere Datenströme parallel zu übertragen. Bei weniger günstigen Übertragungsbedingungen sind Beamforming und STBC das bessere Mittel, so sie denn von Sender und Empfänger

unterstützt werden. Die damit erreichbaren Datenraten sind natürlich geringer als mit MIMO Spatial Multiplexing, da nur ein Datenstrom verwendet wird. Welches der Verfahren für eine Übertragung angewandt wird, muss der Sender selbständig anhand der Übertragungssituation entscheiden, sowie mit dem Wissen, welche MIMO-Arten die Gegenstelle unterstützt. Bei mehreren Endgeräten im Netzwerk können alle Verfahren nebeneinander koexistieren. Ein Gerät mit guten Empfangsbedingungen wird dann vom Access Point mit Spatial Multiplexing bedient, während das nächste Paket an ein weiter entferntes Gerät mit STBC kodiert wird.

Eine weitere optionale Funktionalität, die in der 802.11n-Arbeitsgruppe definiert wurde, ist das Modulation and Coding Scheme (MCS) Feedback. Ohne dieses Verfahren müssen Sender z. B. anhand eingehender ACKs oder NACKs entscheiden, welche Modulation und Kodierung sie für die Übertragung zu einem Endgerät verwenden. In der Praxis kann beobachtet werden, dass ein Endgerät ein nicht korrekt empfangenes Paket zunächst mehrmals mit dem ursprünglich gewählten MCS versendet. Ist dies nicht erfolgreich, wird das Paket dann stufenweise mit einem kleineren MCS versendet bis das Paket erfolgreich empfangen wurde. Dies ist natürlich nicht optimal und führt dazu, dass ein MCS verwendet wird, der die Empfangsbedingungen nicht optimal ausnutzt und Daten zu langsam übertragen oder mehrmals wiederholt werden müssen. Mit MCS Feedback wurde eine Möglichkeit geschaffen, dass ein Sender von einem Empfänger Feedback über seine Empfangsbedingungen anfordern kann. Der Empfänger liefert dann Informationen im MAC Header in darauf folgenden Übertragungen implizit zurück.

4.6.5 IEEE 802.11ac mit bis zu 6,8 Gbit/s

In der Praxis wurden in 802.11n Geräten nur die Grundfunktionen des Standards wie mehrere MIMO Kanäle, 40 MHz Bandbreite und die Nutzung des 5 GHz Bandes implementiert. Weitere Funktionen wurden dann in der Praxis erst in Geräten umgesetzt, die die 802.11ac Spezifikation unterstützen. Dieser ist der direkte Nachfolger von 802.11n und erste Geräte, die den Very High Throughput (VHT) Physical Layer (PHY) unterstützten, kamen in 2013 auf den Markt. Wie auch bei den vorangegangenen Versionen des Standards wurde in den ersten Endgeräten nur ein kleiner Teil der neuen Funktionen verwirklicht (802.11ac Wave 1). Diese hatten das Ziel, die maximal mögliche Übertragungsrate bei guten Empfangsbedingungen weiter zu steigern. Ab 2016 kamen auch erste Endgeräte mit 802.11ac Wave 2 Unterstützung auf den Markt, die durch Unterstützung von Multi-User MIMO (MU-MIMO) zwar nicht die Übertragungsgeschwindigkeit eines einzelnen Endgerätes weiter steigern, jedoch den Gesamtdurchsatz im Netzwerk erhöhen können, sofern auch der Access Point und andere Geräte im Netzwerk diese Funktion unterstützen.

Die einfachste Art die theoretisch möglichen Datenraten weiter zu erhöhen ist, das Channel Bonding von 2×20 MHz Kanälen von 802.11n zu erweitern. Umgesetzt wurde zunächst in 802.11ac Wave 1 Geräten ein 80 MHz Kanal, der die mit 802.11n

erreichbaren Datenraten verdoppelt. Die Spezifikation enthält auch eine Option, zwei 80 MHz Kanäle zu einem 160 MHz breiten Kanal zu kombinieren. Eine weitere Option sieht vor, zwei 80 MHz Kanäle an unterschiedlichen Stellen im Frequenzband zu verwenden, um reservierte Regionen umgehen zu können. Dies ist nötig, da in manchen Ländern Anwendungen wie das Wetterradar das 5 GHz Band in zwei Teile spaltet. Einen Überblick, welche Teile des 5 GHz Bandes in unterschiedlichen Ländern verwendet werden kann, findet sich auf Wikipedia¹¹. Da nur das 5 GHz Band genug Platz bietet (insgesamt 350 MHz), ist die 802.11ac Spezifikation nur für dieses Band ausgelegt.

In den meisten Ländern der Welt sind ungefähr 400 MHz im 5 GHz Band für Wi-Fi zugänglich. Dies erlaubt 4 bis 5 nicht überlappende 80 MHz Kanäle. Wie auch bei den früheren PHYs ist es auch hier möglich, mehrere Netzwerke unabhängig auf dem gleichen Kanal zu betreiben und die Bandbreite somit zu teilen. Wenn mehrere Netzwerke auf dem gleichen Kanal betrieben werden, können Endgeräte die Übertragungen von Endgeräten des anderen Netzwerkes, die sich in seiner Nähe befinden, detektieren. Der bereits vorgestellte Collision Avoidance Mechanismus funktioniert also auch in diesem Fall. In der Praxis überlappt sich die geografische Abdeckung von Netzwerken nicht immer vollständig. Das bedeutet, dass manche Geräte in einem Netzwerk die Übertragungen von Geräten in einem anderen Netzwerk auf der gleichen Frequenz nicht bemerken. In diesem Fall ist die Collision Avoidance nicht ganz so effektiv, wie wenn nur ein Netzwerk auf einem Kanal sendet.

Um die Zusammenarbeit zwischen Netzwerken auf dem gleichen Kanal zu verbessern, wurde eine Anzahl an Funktionalitäten mit 802.11ac eingeführt. Wie zuvor schon beschrieben, können Endgeräte erkennen, wenn ein anderes Endgerät Daten überträgt und können somit mit ihrer eigenen Übertragung warten. Dies ist als Collision Avoidance bekannt. In der Praxis wird dies mit zwei Methoden erreicht, durch Erkennen eines Signals und durch Erkennen einer Sendeleistung auf dem Kanal. Erkennen eines Signals bedeutet, dass ein Endgerät den Beginn eines Frames empfangen kann und somit durch Dekodieren des Headers herausfinden kann, wie lange der Kanal belegt sein wird. Erkennen der Sendeleistung andererseits bedeutet, dass das Endgerät nur einen Signalegel auf dem Kanal detektieren kann, die Übertragung jedoch nicht dekodieren kann.

Um abwärtskompatibel zu bleiben, teilt 802.11ac die Kanalbandbreite in 20 MHz Stücke (Chunks) auf. Ein 80 MHz Kanal hat somit 4 Chunks. Wird der RTS/CTS Mechanismus verwendet, um den Kanal zu reservieren, werden entsprechende Pakete auf allen Chunks gleichzeitig gesendet. Dies erlaubt 802.11n und 11a Geräte, die Reservierung des Kanals zu erkennen, obwohl diese nicht in der Lage sind, einen 80 MHz Kanal zu verwenden.

Dieser Mechanismus kann auch verwendet werden um die Kanalbreite für eine Übertragung zu reduzieren, wenn z. B. das CTS nur auf zwei 20 MHz Chunks statt auf allen vier empfangen wird, weil ein Nachbarnetzwerk Interferenz auf den anderen Chunks erzeugt. Diese Entscheidung kann für jedes Paket von neuem getroffen werden, was das System sehr flexibel macht.

In der Praxis kann es vorkommen, dass nicht alle Endgeräte die Übertragung in einem 80 MHz Kanal unterstützen. Darum wurde spezifiziert, dass mehrere Netzwerke, die den gleichen 80 MHz breiten Kanal verwenden, sich untereinander koordinieren können, um eine gleichzeitige Übertragung in den unterschiedlichen Netzwerken zu ermöglichen. Dies wird durch die Aufteilung des Kanals in einen primären Kanal (Primary Channel) und einen nicht-primären Kanal (Non-Primary Channel) erreicht. Die Kanalbandbreite ist dabei die Hälfte der vollen Kanalbandbreite, also 40 MHz in einem 80 MHz Netzwerk. Zwei sich überlappende 80 MHz Netzwerke konfigurieren somit ihren primären Kanal auf den jeweils anderen Teil des ganzen Kanals. Wenn zwei 40 MHz Endgeräte, die zu zwei unterschiedlichen 80 MHz Netzwerken gehören, gleichzeitig senden wollen, verwenden sie jeweils den primären Kanal des eigenen Netzwerk, der sich nicht mit dem primären Kanal des anderen Netzwerks überlappt. Jedes Endgerät findet einen freien Kanal vor und es kommt zu keiner Kollision. Wenn schon eine Übertragung im nicht-primären Kanal läuft, kann ein Endgerät entweder auf die nächste Übertragungsmöglichkeit warten oder seine Daten nur auf dem primären Kanal übertragen. Auch diese dynamische Bandbreitennutzung trägt dazu bei, dass Endgeräte, die nicht die volle Kanalbandbreite unterstützen, trotzdem das vorhandene Spektrum möglichst gut ausnutzen können. Außerdem optimiert diese Maßnahme die Bandbreitennutzung, wenn 802.11ac Netzwerke im gleichen Spektrum betrieben werden wie ältere 802.11a und 802.11n Access Points, die nur 20 MHz oder 40 MHz Kanäle verwenden.

Da das 2,4 GHz Band viel zu schmal für 80 oder 160 MHz Kanäle ist, wurde der 802.11ac Standard nur für den 5 GHz Bereich spezifiziert. In der Praxis unterstützen 802.11ac kompatible Chipsätze natürlich auch die 802.11b, g und n Übertragungsmodi im 2,4 GHz Band und sind somit abwärtskompatibel und universell einsetzbar. Zusätzlich zu den breiteren Bändern wurde auch ein neues Modulationsverfahren spezifiziert, das bei besonders guten Kanalbedingungen verwendet werden kann. Während 802.11n die Übertragung von bis zu 6 Bits pro Übertragungsschritt mit der 64 QAM Modulation erlaubt, wurde mit 802.11n nun auch die Übertragung von 8 Bits pro Übertragungsschritt, also 256 QAM spezifiziert. Zusammen mit einer Kodierung von 5/6, (5 Nutzdatenbits, 1 Fehlerkorrekturbit) benötigt 256 QAM einen um 5 db besseren Kanal als 64 QAM. Das bedeutet, dass weniger Rauschen auf dem Kanal vorhanden sein darf und dass mehr Signalenergie empfangen werden muss. Dies kann teilweise durch immer bessere Empfängerbaugruppen erreicht werden. Durch die zusätzlichen zwei Bits pro Übertragungsschritt kann die maximale Datenrate um 30 % gegenüber einer 64 QAM Übertragung gesteigert werden. Die folgende Tabelle gibt einen Überblick über die Modulations- und Kodierraten, die für 802.11ac spezifiziert wurden. MCS 0 ist z. B. die Kombination aus einer sehr konservativen Modulationsart (BPSK), die nur ein Bit pro Übertragungsschritt sendet und einer Kodierrate von 1/2, d. h. für jedes Datenbit wird auch ein zusätzliches Fehlerkorrekturbit übertragen.

MCS	Modulation	Kodierrate
0	BPSK	1/2
1	QPSK	1/2
2	16QAM	1/2
3	16QAM	1/2
4	64QAM	3/4
5	64QAM	2/3
6	64QAM	3/4
7	64QAM	5/6
8	256QAM	3/4
9	256QAM	5/6

802.11ac verwendet das auch schon von 802.11n bekannte kurze Guardintervall von 400 ns zwischen OFDM Symbolen, da sich dies in der Praxis bewährt hat. Dies ermöglicht einen Geschwindigkeitsvorteil gegenüber 802.11a und 11g Netzwerken von etwa 10 %. Eine weitere Möglichkeit, die theoretisch maximal Datenrate zu steigern ist, die Anzahl der MIMO Datenströme weiter zu erhöhen. Der 802.11ac Standard unterstützt bis zu 8×8 MIMO, verdoppelt also die Datenrate des auf 4×4 MIMO begrenzten 802.11n Standards. In der Praxis ist es jedoch auch heute noch schwierig, mehr als drei bis vier Antennen in Endgeräten unterzubringen und die Kombination aus 4×4 MIMO und 256 QAM benötigt einen noch robusteren Kanal. Es hat sich gezeigt, dass in den meisten Anwendungsszenarien eine heute übliche 3×3 MIMO Übertragung mit 64 QAM nur unwesentlich mehr Durchsatz ermöglicht, als eine 2×2 MIMO- oder sogar eine Single Stream Übertragung. Weitere MIMO Pfade könnten jedoch nützlich sein, wenn sie in Kombination mit Beamforming verwendet werden, um mehrere Endgeräte gleichzeitig anzusprechen.

Beamforming ist eine weitere Option des 802.11ac Standards, um die Signalenergie in die Richtung eines Endgerätes zu fokussieren. Damit dies funktioniert, muss der Access Point wissen, in welche Richtung ein Signal konzentriert werden soll. Dazu sendet der Access Point sogenannte Channel Sounding Announcements, auch Null Data Packet (NDP) Announcements genannt. Der Name basiert auf der Channel Sounding Methode, die leere Daten Frames nutzt, deren OFDM Symbole dann vom Empfänger auf Veränderungen analysiert werden, die während der Übertragung auftraten. Zu diesem Zweck sendet der Access Point NDP Announcement Nachrichten, um Endgeräte, die Beamforming unterstützen, aufzufordern, sich zu melden. In einem zweiten Schritt werden dann NDP Pakete vom Access Point gesendet und von einem Endgerät empfangen. Dieses analysiert dann die OFDM Symbole der Pakete und berechnet aufgrund der festgestellten Veränderungen eine Antwort die beschreibt, wie die OFDM Symbole während der Übertragung verändert wurden. Aufgrund dieser Rückmeldung kann der Access Point dann eine Steuerungsmatrix (Steering Matrix) für jedes Endgerät

berechnen, die dann auf die Datenpakete vor der Übertragung angewandt wird. Die Steering Matrix beschreibt dabei, wie ein Signal über individuelle Sendepfade/Antennen verteilt werden muss und welche Phasenverschiebung (Verzögerung) jeweils anzuwenden ist. Auf diese Weise wird erreicht, dass die Interferenz der unterschiedlichen Wellenfronten das Signal an einem Ort verstärken, während das Signal anderen Orten, an denen sich das Endgerät nicht befindet, abgeschwächt wird. Um den sich ständig ändernden Signalbedingungen entgegenzuwirken, muss die Sounding Prozedur mehrere Male pro Sekunde durchgeführt werden. Ein typischer Wert ist alle 100 ms¹². Da das Signal an einem Ort verstärkt wird, ist es aufgrund der Vorschriften zur maximalen Sendeleistung notwendig, das Ausgangssignal des Access Points während einer solchen Übertragung um 3 dB abzuschwächen. Es bleibt abzuwarten, welchen Einfluss diese Leistungsreduktion auf die Effektivität des Beamformings haben wird.

Beamforming kann auch mit MIMO kombiniert werden, um mehrere Datenströme zu einem Gerät zu senden. Dies wird als Single-User (SU) Beamforming bezeichnet. In der Praxis ist heute zu beobachten, dass die Kombination von MIMO und Beamforming verwendet wird. Ein noch anspruchsvollereres Verfahren ist die gleichzeitige Übertragung von einem oder mehreren Datenströmen an mehrere Endgeräte. Dies wird als Multi-User (MU) Beamforming bezeichnet. Bis zu vier Endgeräte können gleichzeitig angesprochen werden, was vor allem dann interessant ist, wenn der Access Point mehr MIMO Antennen zur Verfügung hat als ein Endgerät. Das könnte z. B. der Fall sein, wenn der Access Point 4 MIMO Streams unterstützt, während z. B. batteriebetriebene Endgeräte, um Strom zu sparen, nur einen MIMO Stream nutzten wollen. Durch den Versand von mehreren Datenströmen an mehrere Endgeräte kann so eventuell der Übertragungskanal besser ausgenutzt werden und ein höherer Gesamtdurchsatz erreicht werden. Die Präambel eines solchen Multi-User Frames enthält dazu die Information, an welche Endgeräte sich dieser richtet. Danach enthält der Multi-User Frame die Datenströme für jeden User, die sozusagen aufeinandergestapelt werden, also in unterschiedlichen MIMO Strömen gleichzeitig übertragen werden. Beamforming wird verwendet, um die individuellen Datenströme jeweils in die Richtung der unterschiedlichen Endgeräte zu senden. Auf diese Weise sieht ein Endgerät nur seinen eigenen Datenstrom und es entsteht weniger Interferenz durch die Datenströme, die gleichzeitig an andere Geräte geschickt werden, da diese in „andere Richtungen“ gesendet werden. Dazu ist es natürlich notwendig, dass die unterschiedlichen Endgeräte sich auch an unterscheidbaren Orten befinden. Da sich die Datenmenge und auch die Modulation und die Kodierung der unterschiedlichen Ströme unterscheiden können, kann es vorkommen, dass Teile mancher Ströme leer sind und mit Padding Bits versehen werden. Eine Frage, die sich hier stellt, ist, wie die Empfänger eines Multi-User Frames den korrekten Empfang bestätigen können, da normale Acknowledgement Frames direkt auf die Übertragung eines Datenpaketes gesendet werden müssen. Dies ist bei Multi-User Frames nicht möglich, da nur ein Acknowledgement Frame zu einer Zeit gesendet werden kann. Darum

müssen die Acknowledgement Frames nacheinander gesendet werden. Dazu wird der Deferred Block Acknowledgement Mechanismus, verwendet, der schon in 802.11n eingeführt wurde. Dort wurde er ursprünglich verwendet, um die Übertragung von mehreren Datenblöcken auf einmal und mit nur einer Nachricht zu bestätigen. Beim Multi-User Beamforming wird nur ein einziger Multi-User Frame gesendet und der Delayed Block ACK Mechanismus dient in diesem Fall dazu, die Antworten der unterschiedlichen Endgeräte über die Zeitachse zu verteilen.

Alle Funktionen gemeinsam ergeben eine theoretische Spitzendatenrate von 6.93 Gbit/s. Dazu wäre ein 160 MHz Kanal notwendig, 8 MIMO Ströme, 256 QAM Modulation, sowie in kurzes Guard Intervall zwischen den Paketen. In der Praxis ist dies jedoch offensichtlich schwer zu erreichen. Aktuell unterstützen in der Praxis 802.11ac kompatible Access Points 3 bis 4 MIMO Ströme über einen 80 MHz Kanal mit einer theoretischen Spitzendatenrate von 1,3, bzw. 1,7 Gbit/s. Viele Hersteller verwenden in der Praxis diese Werte, um eine Aussage über die Leistungsfähigkeit ihrer Produkte zu machen und addieren zusätzlich noch die theoretisch maximal mögliche Datenrate zu diesem Wert, die mit dem 802.11n Standard im 2,4 GHz Band möglich ist. Des Weiteren bezeichnen manche Hersteller Geräte, die einen bis zu 80 MHz breiten Kanal unterstützen als Wave 1 Endgeräte. Wave 2 Endgeräte hingegen unterstützen eine Kanalbandbreite bis zu 160 MHz.

Auf dem IP Layer sind die erreichbaren Datenraten jedoch auch unter optimalen Bedingungen wesentlich niedriger. Die folgende Tabelle gibt einen Überblick über typische Geschwindigkeiten, die mit Messungen in der Praxis ermittelt wurden. Während diese Datenraten wesentlich höher als noch bei 802.11n sind, bleiben diese jedoch weit hinter den theoretischen Spitzendatenraten zurück.

Aktuell verwenden nicht alle 802.11ac fähigen Endgeräte und Access Points das gesamte 5 GHz Band. Manche Modelle unterstützen nur 80 MHz am unteren Teil des Bandes (Kanalnummern 36–48), da sie keine dynamische Frequenzwahl (Dynamic Frequency Selection, DFS) unterstützen und keine anderen Nutzer des Bandes (z. B. Wetterradar) erkennen können. In Deutschland müssen beispielsweise Geräte DFS unterstützen, falls sie Kanäle oberhalb von Kanal 48 verwenden wollen. Wird ein höherer Kanal verwendet, muss der AP den Kanal auf Signale von „höher priorisierten Nutzern“ in einem Zeitraum von 10 min untersuchen und darf den Kanal erst danach nutzen. Während dieser Zeit, z. B. nach Änderung der Konfiguration, ist der Kanal im 5 GHz Band nicht verfügbar. Manche APs umgehen diese Verzögerung, in dem sie den 80 MHz breiten Kanal 36–48 in diesem Zeitraum verwenden. Werden keine „höher priorisierten Nutzer“ gefunden, wird der Kanal dann entsprechend verschoben. Der AP kann dies den Endgeräten mit einer Channel Switch Operation Nachricht mitteilen, damit diese während dem Kanalwechsel das Netzwerk nicht verlieren.

Datenrate	Netzwerkeigenschaften und Umgebungsbedingungen
700 Mbit/s	2 × 2 MIMO, High-End Access Point, sowie ein High-End Wi-Fi Chipsatz im Notebook, sehr kurze Distanz ¹³
300–400 Mbit/s	2 × 2 MIMO, High-End Access Point und Wi-Fi Chipsatz im Notebook, kurze Distanz, ~5–8 m, 1 Wand, 80 MHz Kanal
100 Mbit/s	2 × 2 MIMO, High-End Access Point, 20 m Distanz zwischen AP und Endgerät, 80 MHz Kanal

4.6.6 IEEE 802.11ax – Wi-Fi 6 – High Efficiency Erweiterungen

Die Evolution von 802.11n nach 802.11ac brachte einen deutlichen Geschwindigkeitsgewinn für Heimnetze in der Praxis, der hauptsächlich der Erweiterung der Kanalbandbreite auf 80 MHz im 5 GHz Band, sowie effizienteren WLAN Chips geschuldet ist. Eine weitere Steigerung der Geschwindigkeit kann mit der Einführung von 160 MHz Kanälen erreicht werden, die von manchen Endgeräten unterstützt werden. Diese Erweiterungen sind auch für Netze nützlich, in denen sich viele Geräte befinden und viele Daten in beide Richtungen ausgetauscht werden. Beispiele hierfür sind stark frequentierte Orte wie Stadien, Flughäfen, Bahnhöfe und Messehallen. An diesen Orten installieren Netzbetreiber nicht nur LTE und 5G in vielen Bändern, sondern verwenden auch WLAN, um die nötige Kapazität bereitzustellen. Außerdem ist zu beobachten, dass sich die Netzwerknutzung in Bürogebäuden in den letzten Jahren stark verändert hat. Während in der Vergangenheit oftmals stationäre PCs und Ethernetverkabelung verwendet wurde, ist mit dem heutigen Desk-Sharing zunehmend Wi-Fi als Netztechnologie gefragt. Auch werden zunehmend Smartphones und Tablets im Firmenumfeld verwendet, die zusätzlichen Netzwerkverkehr erzeugen. Umgebungen mit sehr vielen Endgeräten und hohem Datenaufkommen ist der Fokus des 802.11ax Standards, der von der Wi-Fi Association auch als Wi-Fi 6 bezeichnet wird. Wie auch schon 802.11ac (Wi-Fi 5), ist dieser Standard in zwei Wellen (Waves) aufgeteilt, um die wichtigsten Funktionen als erstes zu implementieren. Die nachfolgende Tabelle gibt einen Überblick über die wichtigsten Wi-Fi 6 Funktionen, die nachfolgend dann genauer beschrieben werden.

802.11ax Funktion	Beschreibung
1024-QAM Modulation	25 % Geschwindigkeitssteigerung gegenüber 802.11ac mit 256-QAM Modulation unter idealen Bedingungen
OFDMA	Orthogonal Frequency Division Multiple Access (OFDMA) im Downlink und Uplink. Daten können an mehrere Geräte gleichzeitig gesendet werden. Datenübertragung an ein Endgerät mit OFDM weiterhin möglich
Multi-User MIMO Verbesserungen	Verbesserungen im Downlink gegenüber 802.11ac. Neu: MU-MIMO im Uplink, Wave-2 Funktion

802.11ax Funktion	Beschreibung
Verbesserter Spatial Re-Use	BSS Coloring, damit Endgeräte zwischen Übertragungen von anderen Geräten an den eigenen und andere APs unterscheiden können. Abhängig vom Signalpegel werden nun gleichzeitige Übertragungen zu unterschiedlichen APs möglich
Bessere IoT Geräteunterstützung	20 MHz-only Modus wird nun zusammen mit einem neuen Sleep Mode unterstützt. Dieser erlaubt Endgeräten, den Transceiver für Stunden oder Tage zu deaktivieren

Während Wi-Fi 5 nur das 5 GHz Band adressiert, bringt 802.11ax auch Neuerungen für das 2,4 GHz Band. Zwar wird das 2,4 GHz Band heute neben WLAN auch sehr stark von anderen Anwendungen wie Bluetooth, drahtlosen Tastaturen, Mäusen, Baby Monitoren, etc., genutzt, es ist jedoch aufgrund der deutlich größeren Reichweite nach wie vor auch für WLAN interessant. Wie nachfolgend gezeigt wird, bringt 802.11ax viele Verbesserungen, um den Durchsatz von Wi-Fi Netzwerken auch in solch stark ausgelasteten Bereichen zu verbessern.

Eine Neuerung, die für alle Anwendungsfälle, inklusive Heimnetzwerke mit nur wenigen Teilnehmern nützlich ist, ist die Einführung der 1024-QAM Modulation. Diese erhöht die theoretische Höchstgeschwindigkeit gegenüber der bisher verwendeten 256-QAM Modulation bei Wi-Fi 5 um 25 %. Wie bei jeder Einführung einer höheren Modulation verringert sich der Radius, in dem diese Modulation verwendet werden kann, weiter, da der Signal- zu Rauschabstand größer als bisher sein muss. Mit 1024-QAM Modulation, einer Coding Rate von 5/6 und zwei MIMO Streams, die typischerweise Notebooks heute unterstützen, ergibt sich auf einem 80 MHz Kanal eine maximale Datenrate von 1201,0 Mbit/s. Zum Vergleich: Wi-Fi 5 mit 256-QAM Modulation erreicht hier nur 866,7 Mbit/s auf dem PHY Layer und etwa 600 Mbit/s auf dem IP Layer. Durch Multi-User MIMO kann der Gesamtdurchsatz des Access Points noch darüber hinausgehen, falls mehrere Geräte gleichzeitig Daten übertragen und die Kanaleigenschaften für die simultanen Übertragungen eine gute Trennung der Datenströme erlauben. Mit 4 MIMO Streams zu unterschiedlichen Geräten kann in einem 80 MHz Kanal eine theoretischen Spitzendatenrate von 2402,0 Mbit/s erreicht werden. Mit 8 MIMO Streams käme man so auf 4803,9 Mbit/s in einen 80 MHz Kanal. In einem 160 MHz Kanal wären bis zu 9607,8 Mbit/s auf dem PHY möglich. Dafür sind jedoch optimale Empfangsbedingungen für jedes Gerät nötig, sowie eine optimale Verteilung in der Fläche, damit die Streams unabhängig voneinander übertragen werden können. In der Praxis ist dies natürlich sehr unwahrscheinlich.

Eine typische Datenrate eines einzelnen Gerätes mit zwei Antennen und einem sehr guten 80 MHz breiten Übertragungskanal beträgt etwa 750 Mbit/s. Mit 160 MHz Kanälen kann dieser Wert nochmals verdoppelt werden. Hier werden dann jedoch die Ethernet Ports des Access Points zum Flaschenhals, da die heute typisch verwendeten 1 Gbit/s Ports nicht mehr ausreichen. High-End Wi-Fi 6 APs haben deswegen heute oft

zumindest einen 2,5 Gbit/s Ethernet Port. Bei anderen Geräten, wie z. B. Notebooks und teilweise auch bei Servern ist dies jedoch noch nicht üblich.

Einen deutlichen Unterschied in dieser Version des Standards ist die neue Symbol Rate und der Subcarrier Abstand. Für Wi-Fi 6 wurde die Symbol Rate von 3,6 µs auf 12,8 µs geändert. Daraus resultiert eine Änderung des Subcarrier Abstands von 312,5 kHz auf 78,125 kHz (vgl. den LTE Subcarrier Abstand von 15 kHz). Somit wird nun eine 1024 FFT Matrix für einen 80 MHz Kanal verwendet, sowie eine 2048 FFT Matrix für 160 MHz Kanäle. Rückwärtskompatibilität mit älteren Geräten wird gewährleistet, indem die bisherige Modulation und Kodierung für die Paket-Präambeln und CTS Kanalreservierungen verwendet werden. Auf diese Weise können ältere Geräte den Start einer Wi-Fi 6 Übertragung erkennen und der CSMA/CA Algorithmus hält dann die Übertragung neuer Pakete entsprechend an. Die nachfolgende Tabelle vergleicht die PHY Konfiguration von 802.11ax mit den entsprechenden Parametern von 802.11ac und LTE.

Parameter	802.11ac (Wi-Fi 5)	802.11ax (Wi-Fi 6)	LTE
Symbollänge	3,2 µs	12,8 µs	66,6 µs
Subcarrier Abstand	312,5 kHz	78,125 kHz	15 kHz
Anzahl Subcarrier in 80 MHz	208	936	4800
FFT Größe für einen 80 MHz Kanal	256	1024	8192 (Mit 4 Carrier Aggregation)
Höchste Modulation	256-QAM	1024-QAM	256-QAM
Sendeleistung AP/ Basestation für 80 MHz Bandbreite	0,2 W, Omni-direktional	0,2 W, Omni-direktional	100–200 W, sektorisiert
Sendeleistung der Endgeräte	0,2 W	0,2 W	0,2 W
Anzahl der MIMO Streams in der Praxis im Downlink	2	2	2–4
Anzahl der MIMO Streams in der Praxis im Uplink	1	1–2	1
Multiple Access Verfahren	Ein Gerät zu einer Zeit	OFDMA oder einzelnes Gerät pro Timeslot	OFDMA
Duplex Verfahren (Trennung von Uplink/ Downlink)	TDD (Time Division Duplex)	TDD	Hauptsächlich Frequency Division Duplex (FDD), TDD wird jedoch auch verwendet

In den IEEE Spezifikationen wird 802.11ax auch als „High Efficiency (HE) PHY“ bezeichnet¹⁴, da die Erweiterung eine Anzahl an Funktionen mitbringt, um Netze mit einer großen Zahl an Geräten, die gleichzeitig Daten übertragen, besser auszulasten. Abgesehen von den Multi-User MIMO Erweiterungen von Wi-Fi 5 konnte bisher in WLAN Netzen nur ein Gerät zu einer Zeit Daten senden und hatte somit auch Zugriff auf die gesamte Kanalbandbreite. In der Downlink Richtung ändert Wi-Fi 6 dies nun, indem es einem Access Point ermöglicht, mehrere Geräte gleichzeitig anzusprechen und deren Pakete an unterschiedlichen Stellen des Kanals über die Frequenz zu platzieren. Dies ist schon von LTE und 5G bekannt und wird als Orthogonal Frequency Division Multiple Access (OFDMA) bezeichnet. Dazu wird bei WLAN nun der Kanal in 2, 4, 8, 20 oder 80 MHz Teilstücke aufgeteilt, die als Resource Units (RU) bezeichnet werden. Hat der Access Point IP Pakete für mehrere Endgeräte im Sendepuffer, kann er den Endgeräten mitteilen, dass ein paralleler Datentransfer in einem Multi-User Paket stattfinden wird. Unterschiedliche Modulation und Kodierungsverfahren können dann in den RUs verwendet werden. Außerdem können individuelle Pilotkanäle pro RU verwendet werden, damit Endgeräte ihren Teil des Kanals besser dekodieren können. Zusätzlich können unterschiedlich breite RUs in einem Multi-User Paket verwendet werden, falls z. B. im Sendepuffer mehr Daten für ein Gerät als für andere vorhanden ist. Auch ACK Pakete, die den erfolgreichen Empfang signalisieren, können parallel übertragen werden. An dieser Stelle sei angemerkt, dass OFDMA im Downlink nur verwendet wird, wenn Daten für mehrere Endgeräte im Sendepuffer des AP vorhanden sind. Ist dies nicht der Fall, läuft die Datenübertragung, wie bisher auch, im Single-User Modus.

Auch Uplink OFDM ist in Wi-Fi 6 beschrieben, wird eventuell aber erst von Wave-2 Geräten unterstützt werden. Gegenüber Downlink OFDMA ist Uplink OFDMA komplexer, da hier der Access Point alle Geräte, die Uplink OFDMA unterstützen, regelmäßig nach ihrem Sendepufferstatus fragen muss (Polling). Erst mit dieser Information kann der AP dann den Endgeräten mitteilen, wann und in welchem Teil des Kanals sie senden sollen. Außerdem wurden spezielle „Random Access“ Gelegenheiten mit einem Backoff Mechanismus spezifiziert, damit Geräte den AP um Uplink OFDMA Zuweisungen bitten können. Der AP nutzt dann diese „Trigger“ Frames, um Ressourcen entsprechend zuzuweisen. Diese Zuweisungen enthalten dann auch Informationen, welche Modulation und Kodierung das Endgerät verwenden soll, sowie die zu verwendende Sendeleistung. Dieses Konzept ist somit der Resource Zuweisung in LTE und 5G Netzwerken sehr ähnlich und sorgt bei einer hohen Anzahl an wartenden Endgeräten für eine effiziente Nutzung des Kanals. Trigger Frames können vom AP auch für folgende Operationen verwendet werden:

- Beamforming Report Polls (BRP)
- Multi-User Block ACK Requests (MU-BAR)
- Multi-User Request to Send (MU-RTS)
- Buffer Status Report Polls (BSRP). Damit erhält der AP Informationen, wie viele Daten im Sendepuffer der Geräte vorhanden ist.

Zusätzlich zu OFDM erweitert Wi-Fi 6 auch die Multi-User MIMO Funktion von Wi-Fi 5. Beispielsweise ist es jetzt möglich, bis zu 8 Geräte gleichzeitig anzusprechen. In Wave-1 Produkten können OFDMA und MU-MIMO nicht gleichzeitig verwendet werden. Auch Uplink MU-MIMO wurde spezifiziert, wird jedoch ebenfalls nicht in Wave-1 Produkten verwendet.

Für Orte, an denen viele WLAN Netze gleichzeitig betrieben werden, wurde für Wi-Fi 6 Wave-2 Produkte die BSS Coloring Funktion spezifiziert. Werden mehrere unabhängige Netze gleichzeitig an einem Ort betrieben, können Geräte in einem Netzwerk auch die Übertragungen von Geräten in anderen Netzen sehen. Auch wenn diese mit geringer Signalstärke empfangen werden, lösen diese trotzdem den CSMA/CA Collision Avoidance Mechanismus aus und das Endgerät wartet dann mit seiner Übertragung. Da das schwache Signal jedoch für einen anderen AP bestimmt ist, ist eine Wartezeit aber unter Umständen nicht notwendig. Mit BSS Coloring gibt ein Endgerät in jeder Paket Präambel an, in welchem Netzwerk es beheimatet ist. Wenn ein anderes Gerät dann Daten zu übertragen hat und nur ein schwaches Signal mit einem BSS Coloring Code für einen anderen AP empfängt, braucht es somit nicht mehr auf das Ende der Übertragung warten.

BSS Coloring kann auch in größeren Netzwerken nützlich sein, die viele APs im gleichen Netzwerk verwenden und die gleiche SSID aussenden. Um eine möglichst nahtlose Mobilität zwischen den APs zu gewährleisten, überlappen sich die Signale der APs. BSS Coloring kann hier verwendet werden, um Übertragungen im gleichen Netzwerk zu unterschiedlichen APs zu unterscheiden. Und schließlich kann mit BSS Coloring auch der Stromverbrauch eines Gerätes gesenkt werden, da ein Endgerät Pakete zu von und zu Nachbar-APs erkennt und diese somit nicht dekodieren muss.

802.11ax enthält auch eine Anzahl Funktionen für das Internet der Dinge (IoT), also Geräte, die typischerweise nur sehr selten und sehr wenig Daten übertragen. Diese werden jedoch oft von einer Batterie gespeist, sind nicht leicht zugänglich und müssen deshalb so stromsparend wie möglich sein. 802.11ax bietet solchen Geräten eine 20 MHz-only Option. Auch wurde ein neuer Power Save Mode eingeführt, der als Target Wait Time (TWT) bezeichnet wird. Die Idee hinter TWT ist, dass Endgerät und AP sich auf einen Zeitraum von Minuten oder Stunden einigen können, in dem das Endgerät nicht erreichbar ist.

Zusammenfassend kann gesagt werden, dass Wi-Fi 6 in Netzen mit wenigen Endgeräten eine etwas höhere Spitzendatenrate gegenüber Wi-Fi 5 bieten kann. Wi-Fi 6 bietet jedoch eine bessere Effizienz und einen höheren Datendurchsatz in alltäglichen Szenarien mit vielen Geräten in einem Netzwerk, sowie an Orten, in denen sich viele Netze überlappen. Rückwärtskompatibilität zu älteren Geräten und Netzwerken bleiben jedoch erhalten. Dies wird hauptsächlich durch die CTS Kanalreservierung gewährleistet, sowie durch die Verwendung von älteren Modulations- und Kodierungsverfahren in Präambeln, sowie in Management- und Beacon Frames, die auch von älteren Geräten dekodiert werden können.

4.7 Wireless LAN-Sicherheit

Sicherheit ist bei Wireless LAN vor allem deswegen ein sehr intensiv diskutiertes Thema, da die Verwendung eines nicht verschlüsselten Netzwerks ein großes Risiko darstellt.

Im Auslieferungszustand ist die Verschlüsselung noch immer bei manchen Access Points deaktiviert. Wird die Verschlüsselung nicht explizit konfiguriert, kann jedes WLAN-fähige Endgerät diesen Access Point ohne vorherige Erlaubnis des Besitzers verwenden. Diese Konfiguration eignet sich vor allem für öffentliche Hotspots, da hier stets wechselnde Teilnehmer einen Hotspot verwenden. Da die Daten aber unverschlüsselt übertragen werden, können diese sehr leicht von anderen abgehört werden und es bleibt den Anwendern überlassen, sich mit VPN-Tunnels und anderen Maßnahmen zu schützen.

Noch bedenklicher ist eine offene Konfiguration für private Heimnetzwerke, die über den Access Point einen Zugang zum Internet herstellen. Wurde die Verschlüsselung nicht konfiguriert, können Nachbarn ohne Wissen des Besitzers seine Internetverbindung nutzen. Außerdem ist es wie bei öffentlichen Wi-Fi Hotspots möglich, den Datenverkehr abzuhören und so z. B. Passwörter etc. auszuspähen. Da auf alle Rechner des Netzwerks Zugriff besteht, können Angreifer auch direkt Schwachstellen der Betriebssysteme ausnutzen, um Daten auf den Rechnern ausspähen.

4.7.1 Wired Equivalent Privacy (WEP) und frühere Sicherheitsverfahren

Um WLAN Netzwerke vor nicht autorisierten Zugriffen zu schützen, wurde ursprünglich die Wired Equivalent Privacy (WEP) Authentifizierung und Verschlüsselung verwendet. Diese war Teil der 802.11b, g und a Standards. Über die Jahre wurden jedoch eine Anzahl von Sicherheitsproblemen gefunden und WEP wurde seither von WPA, WPA-2 und WPA-3 ersetzt. Diese Verfahren werden nachfolgend beschrieben. Da WEP heute typischerweise nicht mehr verwendet wird, wird das Verfahren nachfolgend auch nicht weiter betrachtet.

4.7.2 WPA und WPA-2 Personal Mode-Authentifizierung

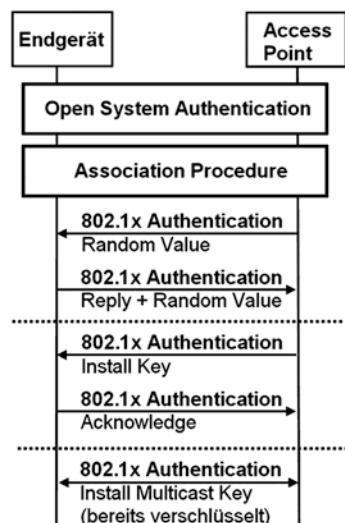
Aufgrund der oben genannten Sicherheitsprobleme wurde von der IEEE 802.11i-Arbeitsgruppe der 802.1x Standard erarbeitet. Da sich jedoch die Verabschiedung des Standards damals beträchtlich hinauszögerte, wurde die Industrie ihrerseits selbst aktiv und entwickelte in der Wi-Fi Alliance die Wireless Protected Access (WPA)-Spezifikation. WPA enthält alle wichtigen Funktionalitäten von 802.11i und

wurde so spezifiziert, dass die neuen Funktionen auch mit Hardware funktionierte, die ursprünglich nur für WEP-Verschlüsselung entwickelt wurde.

Die Schwächen von WEP werden von WPA durch verbesserte Authentifizierung während der Verbindungsaufnahme und einer neuen Verschlüsselung gelöst. Wie in Abb. 4.8 gezeigt, meldet sich ein Teilnehmer bei einem Netzwerk durch eine Pseudo-Authentifizierung und eine Association-Prozedur am Netzwerk an. Bei WPA folgt darauf eine weitere Authentifizierung und danach eine sichere Schlüsselübergabe für die Chiffrierung der Nutzdaten. Die erste Authentifizierung ist damit überflüssig, wurde aber dennoch aus Kompatibilitätsgründen beibehalten. Um Endgeräten mitzuteilen, dass ein Netzwerk WPA statt WEP unterstützt, enthalten Beacon Frames einen zusätzlichen WPA-Parameter. Dieser informiert Endgeräte, dass ein zusätzlicher Authentifizierungsschritt und Schlüsselaustausch nach der Association-Prozedur notwendig ist. Der WPA-Parameter enthält auch zusätzliche Informationen über den für die Authentifizierung und Verschlüsselung zu verwendenden Algorithmus. Erste WPA-Endgeräte verwendeten zunächst nur TKIP (Temporal Key Integrity Protocol) für die Verschlüsselung. Aktuelle Endgeräte unterstützen auch AES (Advanced Encryption Standard), welcher bei WPA-2 fest vorgeschrieben ist. Weitere Details dazu folgen im Laufe dieses Kapitels.

Abb. 4.20 zeigt die vier neu hinzugekommenen Schritte für die WPA Pre-Shared Key (PSK)-Methode, um Endgeräte gegenüber dem Access Point zu authentifizieren und umgekehrt. Pre-Shared Key bedeutet in diesem Zusammenhang, dass im Access Point und im Endgerät das gleiche Passwort hinterlegt wurde. Außerdem einigen sich Endgeräte und Access Point während dieses Prozesses auf ein gemeinsames Schlüsselpaar für die Chiffrierung der Nutzdaten, die Session Keys. In der ersten Nachricht sendet der Access Point eine Zufallszahl an das Endgerät. Dieses verwendet dann die Zufallszahl

Abb. 4.20 WPA-PSK-Authentifizierung und Schlüsselaustausch



und das gemeinsame Passwort (Pre-Shared Key), um eine Antwort zu generieren. Das Passwort hat eine Länge von 8 bis 64 Zeichen. Die Antwort wird dann zusammen mit einer weiteren Zufallszahl zurück an den Access Point geschickt. Der Access Point vergleicht im nächsten Schritt die Antwort mit der zuvor selber berechneten Antwort. Diese können nur identisch sein, wenn beide Seiten für die Berechnung der Antwort das gleiche Passwort verwendet haben. Stimmen die Antworten überein, ist das Endgerät authentifiziert. Im nächsten Schritt generiert der Access Point einen Sitzungsschlüssel (Session Key), welcher mit dem gemeinsamen Passwort verschlüsselt wird und zum Endgerät geschickt wird. Das Endgerät entschlüsselt den Sitzungsschlüssel mit dem gemeinsamen Passwort und bestätigt dem Access Point den korrekten Empfang der Nachricht. Diese Bestätigung aktiviert auch implizit die Verschlüsselung in beiden Richtungen. In einem letzten Schritt teilt dann der Access Point dem Endgerät noch den Schlüssel für die Dechiffrierung von Broadcast Frames mit. Diese Nachricht ist bereits verschlüsselt. Während der Sitzungsschlüssel für jedes einzelne Endgerät individuell ist, ist der Broadcast-Schlüssel für alle Endgeräte gleich, da ein Broadcast-Paket von allen Endgeräten gleichzeitig entschlüsselt werden muss.

Der Vorteil der Verwendung von individuell generierten Sitzungsschlüsseln gegenüber der direkten Verwendung des Passworts für die Verschlüsselung ist, dass dieser während einer laufenden Verbindung geändert werden kann. Dies verhindert sogenannte „Brute Force-Attacken“, die versuchen, den Schlüssel durch ausprobieren zu erraten. Ein typischer Wert für das Austauschen des Sitzungsschlüssels ist eine Stunde.

Während WPA-PSK ein WLAN Netzwerk gegen externe Angreifer schützen kann, hat diese Art der Authentifizierung jedoch eine interne Schwachstelle. Ein Angreifer, dem das gemeinsame Passwort bekannt ist und die Prozedur zur Erzeugung des individuellen Session Keys während der Authentifizierung beobachtet, kann im Anschluß alle Datenpakete des Endgerätes entschlüsseln. Netzwerkanalysesoftware wie z. B. Wireshark haben diese Funktionalität bereits eingebaut. Dies bedeutet, dass in einem WPA-PSK geschützten Netzwerk neben dem gemeinsam genutzten Passwort nur wenig Wissen notwendig ist, um Pakete von anderen Geräten abzufangen und zu entschlüsseln.

4.7.3 WPA und WPA-2 Enterprise Mode Authentifizierung – EAP-TLS

Zusätzlich zur WPA-PSK-Authentifizierung, für die ein gemeinsamer Schlüssel (Pre-Shared Key) im Access Point und in den Endgeräten gespeichert werden muss, gibt es bei WPA und WPA-2 auch einen Enterprise Mode für Firmen. Hier werden die Passwörter oder Zertifikate in einem zentralen Authentifizierungsserver gespeichert. Dies ermöglicht es Firmen, mehrere Access Points zu betreiben, ohne die Authentifizierungsinformationen in jedem einzelnen Access Point separat vorrätig zu halten. Außerdem ermöglicht diese Methode, Nutzer individuell zu authentifizieren. Somit kann für

jeden Benutzer individuell eine Zugangsberechtigung erteilt und auch wieder entzogen werden. Für die Kommunikation zwischen Access Point und Authentication Server wird z. B. das RADIUS (Remote Authentication Dial In User Service) Protokoll verwendet.

WPA Professional unterstützt unterschiedliche Authentifizierungsprotokolle, die Extensible Authentication Protocols (EAP) genannt werden. Ein sehr häufig genutztes EAP Protocol ist EAP Transport Layer Security (EAP-TLS) von Haverinen and Salovei, das in RFC 5216 beschrieben ist. Dieses Protokoll verwendet Zertifikate, die im Endgerät und Authentifizierungsserver gespeichert werden. Der wichtigste Teil des Zertifikats sind die öffentlichen Schlüssel (Public Keys) des Endgeräts und des Authentifizierungsservers. Diese Schlüssel werden verwendet, um die Sitzungsschlüssel (Session Keys) zu chiffrieren, die zwischen Endgerät und Netzwerk ausgetauscht werden. Wie zuvor beschrieben, werden dann die Sitzungsschlüssel verwendet, um die Nutzdaten zwischen Access Point und Endgerät zu verschlüsseln.

Nachdem der Sitzungsschlüssel mit dem öffentlichen Schlüssel chiffriert wurde, kann dieser nur mit dem privaten Schlüssel (Private Key) der Gegenstelle wieder dechiffriert werden. Dieser Vorgang wird in Abb. 4.21 gezeigt. Da die privaten Schlüssel niemals zwischen Endgerät und Netzwerk ausgetauscht werden, kann durch Abhören der Verbindung der Session Key nicht kompromittiert werden. Somit kann auch die Übertragung der Nutzdaten später nicht von einem Angreifer dechiffriert werden. Einziger Nachteil der Nutzung von Zertifikaten ist der Umstand, dass diese einmalig auf dem Endgerät installiert werden müssen. Dies ist etwas komplizierter, als einfach ein Passwort zu vergeben, jedoch wesentlich sicherer, wenn die Zertifikate korrekt verteilt und installiert werden. Nicht gezeigt wird in Abb. 4.21 die Übertragung des

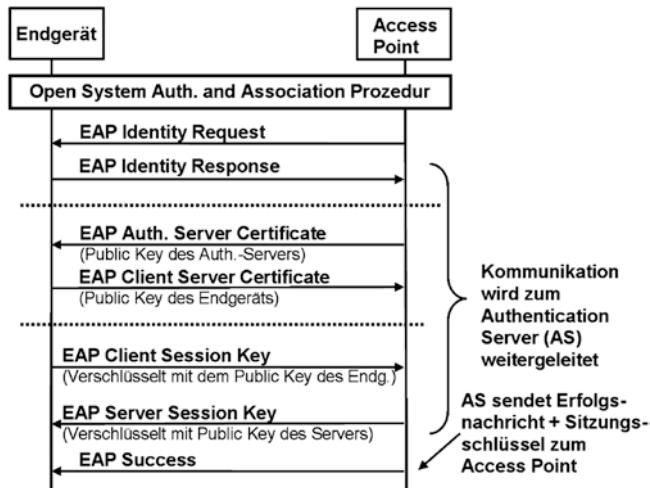


Abb. 4.21 EAP-TLS-Authentifizierung

Sitzungsschlüssels für Broadcast-Pakete, die unmittelbar nach der erfolgreichen Authentifizierung übertragen werden.

In Abb. 4.21 ist außerdem zu sehen, dass der Access Point in der Authentifizierungsphase nur den Datenaustausch mit dem Authenticationserver zulässt. Erst nachdem die Authentifizierung erfolgreich abgeschlossen wurde und nachdem der Server dem Access Point die Freigabe erteilt hat, erlaubt der Access Point dem Endgerät freien Zugriff auf das Netzwerk. Die Nutzdaten sind dann über die Luftschnittstelle schon verschlüsselt. Üblicherweise ist das erste Nutzdatenpaket eine DHCP (Dynamic Host Configuration Protocol)-Anforderung, um eine IP-Adresse zu erhalten.

Des Weiteren sei angemerkt, dass die EAP-TLS-Authentifizierung große Ähnlichkeit mit TLS und SSL (Secure Socket Layer) hat. Diese Protokolle werden von HTTPS (Secure HTTP) für die Authentifizierung und Generierung von Sitzungsschlüsseln für eine sichere Verbindung zwischen einem Web Server und einem Webbrower verwendet. Der Hauptunterschied zwischen EAP-TLS und der HTTP TLS-Authentifizierungsprozedur ist, dass bei EAP-TLS eine gegenseitige Authentifizierung stattfindet, während sich bei HTTPS TLS nur der Web Server gegenüber dem Webbrower authentifizieren muss. Aus diesem Grund benötigt der Webbrower auch kein Zertifikat für den Aufbau einer verschlüsselten Verbindung.

4.7.4 WPA und WPA-2 Enterprise Mode Authentication – EAP-TTLS

Eine andere, in der Praxis ebenso gebräuchliche EAP Methode ist EAP-TTLS (Tunneled Transport Layer Security). Statt Zertifikaten auf Netzwerk- und Endgeräteseite verwendet diese EAP Methode nur ein Zertifikat auf der Netzwerkseite und auf der Endgeräteseite eine Kombination aus Nutzernamen und Passwort. Somit ist es nicht notwendig, Zertifikate auf dem Endgerät zu installieren. Diese Authentifizierung wurde z. B. vom Autor bei Konferenzen beobachtet, auf denen die Organisatoren einen geschützten Internetzugang über WLAN anboten. Die Vorteile dieser Methode zu WPA-PSK sind wie folgt:

1. Individueller Benutzernamen und Passwort pro Endgerät stellen sicher, dass passive Angreifer keine Datenpakete eines Endgerätes entschlüsseln können, auch wenn sie die initiale Authentifizierung abhören konnten.
2. Endgeräte können während des Verbindungsaufbaus überprüfen, ob sie sich mit dem richtigen Netzwerk verbinden und nicht ein Angreifer versucht, mit einem Access Point mit gleicher SSID wie das eigentliche Netzwerk die Verbindung zu kapern. Dies ist durch die Verwendung des Zertifikats möglich, da nur das richtige Netzwerk den dazugehörigen privaten Schlüssel besitzt, der den Austausch von Username und Passwort schützt.

Abb. 4.22 zeigt den Ablauf einer EAP-TTLS Certificate Authentifizierung. Nach der Association Prozedur mit dem Netzwerk fragt der WLAN Access Point zunächst nach einem Nutzernamen, der anonym sein kann und der dazu dient, einen Authentifizierungsserver im Netzwerk auszuwählen, falls dort mehrere zur Auswahl stehen. Danach beginnt die TTLS-EAP Authentifizierungsprozedur. Das Endgerät antwortet mit einem „Client Hello“ Paket, das eine Liste aller unterstützten Verschlüsselungsalgorithmen enthält. Das Netzwerk sucht sich daraus einen passenden Algorithmus aus und sendet dann sein signiertes Zertifikat, das den Public Key für die Authentifizierungsprozedur enthält, an das Endgerät.

In Firmennetzwerken wird das WLAN Zertifikat üblicherweise von einer nicht-öffentlichen Zertifikatsinstanz ausgestellt. Aus diesem Grund muss das zu dieser nicht öffentlichen Instanz gehörende „Root Zertifikat“ zuvor manuell auf dem Endgerät installiert werden, damit das Endgerät die Gültigkeit des WLAN Zertifikats beim Verbindungsaufbau überprüfen kann. Wenn eine öffentliche Zertifizierungstelle das WLAN Zertifikat signiert hat, ist dies nicht nötig, da das „Root Zertifikat“ der öffentlichen Zertifizierungsstelle schon Teil des Betriebssystems des Endgeräts ist. Wichtig ist in beiden Fällen, dass das Endgerät nicht nur prüft, ob das Zertifikat gültig ist, sondern auch, ob es den Domain Namen enthält, der für das Netzwerk gültig ist. Der Domain Name, der für das Zertifikat gültig ist, wird im Zertifikat im „alt-subject“ Parameter hinterlegt und muss vom Endgerät geprüft werden. Dieser Parameter kann z. B. über eine Konfigurationsdatei festgelegt werden oder manuell durch Eintrag des Domänenamens in

No.	Time	Source	Destination	Protocol	Length	Info
1	11:13:48,39	ArubaNet 8a:26:e0		EAP	62	Request, Identity
2	11:13:48,39	Woonsang 04:05:06		EAP	26	Response, Identity
3	11:13:48,44	ArubaNet 8a:26:e0		EAP	62	Request, Tunneld TLS EAP (EAP-TTLS)
4	11:13:48,44	Woonsang 04:05:06		TLSv1	225	Client Hello
5	11:13:48,47	ArubaNet 8a:26:e0		TLSv1	1024	Server Hello, Certificate, Server Key Exchange, Server Hello Done
6	11:13:48,47	Woonsang 04:05:06		EAP	26	Response, Tunneld TLS EAP (EAP-TTLS)
7	11:13:48,53	ArubaNet 8a:26:e0		TLSv1	1024	Server Hello, Certificate, Server Key Exchange, Server Hello Done
8	11:13:48,53	Woonsang 04:05:06		EAP	26	Response, Tunneld TLS EAP (EAP-TTLS)
9	11:13:48,66	ArubaNet 8a:26:e0		TLSv1	1024	Server Hello, Certificate, Server Key Exchange, Server Hello Done
10	11:13:48,66	Woonsang 04:05:06		EAP	26	Response, Tunneld TLS EAP (EAP-TTLS)
11	11:13:48,68	ArubaNet 8a:26:e0		TLSv1	1024	Server Hello, Certificate, Server Key Exchange, Server Hello Done
12	11:13:48,68	Woonsang 04:05:06		EAP	26	Response, Tunneld TLS EAP (EAP-TTLS)
13	11:13:48,69	ArubaNet 8a:26:e0		TLSv1	95	Server Hello, Certificate, Server Key Exchange, Server Hello Done
14	11:13:48,69	Woonsang 04:05:06		TLSv1	160	Client Key Exchange, Change Cipher Spec, Encrypted Handshake Message
15	11:13:48,78	ArubaNet 8a:26:e0		TLSv1	89	Change Cipher Spec, Encrypted Handshake Message
16	11:13:48,79	Woonsang 04:05:06		TLSv1	132	Application Data, Application Data
17	11:13:48,79	ArubaNet 8a:26:e0		EAP	62	Success
18	11:13:48,79	ArubaNet 8a:26:e0		EAPOL	137	Key (Message 1 of 4)
19	11:13:48,80	Woonsang 04:05:06		EAPOL	137	Key (Message 2 of 4)
20	11:13:48,80	ArubaNet 8a:26:e0		EAPOL	171	Key (Message 3 of 4)
21	11:13:48,80	Woonsang 04:05:06		EAPOL	115	Key (Message 4 of 4)
22	11:13:49,90	151.217.197.162	DHCP	344	DHCP Offer	- Transaction ID 0x5f094dd4
23	11:13:49,92	151.217.197.192,5	DHCP	344	DHCP ACK	- Transaction ID 0x5f094dd4

Version: 802.1X-2001 (1)
Type: EAP Packet (0)
Length: 1004
▼ Extensible Authentication Protocol
Code: Request (1)
Id: 3
Length: 1004
Type: Tunneld TLS EAP (EAP-TTLS) (21)
▼ EAP-TLS Flags: 0xc0

Abb. 4.22 EAP-TTLS Authentifizierung mit Zertifikat

die Verbindungseinstellungen. In Ubuntu Linux wird diese Einstellung im Verbindungsprofil im 802.11x EAP-TTLS Bereich vorgenommen:

```
[802-1x]
eap=ttls;
identity=x
ca-cert=/etc/ssl/certs/StartCom_Certification_Authority.pem
altsubject-matches=DNS:radius.c3noc.net;
phase2-auth=pap
password-flags=1
```

Nachdem das Endgerät das Server Zertifikat akzeptiert hat (Paket 14 in der Abbildung), wird ein verschlüsseltes Handshake Paket ausgetauscht, das Client spezifisch ist. Für diesen Dialog verwendet der Client den Public Key, der im Zertifikat enthalten war, und verschlüsselt damit die Nachricht. Auf der Netzwerkseite wird dann der Private Key verwendet, um die Nachricht wieder zu entschlüsseln. Da der Private Key niemals zu Clients übertragen wird, kann ein Angreifer keinen eigenen Access Point mit gleicher SSID aufsetzen, da er keine Kopie des Private Keys hat und somit eingehende Pakete nicht entschlüsseln kann.

Danach wird eine normale EAPOL WLAN Authentifizierung durchgeführt und die Link Level Verschlüsselung aktiviert. Diese basiert auf einem individuellen „Secret“, das während der TTLS Prozedur ausgetauscht wurde. Paket 22 zeigt das erste verschlüsselte Paket, das zwischen Client und Access Point ausgetauscht wird und eine DHCP Anfrage enthält, um eine IP Adresse zu erhalten. Da der Trace auf dem Client Gerät gemacht wurde, ist die entschlüsselte Variante des Pakets zu sehen. Nachdem der Client eine IP-Adresse bekommen hat, ist die Verbindung fertig aufgebaut und Nutzdatenpakete können ausgetauscht werden.

4.7.5 WPA und WPA-2 Enterprise Mode Authentication – EAP-PEAP

Eine weitere Wifi Authentifizierungsprozedur die in der Praxis verwendet wird ist EAP – Protected Extensible Authentication Protocol (EAP-PEAP). Verwendet wird EAP-PEAP z. B. von manchen Universitäten die dem Eduroam Wifi System (<https://www.eduroam.org>) angeschlossen sind. Eduroam bietet Studenten und Mitarbeitern einen Internet Zugang nicht nur an ihrer eigenen Universität, sondern auch bei allen anderen teilnehmenden Institutionen (Roaming), ohne dass dabei die Netzwerkeinstellungen geändert werden müssen.

Wie das zuvor beschriebene EAP-TTLS verwendet EAP-PEAP ein Zertifikat, um das Netzwerk gegenüber dem Endgerät zu authentifizieren und einen Nutzernamen und Passwort in der Gegenrichtung. Um eine Man-in-the-Middle Attacke zu verhindern,

muss das Endgerät so konfiguriert sein, dass es auch im Roamingfall nur das Zertifikat der Heimatinstitution akzeptiert.

Da Eduroam ein verteiltes System ist, gibt es kein zentrales Authentifizierungssystem. Jede Institution (z. B. Universität) hat ihre eigenen Authentifizierungsserver, die über das Internet von andern teilnehmenden Institutionen erreicht werden können. Zu Beginn des Authentifizierungsprozesses muss das Endgerät eine anonyme Identität senden, die eine Information über die Heimatinstitution enthält. Mit dieser Information findet das Eduroam System dann den Authentifizierungsserver des Teilnehmers im lokalen Netzwerk, oder im Falle eines Roamers bei einer anderen Universität und fordert dort das entsprechende Zertifikat für die Authentifizierung an. Der Access Point leitet dann das Zertifikat an das Endgerät weiter, das dann überprüfen muss, ob es gültig ist und ob es von der Heimatuniversität stammt. Danach verschlüsselt das Endgerät seinen Usernamen und Passwort mit dem Public Key der im Netzwerkzertifikat enthalten ist und sendet das Ergebnis an den Access Point. Von dort aus wird das Paket dann entweder an den lokalen Authentifizierungsserver oder, im Falle eines Roamers, zum Server der Heimatinstitution gesendet. Interessant ist es in diesem Zusammenhang, dass das lokale Netzwerk den Nutzernamen und das Passwort eines Roamers nicht sieht, da dieses ja verschlüsselt ist und nur vom Authentifizierungsserver im Heimatnetzwerk entschlüsselt werden kann. Sind Nutzernamen und Passwort gültig, informiert der Server den Wifi Access Point, der dann den Zugang zum Netzwerk freigibt. Der Internetzugang wird dann nicht nur für lokale Nutzer direkt vom lokalen Netzwerk bereitgestellt, sondern auch für Roamer. Das bedeutet, dass die Datenpakete von Roamern nicht wie im Mobilfunk erst zurück zum Heimatnetzwerk geschickt werden und erst von dort aus ins Internet.

Typischerweise stellen Universitäten ihren Mitarbeitern und Studenten Konfigurationsdateien oder Installationsprogramme bereit, um die Netzwerkeinstellungen von Endgeräten automatisch zu konfigurieren. In Ubuntu Linux sind z. B. für die Universität Wien folgende Einstellungen notwendig:

```
[802-1x]
eap=peap;
anonymous-identity=@univie.ac.at
identity=a8398493@univie.ac.at
ca-cert=/home/..../eduroam-full-certificate-chain-university-of-
vienna.pem

domain-suffix-match=univie.ac.at

phase2-auth=mschapv2
```

Neben dem Pfad zu den benötigten Root-Zertifikaten, mit denen das Zertifikat des Heimatnetzwerkes überprüft werden kann, ist die Zeile „domain-suffix-match“ von besonderer Bedeutung. Diese Zeile ist notwendig, damit das Endgerät nur Zertifikate der Heimatuniversität während des Authentifizierungsprozesses zulässt und somit

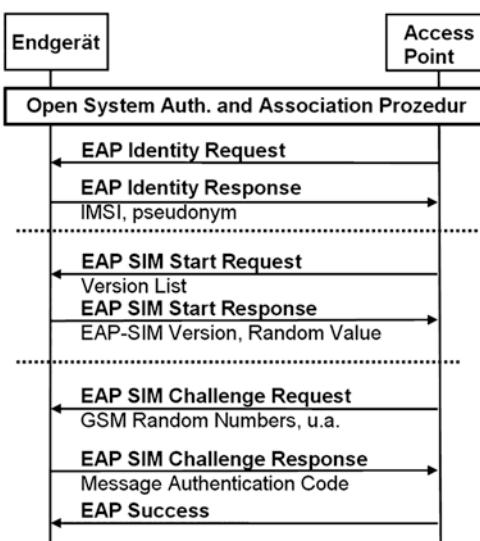
Man-in-the-Middle Attacken erkannt werden, die ein gültiges Zertifikat verwenden, das aber nicht von der Heimatinstution des Teilnehmers ausgestellt wurde¹⁵.

4.7.6 WPA und WPA Enterprise Mode Authentifizierung – EAP-SIM

Viele Mobilfunknetzbetreiber betreiben heute auch ein Wi-Fi Hotspot Netzwerk, um an besonders frequentierten Orten wie Bahnhöfen, Flughäfen und Hotels zusätzliche Netzkapazität zur Verfügung zu stellen. Ein Problem dabei ist, wie sie ihre Kunden auch in ihrem Wi-Fi Netzwerk authentifizieren können. Zwar gibt es hier heute auf dem Markt schon einige Lösungen, die jedoch alle eine Interaktion mit dem Nutzer vorsehen. Da dies umständlich und für viele Applikationen hinderlich ist, wurde das EAP-SIM-Protokoll in RFC 4186 für GSM-SIMs spezifiziert und das EAP-AKA-Protokoll in RFC 4187 für UMTS/LTE-SIMs. Bei dieser Art der Authentifizierung ist wie bei GSM, UMTS oder LTE keine Interaktion mit dem Nutzer nötig, da alle Informationen von der SIM-Karte abgefragt werden. Diese Art der Authentifizierung bietet sich somit besonders für Smartphones und Tablets an, die neben einer Wi-Fi Schnittstelle auch über eine Mobilfunkchnittstelle und die dazugehörige SIM Karte verfügen.

EAP-SIM und EAP-AKA verwenden die gleiche Authentifizierungsmethode wie bereits im Kapitel über WPA Personal und Enterprise Authentifizierung beschrieben. Abb. 4.23 zeigt die Nachrichten, die während der Authentifizierung zwischen dem mobilen Endgerät und dem Authentifizierungsserver über einen EAP-SIM bzw. EAP-AKA kompatiblen Access Point übertragen werden. Nach einer Wi-Fi Open System-Authentifizierung und der Association-Prozedur startet das Netzwerk die EAP-Prozedur durch Senden einer EAP Identity Request-Nachricht, auf die das mobile Endgerät mit

Abb. 4.23 EAP-SIM-Authentifizierung



einer EAP Identity Response-Nachricht antworten muss. Die Identität, die in dieser Nachricht zurückgegeben wird, besteht aus einem Identity Type Identifier, der IMSI aus der SIM-Karte und einem spezifischen Postfix (Anhang) des Mobilfunkbetreibers. Alternativ kann das mobile Endgerät auch eine temporäre Identität (Pseudonym) an das Netzwerk schicken, das während einer früheren Authentifizierung ausgehandelt wurde. Das Pseudonym hat die gleiche Aufgabe wie die TMSI (Temporary Mobile Subscriber Identity) in GSM und UMTS Netzwerken, nämlich der Verschleierung der Identität gegenüber Abhörversuchen auf der Luftschnittstelle, hat jedoch ein anderes Format.

Im nächsten Schritt sendet dann das Netzwerk eine EAP-SIM bzw. EAP-AKA Start Request-Nachricht. Diese enthält Informationen über die unterstützten EAP-SIM/AKA-Authentifizierungsalgorithmen. Das mobile Endgerät wählt dann einen dieser Algorithmen aus und antwortet mit einer EAP-SIM/AKA Start Response-Nachricht. Diese enthält eine Zufallszahl, die später im Netzwerk zusammen mit dem geheimen GSM-Schlüssel Kc für diverse Berechnungen verwendet wird. Da der geheime GSM-Schlüssel Kc sowohl dem Netzwerk als auch der SIM-Karte bekannt ist, kann sich auf diese Weise nicht nur das Endgerät gegenüber dem Netzwerk authentifizieren, sondern das Netzwerk auch gegenüber dem Endgerät.

An diesem Punkt verwendet der Authentifizierungsserver die IMSI des Teilnehmers, um vom Home Location Register (HLR)/Authentication Center (AuC) Authentication Triplets anzufordern. Das HLR/AuC antwortet auf diese Anfrage mit zwei oder drei Triplets, die jeweils eine Zufallszahl und Kc-Chiffrierungsschlüssel enthalten. Diese werden dann verwendet, um die EAP-SIM-Sitzungsschlüssel und andere Parameter für den Authentifizierungsprozess zu erzeugen. Diese werden dann in verschlüsselter Form zusammen mit den zwei oder drei GSM Zufallszahlen im Klartext zum mobilen Endgerät in der SIM Challenge Request-Nachricht geschickt. Wird EAP-AKA verwendet, werden an dieser und anderer Stelle die entsprechenden UMTS Authentifizierung und Verschlüsselungsparameter verwendet.

Das Endgerät schickt nach Empfang der Nachricht die GSM Zufallszahlen weiter zur SIM-Karte. Die SIM-Karte erzeugt mit diesen die GSM Signed Response (SRES) und die GSM Chiffrierschlüssel (Kc), die im folgenden verwendet werden, um die zuvor erhaltenen EAP-SIM/AKA Parameter zu entschlüsseln. Stimmt nach der Entschlüsselung die Signed Response vom Netzwerk mit der der SIM-Karte überein, ist das Netzwerk authentifiziert und das Endgerät kann eine korrekte Antwort zurückschicken. Im Netzwerk wird diese Nachricht dann verifiziert und im Erfolgsfall eine EAP Success-Nachricht an das Endgerät zurückgeschickt. Ab diesem Zeitpunkt hat das Endgerät dann Zugriff auf das Netzwerk.

Abb. 4.24 zeigt die bei der EAP-SIM/AKA-Authentifizierung beteiligten Komponenten und Protokolle. Links ist das mobile Endgerät dargestellt, das seine EAP-Nachrichten über das EAPOL Protokoll sendet. Für die Kommunikation zwischen Access Point und dem Authentifizierungsserver wird das RADIUS (Remote Authentication Dial In User Service) Protokoll verwendet. Der Authentifizierungsserver seinerseits kommuniziert mit dem HLR/AuC über das SS-7 Signalisierungsnetzwerk und dem MAP (Mobile Application Part)-Protokoll.

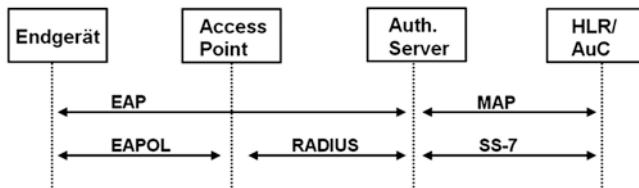


Abb. 4.24 An der EAP-SIM-Authentifizierung beteiligte Komponenten

4.7.7 Verschlüsselung mit WPA und WPA-2

Um die Verschlüsselung gegenüber WEP zu verbessern, führt WPA das Temporal Key Integrity Protocol (TKIP) ein. Bei WEP wurde ein 24 Bit Initial Vector (IV), der WEP-Schlüssel und der RC-4 Algorithmus verwendet, um eine Verschlüsselungssequenz für jedes Paket zu generieren (vgl. Abb. 4.16). TKIP verwendet nun einen 48 Bit Initial Vector, einen Master Key und den RC-4-Algorithmus, um die Verschlüsselungssequenz für jedes Paket zu erzeugen. Dieses Verfahren ist wesentlich sicherer, da der Initial Vector verlängert wurde und der Master Key ständig (z. B. einmal pro Stunde) zwischen Endgerät und Access Point neu ausgehandelt wird.

Die von WPA verwendete Verschlüsselung entspricht nicht ganz den Anforderungen des 802.11i Standards, wird jedoch trotzdem als sicher angesehen. Vorteil des Verfahrens ist, dass TKIP mit Hardware kompatibel ist, die nur für den Einsatz mit WEP vorgesehen war.

Um Attacken zu verhindern, die eine Schwäche ausnützen, die beim Wiedereinspielen von zuvor abgehörten und leicht veränderten Paketen auftreten, wird der Initial Vector bei jedem Paket um 1 erhöht. WPA-kompatible Geräte ignorieren Pakete mit schon verwendeten IV's und sind somit gegen diese Angriffsart immun.

In der Theorie können Access Points gleichzeitig WPA und WEP-Endgeräte unterstützen. In der Praxis bieten dies jedoch nur wenige Access Points an. Dies ist auch sinnvoll, da dies die Sicherheit des Systems stark reduzieren würde.

Als zusätzliche Sicherheit führt TKIP auch einen Message Integrity Code (MIC) für jedes Datenpaket ein. Der Prozess für die Erzeugung des MIC wird manchmal auch als „Michael“ bezeichnet. Im Unterschied zur CRC-Prüfsumme, die weiterhin Teil jedes Datenpakets ist, ist wie folgt: Die CRC-Prüfsumme wird aus dem Inhalt des Pakets mit einem öffentlich bekannten Algorithmus erzeugt. Der Empfänger kann somit prüfen, ob der Inhalt eines Pakets durch einen Übertragungsfehler geändert wurde. Da der Eingangsparameter und Algorithmus bekannt sind, ist die Prüfsumme jedoch nicht geeignet, um zu überprüfen, ob das Paket gezielt durch einen Angreifer verändert wurde, da der Angreifer die Prüfsumme selber ändern könnte. Der MIC andererseits wird ebenfalls mit einem bekannten Algorithmus berechnet, hat jedoch als Eingangsparameter sowohl die Nutzdaten, als auch einen Message Integrity Key, der bei der TKIP-Authentifizierung zusammen mit dem Sitzungsschlüssel erzeugt wurde. Einem Angreifer ist es somit nicht

möglich, einen korrekten MIC zu berechnen und kann somit auch nicht den Inhalt des Pakets verändern. Bleibt anzumerken, dass sowohl der CRC als auch der MIC im verschlüsselten Teil des Datenpakets untergebracht sind. Um also die CRC Prüfsumme oder den MIC zu verändern, müsste ein Angreifer also zunächst einmal die RC-4-Verschlüsselung in Kombination mit den WPA-Sicherheitsvorkehrungen überwinden.

Tritt während der Übertragung ein Fehler auf, sind beim Empfänger sowohl die MIC als auch die CRC-Prüfsumme falsch. Der Empfänger eines Datenpaketes kann somit zwischen Übertragungsfehlern und Angriffen auf die Datenintegrität unterscheiden. WPA schreibt vor, dass Endgeräte, die mehr als einen Frame pro Minute mit falschem MIC und korrektem CRC empfangen, sich vom Netzwerk trennen müssen und danach eine Minute warten, bis sie sich wieder am Netzwerk anmelden. Auf diese Weise werden Angriffe auf die Nutzdatenintegrität effektiv verhindert.

Nach der Verabschiedung des 802.11i Standards passte die Wi-Fi Alliance den WPA-Zertifizierungsprozess entsprechend an. WPA-2 ist eine Implementierung des 802.11i Standards und ist rückwärtskompatibel zu WPA. Das bedeutet, dass ein WPA-2 zertifizierter Access Point auch „nur“ WPA-fähige Endgeräte unterstützt. WPA-2 Access Points können auch ältere WEP Endgeräte unterstützen, wenn WPA/WPA-2 deaktiviert wird. Zusätzlich zum TKIP-Algorithmus, der mit WPA eingeführt wurde, unterstützt WPA-2 nun auch die stärkere AES (Advanced Encryption Standard)-Verschlüsselung. Wie bei WPA gibt es auch bei WPA-2 zwei Ausführungen: Ist ein Gerät für den „Personal Mode“ zertifiziert, erlaubt es eine Authentifizierung mit einem Access Point per Pre-Shared Key (PSK)-Verfahren. Für Firmen, in denen oftmals mehrere Access Points verwendet werden, sollte ein Access Point „WPA-2 Enterprise Mode“ zertifiziert sein. Zusätzlich zum PSK-Verfahren unterstützen solche Access Points auch das 802.1x-Authentifizierungsframework und können mit externen Authentifizierungsservern kommunizieren, wie dies weiter oben beschrieben wurde.

4.7.8 Wi-Fi Protected Setup (WPS)

Die Konfiguration eines Endgeräts für einen Wi-Fi Access Point, der eine WPA/WPA-2 Authentifizierung und Verschlüsselung fordert, ist recht einfach, denn der Nutzer gibt dazu einfach das im Access Point konfigurierte WPA/WPA-2 Passwort im Endgerät ein. Die Wi-Fi Alliance wollte jedoch diesen Prozess noch weiter vereinfachen und schuf dazu die Wi-Fi Protected Setup (WPS) Prozedur. Alle Access Points und Endgeräte müssen diese heute implementieren, um von der Wi-Fi Alliance eine Freigabe für das Anbringen der Wi-Fi Kompatibilitätslogos auf den Geräten und Verpackungen zu erhalten. WPS ist keine neue Verschlüsselungsmethode sondern soll lediglich ein einfaches Verfahren sein, den WPA/WPA-2 Schlüssel bei der ersten Konfiguration eines Endgeräts automatisch vom Access Point zu einem Endgerät zu übertragen, um dem

Nutzer damit die Eingabe eines langen Passworts zu ersparen. WPS spezifiziert unterschiedliche Verfahren, von denen Access Points heute die Pushbutton Methode, sowie die PIN-Methode unterstützen.

Die PIN-Methode kann heute in den meisten Access Points aktiviert und auch deaktiviert werden und funktioniert wie folgt:

Schritt 1: Zunächst findet ein Diffie-Hellman Schlüsselaustausch statt, um einen verschlüsselten Kanal für die Prozeduren herzustellen, die im Anschluß folgen. Der so aufgebaute Kanal dient noch nicht zu Authentifizierung, sondern nur dazu, einen verschlüsselten Kanal zur Verfügung zu stellen, über den dann sensitive Daten übertragen werden können. Ein aktiver Angreifer oder eine Offline Attacke mit aufgezeichneten Daten wird somit unterbunden. Dieser Ansatz ähnelt sehr stark dem HTTPS Ansatz bei Webseiten. Hier wird zunächst mit HTTPS ein verschlüsselter Kanal aufgebaut, über denn dann z. B. Anmelddaten übertragen werden, die der Nutzer auf einer Webseite eingegeben hat.

Schritt 2: Access Point und Endgerät erzeugen unabhängig voneinander unterschiedliche Zufallszahlen, die ‚Nonces‘ genannt werden. Zusammen mit der 8-stelligen PIN wird die Nonce als Eingabeparameter für eine Hash Funktion verwendet. Die Hash Funktion erzeugt aus diesen zwei Parametern ein 256 Bit langes Ergebnis, aus dem weder die PIN noch die Zufallszahl berechnet werden können, da die Hash Funktion nicht umkehrbar ist. Das bedeutet, dass ein Angreifer, der den Netzwerkverkehr abhört, nur aus dem Hash Wert die PIN nicht berechnen kann.

Schritt 3: Access Point und Endgerät schicken sich jeweils das Resultat ihrer Hash Funktion zu.

Schritt 4: Nachdem beide Seiten das Hash Resultat der jeweils anderen Seite erhalten haben, werden dann die Zufallszahlen (die ‚Nonces‘) ausgetauscht.

Schritt 5: Beide Endgeräte verwenden jetzt die Zufallszahl des anderen Geräts, fügen die PIN hinzu und führen die Hash Funktion mit diesen Parametern durch. Wenn die Ergebnisse mit den in Schritt 1 übertragenen Hash Resultaten übereinstimmen, können sich beide Seiten sicher sein, die gleiche PIN verwendet zu haben. Ein Angreifer kann zwar die Übertragung aller Informationen beobachten, aber nicht eingreifen, da die Hash Resultate zuerst übertragen wurden und er die PIN erst mit den Zufallszahlen aus Schritt 3 berechnen kann.

Schritt 6: Nachdem sich beide Seiten sicher sind, dass die PIN auf beiden Seiten identisch war, kann mit den übertragenen Parametern ein verschlüsselter Kanal aufgesetzt werden und der WPA/WPA-2 Schlüssel übertragen werden. Bei diesem Schlüssel handelt es sich um den Parameter, den ein Anwender eingeben würde, wenn er auf WPS verzichtet und stattdessen direkt den WPA/WPA-2 Schlüssel im Endgerät konfiguriert.

Schritt 7: Nachdem das Endgerät das WPA/WPA-2 Passwort empfangen hat, startet es eine WPA/WPA-2 Verbindungsaufnahme.

Während das beschriebene Verfahren im ursprünglichen Design sehr sicher war, wurden in der praktischen Implementierung jedoch einige Schwachstellen eingebaut, die in der Praxis das Protokoll sehr angreifbar machen. Ist WPS ständig aktiv und ist die PIN zusätzlich immer die gleiche, kann ein Angreifer mit einem Brute-Force Angriff alle PIN Werte ausprobieren. Die dafür notwendige Zeit liegt im Stundenbereich, da die PIN Überprüfung in zwei Phasen zu je 4 Ziffern durchgeführt wird. Dadurch muss ein Angreifer nur die WPS Prozedur maximal 10.000 mal starten um die ersten 4 Ziffern zu ermitteln. Im zweiten Schritt sind noch weniger Versuche nötig, da eine Ziffer eine Prüfsumme darstellt, die sicherstellen soll, dass sich der Nutzer bei der PIN Eingabe nicht vertippt hat. Diese Ziffer kann somit aus den anderen Ziffern berechnet werden. Manche Access Points versuchen einen Angriff zu verlangsamen, indem nur wenige WPS Versuche pro Minute akzeptiert werden. Doch auch mit diesem Ansatz lässt sich die benötigte Anzahl an WPS Versuchen in nur wenigen Stunden erreichen.

Die zweite WPS Schwachstelle ist der Umstand, dass ein Angreifer die PIN errechnen kann, wenn er den Datenverkehr der WPS Prozedur aufzeichnen kann. Dies ist möglich, da sowohl der Hash Wert als auch die Zufallszahl am Ende der WPS Prozedur dem Angreifer bekannt sind. Zwar kann er mit diesem Wissen die WPS Prozedur an sich nicht stören, kann aber mit dem Hash Wert und der Nonce die PIN errechnen und danach das WPA/WPA-2 Passwort dekodieren, das der Access Point dem Endgerät am Ende der WPS Prozedur übermittelt hat.

Während die Chance recht gering ist, dass ein Angreifer den WPS Prozess mitschneiden kann, ist der zuvor beschriebene Online Angriff mit durchprobieren aller PINs in der Praxis leicht möglich. Dies kann nur verhindert werden, wenn eine PIN nur einmal verwendet wird. Dies ist jedoch für den Nutzer nicht sehr komfortabel, da die PIN dann nicht auf die Rückseite des Geräts gedruckt werden kann. Bisher hat nur ein namhafter Hersteller die WPS Prozedur mit wechselnden PINs implementiert und ist somit nur noch für die oben beschriebene Offline Attacke anfällig. Außerdem kann argumentiert werden, dass die Eingabe eines WPA/WPA-2 Passworts auch für einen Laien nur unwesentlich schwieriger ist, als die Eingabe einer 8-stelligen WPS PIN, die nur aus Zahlen besteht. Aus diesen Gründen ist es ratsam, WPS grundsätzlich im Access Point zu deaktivieren.

4.7.9 WPA3 Personal Mode Authentication

Wie zuvor besprochen verwendet WPA-2 PSK ein Nutzerpassword als Basis für die Authentifizierung und Verschlüsselung zwischen dem Wi-Fi Access Point und den Clients. Da in vielen Netzwerken kurze und schwache Passwörter verwendet werden, ist es mittlerweile recht einfach geworden, mit offline Brute-Force Angriffen das Passwort zu erraten. Dazu ist keine aktive Interaktion mit dem Netzwerk nötig. WPA-3 löst dieses Problem mit einem Authentifizierungsansatz, der als „Simultaneous Authentication of Equals“ (SAE) bezeichnet wird. SAE basiert auf dem Diffie Hellman Elliptic Curve

Public/Private Key Algorithmus, der auch bei HTTPS für die Schlüsselerzeugung verwendet wird. Die mathematischen Details sind in RFC 7664¹⁶ zu finden.

Die Prozedur zum Erstellen der Keys für die Authentifizierung und Verschlüsselung der Verbindung beginnt durch die Erzeugung je einer Zufallszahl auf dem Access Point und dem Client (X und Y). X und Y bleiben geheim und werden zu keinem Zeitpunkt über die Radioschnittstelle ausgetauscht. In Kombination mit dem Wi-Fi Passwort, das auf beiden Seiten identisch ist, werden X und Y dann als Eingabeparameter für eine mathematische Funktion auf jeder Seite verwendet, um die öffentlichen Werte A und B zu generieren. Diese werden dann der jeweiligen Gegenstelle mitgeteilt. Ein Angreifer kann A und B zwar mitlesen, aber Aufgrund des zur Erzeugung verwendeten Algorithmus ist es nicht möglich, mit diesen Werten auf X, Y oder das Wi-Fi Passwort zu schließen. Auf einem Gerät wird dann X und B verwendet, um das gemeinsame Secret S zu berechnen. Auf dem anderen Gerät werden dazu Y und A verwendet. Wurde auf beiden Seiten das gleiche Passwort verwendet, ist auch das berechnete Secret S identisch. Das bedeutet, dass ohne Wissen um die Zufallszahl eines Gerätes es einem Angreifer nicht möglich ist, aus A und B das Secret S zu berechnen.

Beide Seiten verwenden dann S als Grundlage, einen symmetrischen Sitzungsschlüssel (Session Key) und andere kryptografische Parameter zu berechnen, die dann mit einer EAPOL Nachricht, die schon von WPA-2 bekannt ist, ausgetauscht werden. Im Unterschied zu WPA-2, wo das Wi-Fi Password verwendet wird, wird bei WPA-3 der geheime S Parameter für die Generierung der weiteren Parameter verwendet. Der Schlüssel ist symmetrisch, da auf beiden Seiten der gleiche Schlüssel verwendet wird. Das bedeutet, dass die Verschlüsselung und die Entschlüsselung mit dem gleichen Schlüssel durchgeführt werden.

Der Austausch von A und B wird als SAE „Commit“ Phase bezeichnet und benötigt eine Nachricht von jeder Seite. In einem zweiten Schritt überträgt dann jede Seite einen Hashwert, der auf S basiert. Beide Seiten können dann überprüfen, ob das gleiche Passwort verwendet wurde. Dieser Schritt wird als SAE „Confirm“ Phase bezeichnet. Somit werden 4 Nachrichten ausgetauscht, die die Open System Authentication Phase ersetzt, die bisher zwei Nachrichten benötigte. Danach wird der Verbindungsauflauf wie bei WPA-2 mit einem Association Request und Response fortgeführt. Schließlich erfolgt dann der EAP-PSK (Encapsulated Authentication Protocol – Pre-Shared Key) Austausch. Hier werden jedoch Werte übergeben, die sich aus dem gemeinsamen geheimen S Parameter ableiten und nicht aus dem geheimen Passwort.

Dieser Ansatz hat gegenüber früheren Prozeduren folgende Vorteile:

- Ein Offline Brute Force Angriff zur Ermittlung des Passworts ist nicht möglich. Da die Zufallszahlen einem Angreifer nicht bekannt sind, kann er auch A und B nicht errechnen, auch wenn ihm das Passwort bekannt wäre.
- Ein Angreifer kann die Zufallszahlen X und Y nicht aus A, B und dem Passwort errechnen. Somit kann er auch den individuellen Session Key eines Teilnehmers nicht

berechnen, auch wenn er das Passwort kennt. Dies wird auch als Perfect Forward Secrecy (PFS) bezeichnet.

An dieser Stelle sei angemerkt, dass ein aktiver Brute Force Angriff auf ein schwaches Passwort noch immer möglich ist. Ist es zu einfach, können wenige Versuche das Passwort zu erraten, ausreichen. Aus diesem Grund gibt es zusätzliche Sicherheitsvorkehrungen, wie z. B. eine zunehmende Verzögerungszeit zwischen zwei Anmeldeversuchen, um Brute Force Attacken zu verlangsamen. Trotz alledem sollte in einem Netzwerk noch immer ein starkes Passwort verwendet werden.

4.7.10 Protected Management Frames

Eine Schwachstelle von WLAN Netzwerken in der Praxis ist, dass jedes Gerät Dis-Association Frames an andere Geräte schicken kann und diese somit aus dem Netzwerk drängen kann. Dies wurde in der Vergangenheit z. B. von einer Hotelkette ausgenutzt, um Wi-Fi Tethering zu Smartphones zu unterbinden und den Gästen somit das Hotel WLAN aufzuzwingen. Diese Praxis wurde zwar schnell durch die nationale Regulierungsbehörde verboten, zeigt aber das Missbrauchspotenzial deutlich auf. Aus diesem Grund hat die IEEE schon vor vielen Jahren die sogenannten Protected Management Frames (PMF) in 802.11w spezifiziert. In der Praxis dauerte es jedoch viele Jahre, bis diese Erweiterung auch in Produkten verwendet wurde.

In der Praxis funktioniert PMF wie folgt: Ohne die PMF Erweiterung sind WLAN Management Frames nicht geschützt, Geräte können also z. B. Dis-Association Frames fälschen. Mit 802.11w PMF werden nun nicht nur Frames mit Nutzdaten mit einem Session Key verschlüsselt, sondern auch die Management Frames. Wie in Abb. 4.25 gezeigt, informiert ein AP alle Geräte im Netzwerk über diese Fähigkeit in seinen Beacon Frames im RSN Capabilities Parameter.

Unterstützt auch das Endgerät PMF, informiert es den Access Point im Association Request Frame bei der ersten Verbindungsaufnahme. Dies geschieht durch das Setzen des PMF-Support Bits, auch wenn der Access Point PMF nicht unterstützt. Unterstützt auch der Access Point PMF, sendet das Gerät zusätzlich weitere Informationen im PMKID Parameter der Nachricht. Der Access Point verwendet dann diese Information während der WPA EAPOL Authentifizierungsprozedur, um zusätzliche Informationen für die WPA Verschlüsselung der Management Frames zu senden. Abb. 4.26 zeigt einen Vergleich zwischen dem dritten Frame während des EAPOL Austauschs mit und ohne PMF Ciphering Informationen. Links sieht man die WPA Schlüsseldaten mit einer Länge von 56 Bytes, während auf der rechten Seite mit aktiviertem PMF 88 Bytes vom AP zum Endgerät geschickt werden.

```

▼ Tag: RSN Information
  Tag Number: RSN Information (48)
  Tag length: 20
  RSN Version: 1
  ▼ Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
    Group Cipher Suite OUI: 00:0f:ac (Ieee 802.11)
    Group Cipher Suite type: AES (CCM) (4)
    Pairwise Cipher Suite Count: 1
  ▼ Pairwise Cipher Suite List 00:0f:ac (Ieee 802.11) AES (CCM)
    ▼ Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
      Pairwise Cipher Suite OUI: 00:0f:ac (Ieee 802.11)
      Pairwise Cipher Suite type: AES (CCM) (4)
    Auth Key Management (AKM) Suite Count: 1
  ▼ Auth Key Management (AKM) List 00:0f:ac (Ieee 802.11) PSK
    ▼ Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) PSK
      Auth Key Management (AKM) OUI: 00:0f:ac (Ieee 802.11)
      Auth Key Management (AKM) type: PSK (2)
  ▼ RSN Capabilities: 0x0080
    .... .... ...0 = RSN Pre-Auth capabilities: Transmitter does not support pre-authent
    .... .... ..0. = RSN No Pairwise capabilities: Transmitter can support WEP default k
    .... .... .00. = RSN PTKSA Replay Counter capabilities: 1 replay counter per PTKSA/C
    .... ..00 .... = RSN GTKSA Replay Counter capabilities: 1 replay counter per PTKSA/C
    .... .0... .... = Management Frame Protection Required: False
    .... ...1.... = Management Frame Protection Capable: True
    .... ..0.... .... = Joint Multi-band RSN: False
    .... .0. .... .... = PeerKey Enabled: False

```

A red arrow points to the 'Management Frame Protection Capable' field (bit 1).

Abb. 4.25 Ein Beacon Frame, der Endgeräten den PMF Support des APs signalisiert

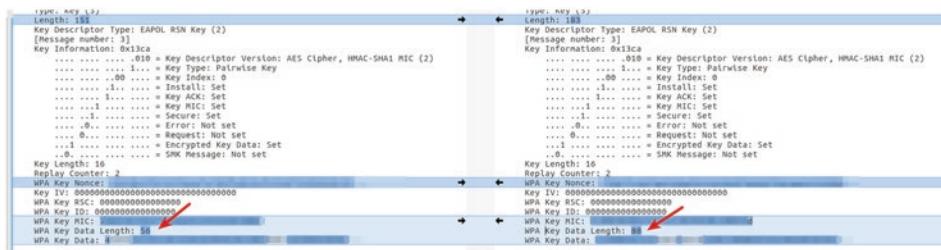


Abb. 4.26 Vergleich der Authentifizierung mit und ohne PMF Unterstützung

Möchte das Endgerät oder der Access Point die Verbindung beenden, wird ein Management Frame gesendet, der PMF geschützt ist. Abb. 4.27 zeigt einen Vergleich zwischen einem solchen Frame mit und ohne PMF Schutz. Nicht verschlüsselte Dis-Association Management Frames werden einfach ignoriert. Für die Rückwärtskompatibilität kann ein AP gleichzeitig PMF- und nicht PMF-fähige Endgeräte unterstützen. Es gibt jedoch im Standard auch die Möglichkeit, nur Geräte mit PMF-Unterstützung im Netz zuzulassen.

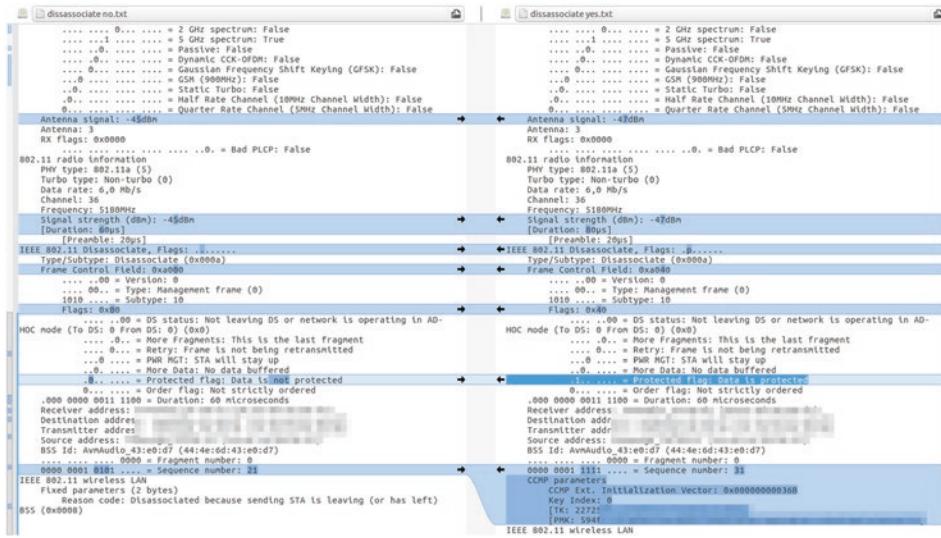


Abb. 4.27 Dis-Association Frame mit und ohne PMF-Unterstützung

4.8 IEEE 802.11e und WMM – Quality of Service

Innerhalb von nur wenigen Jahren haben Wireless LANs die Kommunikation in Büros, Arbeits- und Wohnzimmern revolutioniert. Anfangs wurden Netzwerke hauptsächlich für Anwendungen wie Web Browsing und Zugriff auf Fileserver verwendet. Diese benötigen hohe Bandbreiten, stellen jedoch sonst nur geringe Anforderungen an das Übertragungsmedium in Punkt zu Verzögerungszeit und gleich bleibende Bandbreite. Mehr und mehr werden Wireless LANs heute jedoch auch von Anwendungen wie Voice over IP oder Videostreaming verwendet, die zusätzliche Anforderungen an ein Übertragungsmedium stellen. Videostreaming beispielsweise braucht neben einer hohen Bandbreite ebenso wie Voice over IP eine garantierte Mindestbandbreite und garantierte maximale Verzögerungszeiten beim Kanalzugriff, um Bild- und Tonaussetzer zu verhindern. Dies ist mit den bisher vorgestellten Wireless LAN Standards auch problemlos möglich, solange der Datenverkehr das Netzwerk nicht an seine Leistungsgrenzen bringt. Benötigt jedoch z. B. eine Multimediaübertragung schon einen Großteil der vorhandenen Bandbreite, können weitere Endgeräte, die gleichzeitig spontan Daten z. B. von einem Fileserver oder aus dem Internet abrufen, den Datenfluss der Multimedia-Anwendung stören. Aus diesem Grund wurde mit IEEE 802.11e dem Wireless LAN Standard eine Quality of Service (QoS)-Komponente hinzugefügt. Wie auch bei anderen Erweiterungen gibt es Teile, die von einem Endgerät unterstützt werden müssen, und andere, die nur optional sind.

Um die Markteinführung von 802.11e zu beschleunigen, wurde von der Wi-Fi Alliance die Wi-Fi Multi-Media (WMM)-Spezifikation auf Basis von 802.11e entwickelt. Ist ein Access Point oder ein Endgerät WMM-zertifiziert, enthält es alle von WMM vorgeschriebenen Funktionen und ist mit WMM-zertifizierten Geräten anderer Hersteller kompatibel. Um sicherzustellen, dass QoS-Erweiterungen in Zukunft in den meisten Geräten implementiert werden, schreibt sowohl der IEEE Standard als auch die 802.11n-Zertifizierung der Wi-Fi Alliance vor, dass die WMM QoS-Erweiterungen bei 802.11n Endgeräten zum Funktionsumfang gehören müssen. Nachfolgend werden deshalb zunächst die von WMM verwendeten 802.11e Funktionalitäten beschrieben und danach optionale Komponenten, die zusätzlich unterstützt werden können.

Kern der Quality of Service-Erweiterungen ist eine Erweiterung der Distributed Coordination Function (DCF), die den Zugriff von Endgeräten auf den Übertragungskanal regelt und in Abschn. 4.5.1 beschrieben ist. DCF schreibt vor, dass ein Endgerät vor der Übertragung eines Pakets eine variable Zeit warten muss, bevor es den Funkkanal belegt, um somit Kollisionen von mehreren Endgeräten beim Kanalzugriff zu vermeiden. Die Wartezeit kann beim ersten Versuch bei 802.11b und g bis zu 31 Slots zu je 20 µs betragen. Ermittelt wird dieser Wert durch Erzeugen einer Zufallszahl zwischen 1 und 31. Sollte die Übertragung fehlgeschlagen, vergrößert sich sie Kanalzugriffswartezeit dann auf 63, 127, usw., bis maximal 1023 Slots, was 20 ms entspricht.

802.11e erweitert die DCF zur Hybrid Coordination Function (HCF). HCF umfasst zwei neue Kanalzugriffsverfahren, den Enhanced Distributed Channel Access (EDCA) und den HCF Controled Channel Access (HCCA). Außerdem ist HCF rückwärts-kompatibel zu DCF, es können sich also gleichzeitig HCF und nicht HCF-fähige Endgeräte im Netzwerk befinden. Im folgenden wird nun zunächst das EDCA-Verfahren beschrieben, das Grundlage der WMM-Spezifikation ist.

Statt allen Endgeräten und allen Datenpaketen ein gleich langes Fenster für das Ermitteln der Zufallszahl zu geben, werden vier Quality of Service-Klassen mit je einer Warteschlange eingeführt. Jeder QoS-Warteschlange werden dann unterschiedliche Wartezeitfenster für Datenpakete beim Kanalzugriff zugeordnet. WMM definiert je eine Warteschlange für Voice, Video, Background und Best Effort-Daten. Jede Klasse hat folgende variablen Parameter:

- Anzahl der Slots, die mindestens gewartet werden muss, bevor ein Datenpaket gesendet werden darf (Arbitration Interframe Space Number, AIFSN).
- Kleinstes Contention Window (CWMin), also die Anzahl der Slots, aus denen ein Zufallsgenerator eine Kanalzugriffswartezeit (Backoff) auswählen kann.
- Größtes Contention Window (CWMax), die maximale Anzahl der Slots, aus denen ein Zufallsgenerator eine Wartezeit nach fehlgeschlagenen Übertragungen auswählen kann.
- Transmit Opportunity (TXOP): Maximale Sendezeit. Granularität des Parameters ist 32 µs.
- Admission Control: Zeigt an, ob Endgeräte sich die Verwendung dieser Klasse genehmigen lassen müssen (siehe unten).

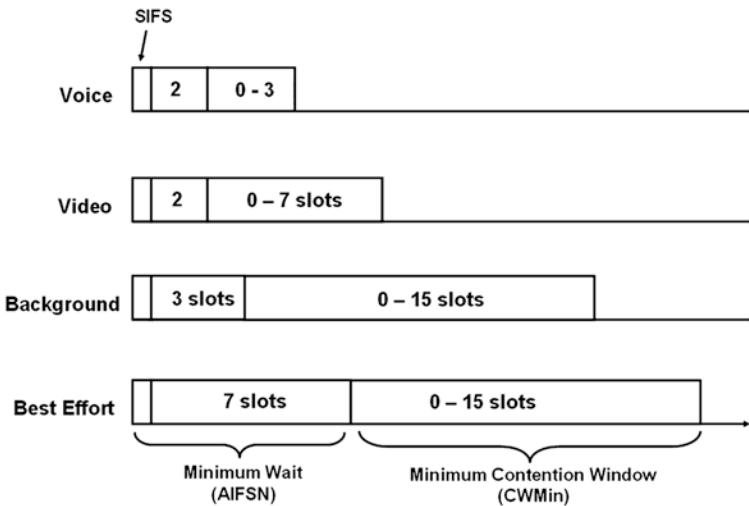


Abb. 4.28 WMM-Prioritätsklassen mit beispielhaften Werten für CWMMin, CWMax und TXOP

Abb. 4.28 zeigt, wie diese Werte in der Praxis für die unterschiedlichen Prioritätsklassen gesetzt werden können. Sprachdaten haben sehr hohe Anforderungen an gleich bleibende Verzögerungszeiten. Deshalb ist es in dieser QoS-Kategorie wichtig, dass ein Datenpaket bei der Backoff-Prozedur bevorzugt wird. Dies wird erreicht, indem die kleinste Wartezeit (AIFSN) nur 2 Slots und das Contention Window nur maximal 3 Slots lang sind. Die maximale Wartezeit beträgt somit nur 5 Slots. Somit werden diese Datenpakete immer vor Best Effort-Daten übertragen, da diese mindestens 7 Slots warten müssen, bevor das Contention Window überhaupt beginnt.

Da die Werte für CWMMin, CWMax und TXOP variabel sind und in Access Points von diversen Herstellern auch manuell gesetzt werden können, werden diese über den WMM-Parameter in Beacon Frames, sowie in Association- und Probe Response Frames den Endgeräten mitgeteilt.

Wichtig bei Quality of Service Implementierungen ist auch, dass Applikationen ihre Daten möglichst einfach und von der Art der Netzwerkschnittstelle unabhängig einer QoS-Klasse zuordnen können. Bei IP Datenpaketen geschieht dies beispielsweise über das Differentiated Services Codepoint-Feld (DSCP) im IP Header. Wenn von einer Applikation nicht speziell angefordert, ist dieses Feld auf „Default“ gesetzt. Abb. 4.29 zeigt den IP Header eines Voice over IP Datenpaket, welches dieses Feld auf „Expedited Forwarding“ gesetzt hat. Der Netzwerktreiber der Wireless LAN-Karte setzt dieses Feld dann entsprechend in den von 802.11e neu definierten QoS-Parameter auf dem Wireless LAN MAC Layer um und stellt das Datenpaket in die Warteschlange für Voice-Pakete.

Die Priorisierung von Datenpaketen auf der Luftschnittstelle wird in den meisten Netzwerkumgebungen die Quality of Service-Anforderungen erfüllen. Befinden sich jedoch zu viele Anwendungen im Netzwerk, die über EDCA eine höhere Priorität für

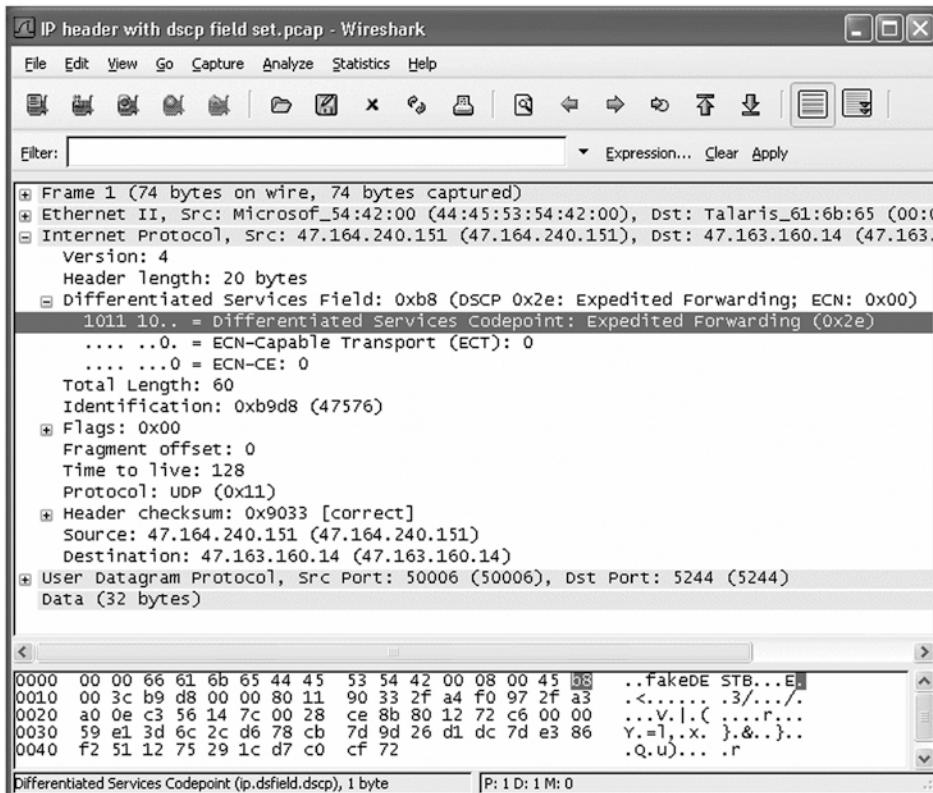


Abb. 4.29 QoS-Markierung in einem IP-Paket

ihre Datenpakete fordern, tritt erneut das schon von DCF bekannte Problem auf, dass die Anzahl der Paketkollisionen steigt. Somit steigt auch die Zugriffszeit auf das Netzwerk sprunghaft an und die Datenrate fällt. Dies kann nur verhindert werden, indem Endgeräte bzw. Anwendungen die Anforderungen eines neuen Datenstroms (z. B. erwartete Datenrate, Paketgröße, etc.) beim Access Point anmelden. Auf diese Weise kann der Access Point weiteren Teilnehmern den Zugang zu einer QoS-Klasse verweigern, sobald die Netzwerklast dies nicht mehr zulässt. Diese Endgeräte oder Applikationen müssen dann eine schlechtere QoS-Klasse verwenden. Im 802.11 Standard gibt es für diesen Zweck den optionalen Admission Control-Mechanismus. In Beacon Frames wird dazu den Endgeräten mitgeteilt, ob für eine QoS-Klasse eine Zugangskontrolle vom Access Point gefordert wird. Unterstützt ein Endgerät keine Admission Control, darf es eine QoS-Klasse, die der Access Point in Beacon Frames nur mit Admission Control zulässt, nicht verwenden.

Um einen neuen Datenstrom anzumelden, sendet ein Endgerät eine Traffic Specification (TSPEC) in einer ADDTS (Add Traffic Specification) Management-Nachricht an den Access Point. Dieser überprüft daraufhin, ob das Netzwerk den zusätzlichen

Anforderungen gerecht werden kann, und erteilt bzw. verweigert in einer Antwortnachricht eine Genehmigung. Wie der Access Point diese Prüfung durchführt, ist vom Standard nicht definiert. In der Praxis fließen in eine solche Entscheidung viele, zum Teil auch dynamische Parameter ein. Ein solch dynamischer Parameter ist z. B. die noch vorhandene Verkehrskapazität, die in einem Netzwerk momentan noch zur Verfügung steht. Dies hängt stark von den Empfangsbedingungen und Fähigkeiten der im Netzwerk befindlichen Endgeräte ab.

Neben Quality of Service-Funktionalitäten führt die 802.11e-Erweiterung auch eine Reihe von optionalen Funktionalitäten ein, um die Kapazität der Luftschnittstelle besser zu nutzen. Wichtigste Funktionalität hierbei ist das Packet Bursting, das auch schon mit proprietären Erweiterungen des 802.11g Standards implementiert wurde und von 802.11e quasi legalisiert wird. Packet Bursting setzt voraus, dass im Sende-puffer eines Endgerätes mehrere Datenpakete auf die Übertragung warten. Statt nach dem Acknowledgement für ein Datenpaket den DCF Backoff Mechanismus zu verwenden, sendet das Endgerät nach einem Short Interframe Space (SIFS) sofort sein nächstes Datenpaket. Zusätzlich gibt es optional noch die Möglichkeit, zwischen Sender und Empfänger einen Block Acknowledgement Mode zu vereinbaren, falls beide Geräte dies unterstützen. Statt jedes Paket einzeln zu bestätigen, sendet der Empfänger, wie in Abb. 4.30 gezeigt, zuerst eine Reihe Datenpakete und fordert dann ein Block Acknowledgement für alle Datenpakete an. Hat der Empfänger alle Datenpakete richtig empfangen, muss dieser nur eine einzige Bestätigung zurückschicken. Dies geschieht

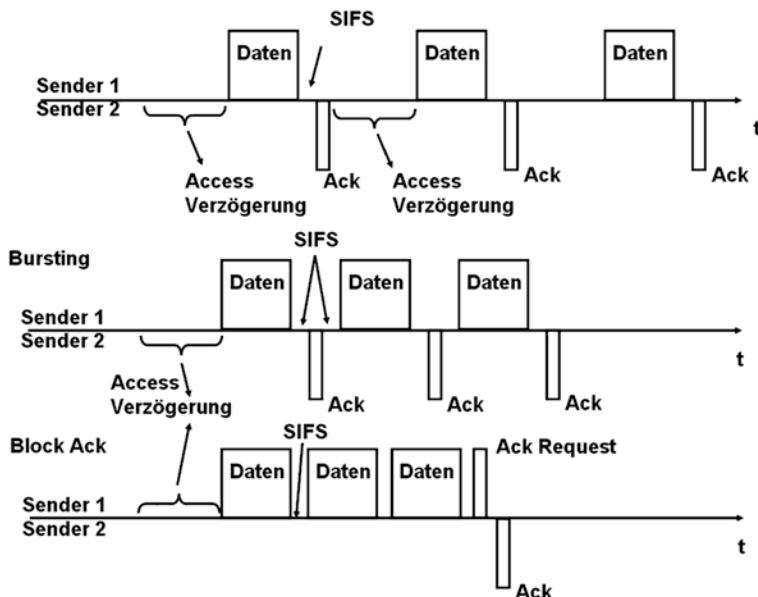


Abb. 4.30 Packet Bursting und Block Acknowledgements

entweder sofort (Immediate Block ACK) oder erst etwas verzögert (Delayed Block ACK), um dem Empfänger mehr Zeit für die Analyse der empfangenen Daten zu geben.

Ob ein Access Point den Block ACK-Mechanismus unterstützt, wird den Endgeräten über Beacon Frames im Capability Information Parameter mitgeteilt, während Endgeräte dem Access Point dies während der Association-Prozedur mitteilen. Die in Abb. 4.16 gezeigte Packet Aggregation, die mit 802.11n eingeführt wurde, kann zusätzlich zu Packet Bursting und Block ACK verwendet werden. Somit gibt es nun zahlreiche unterschiedliche Möglichkeiten, das Übertragungsmedium für die Übertragung von großen Datenmengen sehr effizient im Vergleich zum ursprünglichen Verfahren zu verwenden.

Zusätzlich zu dem in Abschn. 4.4 beschriebenen ursprünglichen Power Save (PS) Mode und dem in Abschn. 4.6.2 beschriebenen Power Save Multi Poll (PSMP) Mechanismus, der mit 802.11n spezifiziert wurde, führt der 802.11e Standard einen weiteren Power Save-Mechanismus ein, das Automated Power-Save Delivery (APSD). Auch hier gibt es wieder mehrere Optionen. Beim Unscheduled-APSD (U-APSD), der von WMM optional unterstützt wird, vereinbaren Endgerät und Access Point, dass das Endgerät in den Schlafzustand überwechselt und dass in dieser Zeit eingehende Pakete im Access Point zwischengespeichert werden. Teil dieser Vereinbarung ist auch, Pakete welcher Prioritätsklasse mit diesem Algorithmus behandelt werden und welche Pakete weiterhin mit dem normalen PS-Modus nach dem Aufwachen des Endgeräts zugestellt werden. Zusätzlich wird eine Service-Periode (SP) vereinbart, in der das Endgerät nach dem Aufwachen aktiv ist, bevor es dann automatisch wieder in den Schlafzustand wechselt.

Bei U-APSD wird keine genaue Zeit vereinbart, nach der das Endgerät wieder aktiv sein muss. Stattdessen schickt das Endgerät einen Trigger Frame an den Access Point, sobald es wieder empfangsbereit ist. Datenpakete von QoS-Klassen, für die U-APSD zuvor aktiviert wurde, werden dann automatisch innerhalb der Service-Periode zugestellt. Am Ende der Service-Periode wechselt das Endgerät wieder automatisch in den Schlafmodus. Pakete von QoS-Klassen, für die kein U-APSD aktiviert wurde, müssen weiterhin mit den für das PS-Verfahren nötigen Poll Frames einzeln angefordert werden. Ob ein Access Point U-APSD unterstützt, teilt dieser den Endgeräten im WMM-Parameter in den Beacon Frames mit. Auf der Endgeräteseite wird der U-APSD Betrieb während der Association-Prozedur über den QoS-Capability Parameter vereinbart oder später während des Betriebs über eine Traffic Specification (TSPEC)-Nachricht.

Zusätzlich spezifiziert der 802.11e Standard auch einen Scheduled-APSD (S-APSD)-Betrieb, welcher allerdings bei WMM nicht vorgesehen ist. Statt eines Trigger Frames vereinbaren hier Endgerät und Access Point ein zirkulisches Aktivitätsintervall.

Während heute meist Endgeräte über den Access Point direkt mit dem Internet kommunizieren, gibt es mehr und mehr Anwendungen im Heimbereich, wie z. B. Video Streaming zwischen einem Notebook und einem Monitor oder Fernseher, die Daten zwischen zwei Endgeräten im gleichen Netzwerk übertragen. Daten können in einem solchen Fall bisher nicht direkt zwischen den Endgeräten ausgetauscht werden, sondern müssen zunächst zum Access Point geschickt werden und von dort dann weiter zum

eigentlichen Empfänger. Da das Datenpaket in einem solchen Fall zweimal über die Luftschnittstelle übertragen werden muss, halbiert sich somit die maximale Bandbreite. Für solche Anwendungen wurde deshalb im 802.11e Standard das Direct Link-Protokoll (DLP) spezifiziert, das jedoch in der Praxis heute noch keine große Rolle spielt. Möchten zwei Endgeräte direkt miteinander kommunizieren, richtet eines der beiden Endgeräte eine Anfrage an den Access Point. Dieser leitet die Anfrage an das andere Endgerät weiter. Befindet sich dieses in Reichweite des ersten Endgerätes und unterstützt ebenfalls das Direct Link Protocol, gibt es eine positive Antwort an den Access Point zurück, der diese wiederum an das anfragende Endgerät weiterleitet. Danach können die zwei Endgeräte dann direkt Verbindung aufnehmen und fortan unter Umgehung des Access Point Datenpakete austauschen.

Der Vollständigkeit halber sei an dieser Stelle auch noch der optionale HCF Controlled Channel Access (HCCA) erwähnt. Dieser Scheduling-Algorithmus kann statt EDCA verwendet werden, ist jedoch optional und nicht Teil der WMM-Spezifikation. Somit ist es unwahrscheinlich, dass HCCA größere Verbreitung finden wird. HCCA ist im Unterschied zu EDCA ein zentraler Scheduling Algorithmus und ermöglicht dem Access Point den Kanalzugriff zu kontrollieren. Dazu sendet der Access Point Poll Frames an jedes Endgerät, das danach die Möglichkeit hat, in einem vorgegebenen Zeitfenster seine Daten zu übertragen. Da der Access Point die Poll Frames schickt, bevor ein anderes Endgerät die Möglichkeit hat, auf den Kanal zuzugreifen, wird auf diese Weise sichergestellt, dass nur Endgeräte, die mit einer ADDTS-Nachricht eine Traffic Specification TSPEC angefordert haben, auch Daten übertragen können. HCCA unterstützt auch die zuvor erwähnten Quality of Service-Klassen und kann somit, wie EDCA, Paketen mit bestimmten Quality of Service-Anforderungen Priorität einräumen.

Um ein Gefühl zu bekommen, welche der in diesem Kapitel vorgestellten Optionen in der Praxis verwendet werden, gibt es eine Anzahl von Netzwerkanalysetools, die sich für eigene Nachforschungen sehr gut eignen. Ein kostenloses und sehr leistungsfähiges Programm ist beispielsweise Wireshark, erhältlich unter <http://www.wireshark.org>. Unter Linux kann mit diesem Programm und diversen Wireless LAN-Karten die Datenübertragung in einem Wi-Fi Netzwerk aufgezeichnet und analysiert werden. Wireshark ist auch unter Windows erhältlich, für das Aufzeichnen von Wireless LAN Paketen ist jedoch ein spezieller Wireless LAN Adapter nötig. Außerdem bietet die Seite auch über das integrierte Wiki diverse Wi-Fi Traces zum Download an, die ebenfalls einen guten Einblick in die Funktionsweise des Wireless LAN Standards bieten. Eine weitere interessante Alternative zum Aufzeichnen von Wi-Fi Paketen ist die Anschaffung eines Linksys WRT54G Access Points, der sich mit einem freien Linux-Betriebssystem namens OpenWRT und Kismet zu einem ausgezeichneten Paketmonitor und Aufzeichnungsgerät umfunktionieren lässt. Weitere Information hierzu finden sich im OpenWRT Wiki auf <http://www.openwrt.org>

4.9 Fragen und Aufgaben

1. Welche Unterschiede gibt es zwischen der Ad-hoc und der BSS-Betriebsart eines Wireless LAN?
2. Welche weiteren Funktionen werden oft zusätzlich in einem Wireless LAN Access Point eingebaut?
3. Was ist ein Extended Service Set (ESS)?
4. Welche Aufgabe hat die SSID und in welchen Frames wird diese verwendet?
5. Welche Stromsparmechanismen gibt es in den Wireless LAN Standards?
6. Warum werden in einem Wireless LAN Acknowledgement Frames verwendet?
7. Aus welchen zwei Gründen wird der RTS/CTS Mechanismus bei 802.11g verwendet?
8. Warum gibt es in einem BSS-Szenario drei MAC-Adressen in einem Wireless LAN MAC Header?
9. Wie wird dem Empfänger die Datenrate des Nutzdatenpakets mitgeteilt?
10. Welche maximale Datenrate kann bei der Kommunikation zwischen zwei 802.11g-Endgeräten in einem BSS erreicht werden?
11. Welche Nachteile hat das DCF-Verfahren für Anwendungen wie Telefonie oder Video Streaming?
12. WPA-2Welche Vorteile bietet WPA3 gegenüber WPA-2?
13. Welche „High Efficiency“ Erweiterungen wurden mit 802.11ax eingeführt?
14. Wie erreicht EDCA eine Priorisierung von Sprachdaten?

Lösungen sind auf der Website zum Buch unter <http://www.cm-networks.de> zu finden.

Anmerkungen

1. IEEE, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, ANSI/IEEE Std 802.11, 1999 Edition (R2003).
2. IEEE, Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications, ANSI/IEEE Std 802.3, March 2002 Edition.
3. IEEE, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-Speed Physical Layer Extensions in the 2.4 GHz Band, ANSI/IEEE Std 802.11b, 1999 Edition (R2003).
4. IEEE, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications – Amendment 4: Further Higher Data Rate Extensions in the 2.4 GHz Band, ANSI/IEEE Std 802.11g, 2003.
5. IEEE, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications – High-Speed Physical Layer Extensions in the 5 GHz Band, ANSI/IEEE Std 802.11a, 1999.

6. IEEE, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications – Amendment: Medium Access Control (MAC) Quality of Service Enhancements, IEEE Std P802.11e/D13, January 2005.
7. IEEE, IEEE Trial-Use Recommended Practice for Multi-Vendor Access Point Interoperability via an Inter-Access Point Protocol Across Distribution Systems Supporting IEEE 802.11 Operation, IEEE Std 802.11 F, 2003.
8. IEEE, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications – Amendment 5: Spectrum and Transmit Power Management Extensions in the 5 GHz Band in Europe, IEEE Std 820.11 h, 2003.
9. IEEE, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications – Amendment 6: Medium Access Control (MAC) Security Enhancements, IEEE Std 802.11i, 2004.
10. R. Droms, RFC 2131 – Dynamic Host Configuration Protocol, RFC 2131, March 1997.
11. Wikipedia, List of WLAN Channels, http://en.wikipedia.org/wiki/List_of_WLAN_channels.
12. M. Gast (2013) 802.11ac – A Survival Guide, O'Reilly, CA, ISBN 978-1-449-34314-9.
13. M. Sauter: An 802.11ac vs. 802.11n Speed Comparison in a Real Life Scenario, October 2016
14. IEEE, Part 11: IEEE Draft Standard for Information Technology [...] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment Enhancements for High Efficiency WLAN, https://standards.ieee.org/project/802_11ax.html, 2020
15. M. Sauter, Eduroam – Wi-Fi with a Certificate and Cool Roaming Features, <https://blog.wirelessmoves.com/2016/02/eduroam-wifi-with-a-certificate-and-cool-roamingfeatures.html>.
16. D. Harkins, RFC 7664, Dragonfly Key Exchange, November 2015



Bluetooth

5

Kabelverbindungen eignen sich vor allem für den stationären Einsatz, sind aber im mobilen Umfeld nicht sehr flexibel. Die Bluetooth Funktechnologie bietet hier eine ideale Lösung für viele Anwendungsfälle. Um aufzuzeigen, welche Möglichkeiten Bluetooth bietet, gibt dieses Kapitel zunächst einen Überblick über die Eigenschaften des Systems, sowie den Aufbau und die Funktionsweise des Protokollstacks. Im weiteren Verlauf führt das Kapitel dann in das Konzept der Bluetooth Profile ein und demonstriert deren praktische Funktionsweise und große Anwendungsvielfalt.

5.1 Überblick und Anwendungen

Durch die fortschreitende Miniaturisierung finden heute zunehmend kleine elektronische Geräte Einzug in das tägliche Leben. Mit Bluetooth können diese Geräte drahtlos und ohne direkte Sichtverbindung miteinander kommunizieren. Während im letzten Jahrzehnt Bluetooth für eine Vielzahl unterschiedlicher Anwendungsfälle eingesetzt wurde, ist aktuell jedoch zu beobachten, dass sich der Einsatz dieser Technologie hauptsächlich auf die folgenden Anwendungsbereiche konzentriert:

- Kabellose Verbindung zwischen einem Smartphone und Audioendgeräten wie z. B. Kopfhörer, Freisprecheinrichtungen, Lautsprechern und Headsets für Sprachtelefonie und Musikwiedergabe.
- Austausch von Dateien (z. B. Bildern) zwischen verschiedenen Endgeräten.
- Anbindung von kabellosen Tastaturen und anderen Eingabegeräten an Notebooks, PCs und Smartphones.

Andere Anwendungen die z. B. die Weitergabe einer Internetverbindung zwischen einem Smartphone und einem Notebook, Kalendersynchronisation, Multiplayer-Spiele zwischen mobilen Endgeräten, etc., wurden mittlerweile durch andere Technologien wie z. B. Wi-Fi Tethering und Cloud Dienste abgelöst.

Da heute eine Vielzahl von unterschiedlichen Herstellern Bluetooth Geräte entwickeln, ist eine einwandfreie Interoperabilität grundlegende Voraussetzung für eine reibungslose Kommunikation. Dies wird durch den Bluetooth Standard und Interoperabilitätstests sichergestellt. Nachfolgende Tabelle zeigt die bisher erschienenen Protokollversionen. Grundlegend gilt, dass jede neue Version zur alten Version abwärts-kompatibel ist. Das bedeutet, dass ein Bluetooth 2.1 Gerät auch mit einem Bluetooth 3.0 Gerät einwandfrei zusammenarbeitet. Funktionalitäten, die mit einer neuen Version eingeführt wurden, können jedoch nicht zusammen mit älteren Geräten genutzt werden.

Version	Erschienen	Kommentar
1.0B	Dez. 1999	Erste Bluetooth Version, die aber nur von den Geräten der ersten Generation verwendet wurde
1.1	Feb. 2001	Diese Version korrigiert eine Reihe von Fehlern und Zweideutigkeiten der vorhergehenden Version des Standards (Errata List). Auf diese Weise wurde die Interoperabilität zwischen Geräten weiter verbessert
1.2	Nov. 2003	Diese Version führt einige neue Funktionalitäten ein. Die wichtigsten sind: Schnelleres Auffinden von Bluetooth Geräten im Empfangsbereich. Gefundene Geräte können jetzt auch nach der Empfangsqualität sortiert werden, siehe Abschn. 5.4.2 Schnellere Verbindlungsaufnahme, siehe Abschn. 5.4.2 Adaptive Frequency Hopping (AFH), siehe Abschn. 5.3 Verbesserte Sprachübertragung z. B. für Headsets (eSCO), siehe Abschn. 5.4.1 und 5.6.4 Verbesserte Fehlererkennung und Flusskontrolle im L2CAP Protokoll Neue Sicherheitsfunktionalität: Anonyme Verbindlungsaufnahme
2.0	2004	Enhanced Data Rate (EDR): Erweitert die Bluetooth 1.2 Spezifikation um schnellere Datenraten. Siehe Abschn. 5.2 und 5.4.1
2.1	2007	Sicherheits- und Detailverbesserungen. Die wichtigsten sind: Secure Simple Pairing: Verbesserung der Sicherheit und Vereinfachung des Pairing Prozesses. Siehe Abschn. 5.5.2 Sniff-Subrating: Weitere Energiesparoption für aktive Verbindungen mit geringem Datenaufkommen. Siehe Abschn. 5.4.2 Erroneous Data Reporting für eSCO Pakete. Siehe Abschn. 5.4.1
3.0+HS	2009	Verbesserungen bei der Sendeleistungssteuerung und Einführung des High Speed (HS) Modus, der für die Verbindlungsaufnahme Bluetooth verwendet, für die Datenübertragung jedoch einen Wi-Fi Kanal. Zwar sind heute die meisten Produkte Bluetooth 3.0 kompatibel, der optionale HS Modus für Wi-Fi Datenübertragung findet bisher jedoch kaum Verbreitung

Version	Erschienen	Kommentar
4.0	2010	Aufnahme von WiBree in den Bluetooth Standard als Bluetooth Low Energy (BLE) Option und unter der Bezeichnung „Bluetooth Smart“ vermarktet
4.1	2013	Führt folgende Verbesserungen ein: LTE Koexistenz in benachbarten Bändern. Automatisches wiederaufnehmen der Verbindung bei kurzen Signalunterbrechungen. Ein Endgerät kann gleichzeitig als Low Energy Hub und Low Energy Endgerät dienen
4.2	2014	BLE Pakete können nun statt 27 bis zu 257 Bytes Nutzdaten enthalten IPv6 über BLE mit dem Internet Protocol Support Profile (IPSP) Zusätzlicher BLE Security Mode um die Verschlüsselung schon während des Connection Setups zu starten
5.0	2016	BLE Erweiterungen: • LE 2M: Datenrate auf 2 Mbit/s erhöht • Höhere Sendeleistung, größere Reichweite
5.1	2019	Positionsbestimmung: Angle of Arrival (AoA) und Angle of Departure (AoD) Feedback. Hierzu werden mehrere Antennen beim Sender und Empfänger benötigt
5.2	2020	Enhanced Attribute Protocol (EATT) sorgt für eine bessere Zusammenarbeit, falls mehrere Applikationen gleichzeitig verwendet werden BLE Power Control: Der Empfänger kann nun die Signalstärke messen und Feedback an den Sender geben. Dieser kann dann gegebenenfalls die Leistung reduzieren und somit die Interferenz im 2.4 GHz band verringern BLE Audioübertragungen: Reduzierte Leistungsaufnahme bei Audioübertragungen gegenüber klassischem Bluetooth über neue isochrone BLE Kanäle. Neuer LC3 Audio Codec und Einführung einer Option für Audioübertragungen von einer Quelle zu mehreren Empfängern

5.2 Physikalische Eigenschaften

Bevor die nächsten Abschnitte näher auf die Funktionsweise von Bluetooth eingehen, folgt hier nun zunächst ein Überblick über die wichtigsten technischen Daten:

Die maximale Datenrate eines Bluetooth Kanals wurde in den ersten Versionen des Standards zunächst auf 780 kbit/s festgelegt. Alle Endgeräte, die direkt miteinander kommunizieren, müssen sich diese Datenrate teilen. Die maximale Datenrate für einen einzelnen Teilnehmer ist deshalb von folgenden Faktoren abhängig:

- Anzahl der Endgeräte, die untereinander gleichzeitig Daten austauschen
- Aktivität anderer Endgeräte

Die höchste Geschwindigkeit aus Sicht eines einzelnen Endgerätes kann erreicht werden, wenn nur zwei Geräte miteinander kommunizieren und nur eines der beiden Geräte eine große Datenmenge zu übertragen hat. In diesem Fall beträgt die maximal mögliche Datenrate 723 kbit/s. Nach Abzug des Overheads ergibt dies eine Datenrate von etwa 650 kbit/s. Dem anderen Endgerät bleibt dann jedoch nur eine Datenrate von etwa 57 kbit/s. Diese Situation gibt es in der Praxis z. B. bei der Dateiübertragung recht oft. Bei dieser Anwendung hat eines der beiden Endgeräte sehr viele Daten zu übertragen, während das andere nur Anfragen und Empfangsbestätigungen schickt. Abb. 5.1 zeigt die möglichen Geschwindigkeiten für dieses Szenario im linken Teil der Grafik.

Möchten beide Endgeräte möglichst schnell senden, liegt die maximal mögliche Geschwindigkeit für beide bei jeweils etwa 390 kbit/s. Abb. 5.1 zeigt diese Situation in der Mitte der Grafik.

Kommunizieren mehr als zwei Endgeräte untereinander, sinkt die maximale Datenrate pro Endgerät weiter, falls alle Endgeräte gleichzeitig mit der maximalen Geschwindigkeit senden wollen. Abb. 5.1 zeigt dies auf der rechten Seite.

Im Jahre 2004 wurde Bluetooth um das Enhanced Data Rate (EDR) Modulationsverfahren erweitert, das Datenraten von bis zu 2178 Mbit/s ermöglicht. Mehr hierzu in Abschn. 5.4.1.

Um diese Übertragungsraten zu erreichen, verwendet Bluetooth einen Kanal im 2,4 GHz ISM (Industrial Scientific and Medical) Band mit einer Bandbreite von 1 MHz. Als Modulationsverfahren wird für normale Pakete das Gaussian Frequency Shift

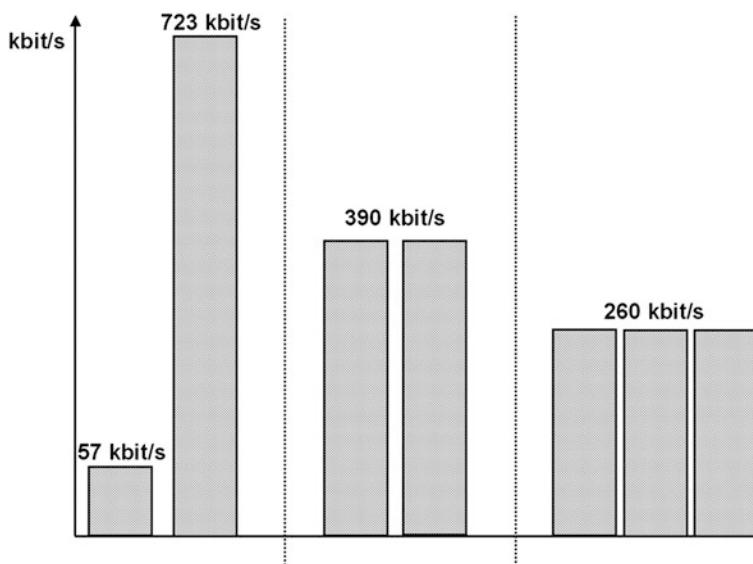


Abb. 5.1 Drei Beispiele für die maximale Geschwindigkeit in Abhängigkeit der Anzahl der Endgeräte und Teilnehmeraktivität

Keying (GFSK) Verfahren verwendet, sowie DQPSK und 8PSK für Enhanced Data Rate Pakete. Die benötigte Bandbreite für eine Bluetooth Übertragung ist verglichen mit Wireless LAN, das für einen Kanal mindestens 22 MHz belegt, sehr gering.

Um eine bidirektionale Übertragung zu ermöglichen, wird ein Übertragungskanal in Zeitschlüsse (Slots) mit einer Länge von je 625 µs unterteilt. Alle Endgeräte, die untereinander Daten austauschen, verwenden diesen Kanal abwechselnd. Dies ist der Grund für die variablen Übertragungsgeschwindigkeiten in Abb. 5.1. Hat ein Endgerät mehr zu senden, kann es bis zu 5 aufeinander folgende Zeitschlüsse belegen, bevor das Senderecht an ein anderes Endgerät übergeht. Hat dieses nur wenig zu senden, belegt es den Übertragungskanal nur für einen Zeitschlitz. Auf diese Weise ist es möglich, die Datenrate in beiden Richtungen dynamisch dem Datenaufkommen anzupassen.

Da sich Bluetooth das 2,4 GHz ISM Frequenzband mit anderen Funktechnologien wie z. B. Wireless LAN teilt, sendet Bluetooth nicht auf einer festen Frequenz, sondern wechselt nach jedem Paket die Frequenz. Ein Paket kann dabei eine Länge von einem, drei oder fünf Slots haben. Dieses Verfahren wird Frequency Hopping Spread Spectrum (FHSS) genannt. In den meisten Fällen können somit gegenseitige Störungen vermieden werden. Sollte die Übertragung in einem Zeitschlitz trotz allem einmal gestört sein, werden die Daten automatisch erneut übertragen. Bei Paketen mit einer Länge von einem Slot (625 µs) ist somit die Hopping-Frequenz 1600 Hz, werden 5 Slot Pakete verwendet, beträgt die Hopping-Frequenz 320 Hz.

Damit mehrere Bluetooth Verbindungen, die auch Piconetze genannt werden, an einem Ort gleichzeitig betrieben werden können, verwendet jedes Piconetz eine eigene Hopping-Sequenz. Für das Frequency Hopping stehen Bluetooth im ISM-Band 79 Kanäle zur Verfügung. Diese Anzahl genügt, um an einem Ort Wireless LAN Netzwerke und viele Bluetooth Netzwerke gleichzeitig und ohne wesentliche gegenseitige Beeinflussung zu betreiben.

Die gegenseitige Beeinflussung von WLAN und Bluetooth in Form einer überlagerten Übertragung auf der gleichen Frequenz bleibt gering, solange Wireless LAN und Bluetooth nur wenig ausgelastet sind. Wie im vorherigen Kapitel gezeigt wurde, werden in einem Wireless LAN bei geringer Aktivität außer kurzen Beacon Frames fast keine Datenpakete gesendet. Ist jedoch ein Wireless LAN stark ausgelastet, wird auch ständig gesendet und eine Bandbreite von 25 MHz, also fast ein Drittel der Bluetooth Kanäle, dauerhaft belegt. In einem solchen Fall ist die Anzahl der gegenseitig zerstörten Pakete recht hoch. Aus diesem Grund wurde mit Bluetooth 1.2 das Adaptive Frequency Hopping (AFH) eingeführt. Sind alle Geräte die in einem Piconetz kommunizieren zu Bluetooth 1.2 kompatibel, führt das Piconetz Master-Gerät (vgl. Abschn. 5.3) für alle Kanäle eine Kanalabschätzung (Channel Assessment) durch. Der Link Manager (vgl. Abschn. 5.4.3) legt dazu eine Liste aller Kanäle an (Channel Bitmap), die für das Frequency Hopping nicht verwendet werden sollen. Diese wird dann an alle Endgeräte des Piconetzes weitergegeben. Wie die Kanalabschätzung gemacht werden soll, wird vom Standard nicht vorgeschrieben. Mögliche Verfahren sind z. B. das Received Signal Strength Indication (RSSI) Verfahren oder der Ausschluss eines Kanals aufgrund einer

hohen Packet Error Rate (PER). Bei Endgeräten, die Bluetooth und WLAN Funk eingebaut haben, bietet der Bluetooth 1.2 Standard auch die Möglichkeit, dass das Endgerät dem Bluetooth Stack Informationen übergibt, welche Kanäle gemieden werden sollen. Dies ist möglich, da das Endgerät weiß, welcher WLAN Kanal aktuell konfiguriert ist und welche Frequenzen somit vom eingebauten Bluetooth Modul vermieden werden sollten.

Da Bluetooth speziell für kleine, mobile und batteriebetriebene Geräte konzipiert wurde, sind im Standard drei verschiedene Sendeleistungen spezifiziert. Endgeräte wie z. B. Mobiltelefone gehören meist zur Leistungsklasse (Power Class) 3 und senden mit einer Leistung von bis zu einem Milliwatt. Endgeräte der Klasse 2 senden mit bis zu 2,5 mW und Endgeräte der Klasse 1 mit bis zu 100 mW. Nur Endgeräte wie z. B. manche USB Sticks für Notebooks und PCs haben einen Sender der Leistungsklasse 1. Deren Energieverbrauch ist jedoch im Vergleich zu Leistungsklasse 3 sehr hoch und sollte deshalb nur von Geräten verwendet werden, bei denen der Energieverbrauch keine entscheidende Rolle spielt. Die Reichweiten der einzelnen Leistungsklassen sind natürlich auch dementsprechend unterschiedlich. Während Klasse 3 Endgeräte eine maximale Distanz von 10 m überbrücken und maximal durch eine Wand senden können, schaffen Klasse 1 Endgeräte bis zu 100 m und können auch mehrere Wände durchdringen. Alle Endgeräte, gleich welcher Leistungsklasse, können miteinander kommunizieren. Da jede Kommunikationsverbindung bidirektional ist, bestimmt jedoch das Endgerät mit der geringeren Leistungsklasse die maximal mögliche Reichweite.

Sicherheitsmechanismen spielen bei Bluetooth eine wichtige Rolle. So wurden in den Standard starke Mechanismen für die Authentifizierung aufgenommen. Diese stellen sicher, dass nur vom Benutzer zugelassene Geräte untereinander kommunizieren können. Auch die Verschlüsselung ist Pflichtbestandteil des Standards und muss in jedem Endgerät integriert sein. Die Verschlüsselungssequenzen sind bei Bluetooth bis zu 128 Bit lang und bilden einen wirksamen Schutz gegen fremdes Abhören.

5.3 Piconetze und das Master Slave Konzept

Bei Bluetooth werden alle Geräte die momentan miteinander kommunizieren in einem sogenannten Piconetz zusammengefasst. Die in Abb. 5.2 beschriebene Frequency Hopping Sequenz des Piconetzes wird durch die HardwareAdresse des Endgerätes berechnet, das als erstes Kontakt zu einem anderen Endgerät aufnimmt und somit das Piconetz aufbaut. Auf diese Weise ist es möglich, viele Piconetze am gleichen Ort ohne gegenseitige Beeinflussung zu betrieben.

Ein Piconetz kann ein Master- und bis zu sieben Slave Endgeräte umfassen. Dies scheint auf den ersten Blick sehr wenig zu sein. Da die meisten Bluetooth Anwendungen, wie in Abschn. 5.1 gezeigt, nur Punkt zu Punkt Verbindungen sind, ist diese Zahl aber vollkommen ausreichend. Jedes Endgerät kann Master oder Slave eines

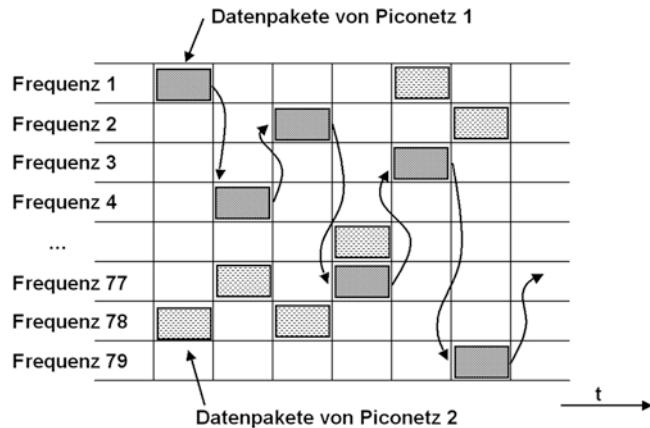


Abb. 5.2 Durch unterschiedliche Hop-Sequenzen können viele Piconetze am gleichen Ort betrieben werden

Piconetzes sein. Per Definition ist immer jenes Endgerät der Master eines Piconetzes, welches dies ursprünglich aufgebaut hat. Folgendes Beispiel verdeutlicht dieses Konzept:

Ein Anwender hat ein Bluetooth fähiges Mobiltelefon und ein Headset. Nachdem diese zwei Geräte anfangs einmal miteinander gekoppelt wurden (Pairing, siehe Abschn. 5.5.1), können diese Geräte jederzeit miteinander Verbindung aufnehmen und somit für die Dauer eines Telefonats ein Piconetz bilden. Nach dem Ende des Telefonats wird die Bluetooth Verbindung zwischen Mobiltelefon und Headset wieder beendet und das Piconetz dadurch wieder abgebaut. Bei einem ankommenden Telefongespräch nimmt das Mobiltelefon mit dem Headset Kontakt auf und ist somit der Master der Verbindung. Möchte im umgekehrten Fall der Anwender ein abgehendes Telefonat führen, betätigt er eine Taste am Headset. Das Headset nimmt daraufhin Verbindung mit dem Mobiltelefon auf. In diesem Fall ist nicht das Mobiltelefon, sondern das Headset der Master des neu aufgebauten Piconetzes. Befindet sich ein anderer Anwender in unmittelbarer Nähe, der auch gerade per Bluetooth Headset telefoniert, führt dies nicht zu Problemen, da die Frequency Hopping Sequenzen der beiden Piconetze unterschiedlich sind. Durch die ursprüngliche Kopplung von Headset und Mobiltelefon ist auch sicher gestellt, dass jedes Headset sein eigenes Mobiltelefon findet und auch nur mit diesem kommunizieren darf.

Der Master eines Piconetzes hat die Kontrolle, wer zu welchem Zeitpunkt Daten auf dem Kanal übertragen darf. Um einem Slave Endgerät das Senderecht zu erteilen, schickt ihm der Master ein Datenpaket. Das Slave Endgerät wird über eine 3-Bit Adresse im Header des Datenpaketes identifiziert, die ihm bei der ersten Kontaktaufnahme zugewiesen wurde. Das Datenpaket des Masters kann je nach Datenaufkommen 1–5 Slots lang sein. Hat der Master keine Daten für den Slave, sendet er ein leeres Paket. Unabhängig, ob das Paket Nutzdaten enthält oder nicht, übergibt der Master dem Slave

auf diese Weise implizit das Senderecht. Der Slave kann dann in den nächsten 1–5 Slots ein Antwortpaket zurückschicken. Bei Bluetooth 1.1 antwortet der Slave auf der nächsten Frequenz in der Frequency Hopping Abfolge. Bei Bluetooth 1.2 wurde dieses Konzept leicht geändert, der Slave antwortet hier auf der zuvor vom Master verwendeten Frequenz. Hat der Slave keine Daten für den Master, antwortet er trotzdem mit einem leeren Paket als Empfangsbestätigung für das zuvor vom Master eingegangene Paket. Nach spätestens 5 Slots geht das Senderecht wieder automatisch an den Master über, auch wenn der Slave noch weitere Daten in seinem Sendepuffer hat. Danach kann der Master entscheiden, ob er wieder diesem, oder einem anderen Slave das Senderecht erteilt. Empfing der Master in den letzten Datenpaketen keine Nutzdaten und ist auch sein Sendepuffer leer, kann er eine Sendepause von bis zu 800 Slots einlegen, um damit Strom zu sparen. Da ein Slot eine Dauer von $625 \mu\text{s}$ hat, entsprechen 800 Slots einer Sendepause von 0,5 s (Abb. 5.3).

Da ein Slave nicht vorhersehen kann, zu welchem Zeitpunkt Datenpakete des Masters eingehen, kann er keine Verbindungen zu weiteren Geräten aufnehmen. In manchen Fällen ist es deshalb notwendig, dass Master und Slave ihre Rollen tauschen können. Diese Funktion ist z. B. notwendig, wenn ein Smartphone mit einem PC Kontakt aufgenommen hat, um mit ihm Daten zu synchronisieren. Da das Smartphone die Verbindung zum PC aufgebaut hat, ist er der Master des Piconetzes. Während die Verbindung besteht, möchte der Nutzer des PCs jedoch ein Bild von einem weiteren Smartphone zu sich übertragen und muss deshalb zusätzlich eine Verbindung zu Smartphone 2 herstellen. Dies ist aber nur möglich, wenn Smartphone 1 (Master) und PC (Slave) die Rollen im Piconetz tauschen. Diese Prozedur wird auch Master–Slave Role Switch genannt. Nach dem Rollentausch ist der PC der Master des Piconetzes zwischen ihm und Smartphone 1. So ist es ihm möglich, zusätzlich Kontakt zu Smartphone 2 aufzunehmen, während die Datenübertragung mit Smartphone 1 noch läuft. Durch die Kontaktaufnahme mit Smartphone 2 und der Übertragung des Bildes verringert sich jedoch die Datenrate zwischen PC und Smartphone 1.

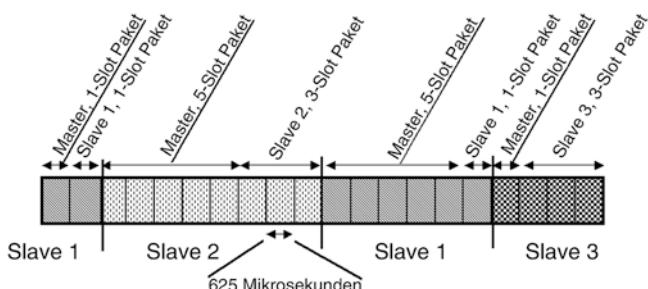


Abb. 5.3 Kommunikation zwischen einem Master und drei Slave Endgeräten

5.4 Der Bluetooth Protokoll Stack

Abb. 5.4 zeigt die unterschiedlichen Schichten des Bluetooth Protokoll Stacks und dient den nachfolgenden Unterkapiteln als Referenz. Die einzelnen Bluetooth Protokollschichten halten sich nur lose an das 7 Schichten OSI Modell, da manche Bluetooth Layer Aufgaben aus unterschiedlichen OSI Schichten übernehmen.

5.4.1 Der Baseband Layer

Die Eigenschaften der physikalischen Schicht, also der Radioübertragung, wurden im vorhergehenden Abschnitt schon beschrieben. Auf den Eigenschaften des physikalischen Kanals setzt dann der Baseband Layer auf, der typische Aufgaben eines Layer 2 Protokolls wie z. B. das Framing von Datenpaketen übernimmt. Für die Datenübertragung bietet der Baseband Layer drei unterschiedliche Frametypen:

Für die Paketdatenübertragung werden bei Bluetooth Asynchronous Connection-Less (ACL) Pakete verwendet. Wie in Abb. 5.5 gezeigt, besteht ein ACL Paket aus einem

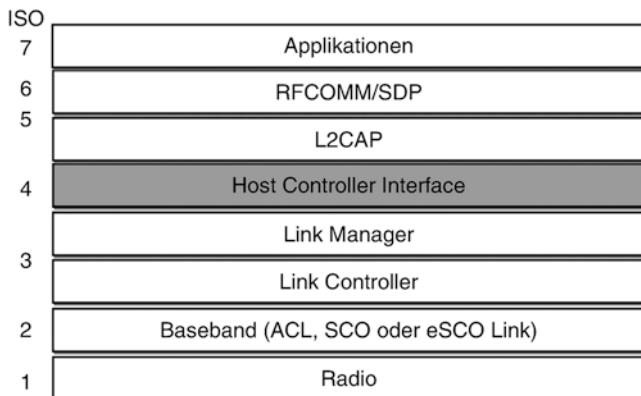


Abb. 5.4 Der Bluetooth Protokoll Stack

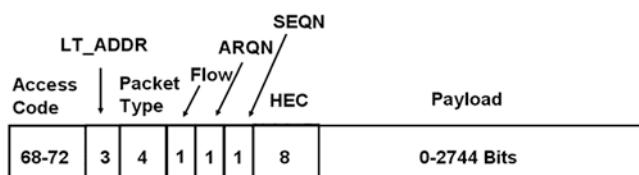


Abb. 5.5 ACL Paket

68–72 Bit langen Access Code, einem 18 Bit Header und einem 0–2744 Bit langen Feld für die eigentlichen Nutzdaten (Payload).

Vor der Übertragung werden die 18 Header-Bits noch durch einen Forward Error Correction Algorithmus in 54 Bits kodiert (1/3 FEC). Dies stellt sicher, dass Übertragungsfehler in den meisten Fällen korrigiert werden können. Je nach Größe des Nutzdatenfeldes benötigt ein ACL Paket 1, 3 oder 5 Slots zu je 625 µs Dauer.

Der Access Code am Anfang des Pakets dient in erster Linie zur Identifikation des Piconetzes, zu dem das aktuelle Paket gehört. Erzeugt wird der Access Code deshalb aus der Geräteadresse des Piconet Masters. Der eigentliche Header des ACL Pakets besteht aus einer Reihe von Bits, die folgende Funktionen haben: Die ersten drei Bits des Headers ist die Logical Transfer Address (LT_ADDR) eines Slaves, die der Master bei der Verbindungsaufnahme zuweist. Über die 3 Bit lassen sich insgesamt bis zu 7 Slaves adressieren.

Daran anschließend folgt der Pakettyp mit 4 Bits, der den Aufbau des restlichen Pakets näher beschreibt. Nachfolgende Tabelle zeigt die unterschiedlichen Möglichkeiten für ACL Pakete. Neben der Anzahl der Slots eines Paketes, unterscheiden sich die Pakettypen auch in der Anwendung einer Forward Error Correction (FEC) für den Nutzdatenteil. Diese ermöglicht es auf der Empfängerseite, Übertragungsfehler zu korrigieren. Nachteil ist jedoch, dass die Anzahl der Nutzdatenbits pro Paket reduziert wird. Mit einer 2/3 FEC wird für zwei Nutzdatenbits ein Bit für die Fehlerkorrektur hinzugefügt. Statt zwei Bits werden dann drei Bits übertragen (2/3). Außerdem wird bei ACL Paketen grundsätzlich eine CRC Checksumme berechnet, um Fehler erkennen zu können.

Paket-typ	Anzahl Slots	Linktyp	Payload (Bytes)	FEC	CRC
0100	1	DH1	0–27	Nein	Ja
1010	3	DM3	0–121	2/3	Ja
1011	3	DH3	0–183	Nein	Ja
1110	5	DM5	0–224	2/3	Ja
1111	5	DH5	0–339	Nein	Ja

Um einen Empfangspufferüberlauf zu vermeiden, kann ein Gerät über das Flow Bit seiner Gegenstelle signalisieren, für den Moment keine weiteren Daten zu senden.

Über das ARQN Bit teilt ein Endgerät seiner Gegenstelle mit, ob das zuvor gesendete Paket korrekt empfangen wurde. Ist dieses Bit nicht gesetzt, sendet die Gegenstelle das zuvor übertragene Paket erneut.

Um auch den kompletten Verlust eines Pakets erkennen zu können, folgt als nächstes Feld im ACL Header das Sequence (SEQN) Bit. Dieses wird bei jeder Übertragung eines neuen Pakets auf den jeweils anderen Bitwert gesetzt. Werden zwei aufeinander folgende Pakete mit identischem SEQN Bit empfangen, bedeutet dies für Endgerät-2, dass sein letztes Paket Endgerät-1 nicht erreicht hat und Endgerät-1 daraufhin sein Paket wiederholt

hat. Endgerät-2 wiederholt daraufhin sein Paket mit Empfangsbestätigung zu Endgerät-1 und ignoriert alle Pakete, bis wieder ein Paket mit korrektem SEQN Bit von Endgerät-1 empfangen wird. Auf diese Weise wird sichergestellt, dass auch bei mehrfachem Paketverlust die Empfangsbestätigung trotzdem zugestellt werden kann.

Als letztes Header-Feld folgt der Header Error Check (HEC). Dieses Feld stellt sicher, dass bei falsch empfangenem Header das Paket beim Empfänger ignoriert wird.

Auf den ACL Header folgt das Payload-Feld. Dieses enthält am Anfang den Payload Header, der folgende Aufgaben erfüllt: Das erste Feld wird L_CH (Logical Channel) genannt. Es gibt an, ob das Payload Feld Nutzdaten (L2CAP Pakete, vgl. Abschn. 5.4.6) oder Signalisierungsdaten in Form einer LMP Nachricht enthält (vgl. Abschn. 5.4.3) (Abb. 5.6).

Mit dem Flow Bit kann ein voller Empfangspuffer auf der L2CAP Nutzdatenschicht gemeldet werden. Schließlich enthält der Payload Header noch ein Längenfeld. Abgeschlossen wird ein ACL Paket immer durch eine 16 Bit Checksumme.

Da bei der Übertragung von ACL Paketen keine Bandbreite garantiert werden kann, eignen sich diese nicht für die Übertragung von Echtzeitdaten wie z. B. Sprache. Für diese Anwendung gibt es auf dem Baseband Layer zusätzlich den Synchronous Connection Oriented (SCO) Pakettyp. Im Unterschied zu ACL Paketen werden SCO Pakete zwischen Master und Slave in fest vorgegebenen Intervallen übertragen. Das Intervall wurde dabei so gewählt, dass die resultierende Bandbreite genau 64 kbit/s beträgt.

Bei SCO Verbindungen ist das Slave Endgerät autonom, es sendet sein SCO Datenpaket auch dann, wenn es zuvor kein Paket vom Master erhalten hat. Dies ist bei einer SCO Verbindung problemlos möglich, da Pakete zu vordefinierten Intervallen gesendet und empfangen werden. Der Slave ist somit also nicht auf eine Sendeerlaubnis des Masters angewiesen und es ist implizit sichergestellt, dass zu dieser Zeit nur er Daten überträgt. Auf diese Weise wird erreicht, dass trotz eines nicht erhaltenen Pakets in Empfangsrichtung das eigene Sprachpaket trotzdem übertragen wird.

Der Header eines SCO Paketes entspricht dem eines ACL Paketes, die Flow, ARQN und SEQN Felder werden bei SCO Paketen jedoch nicht verwendet. Die Länge des Nutzdatenfeldes beträgt immer genau 30 Bytes. Je nach verwendetem Fehlerkorrekturverfahren entspricht dies 10, 20 oder 30 Nutzdatenbytes. Nachfolgende Tabelle gibt einen Überblick über die möglichen SCO Pakettyphen.

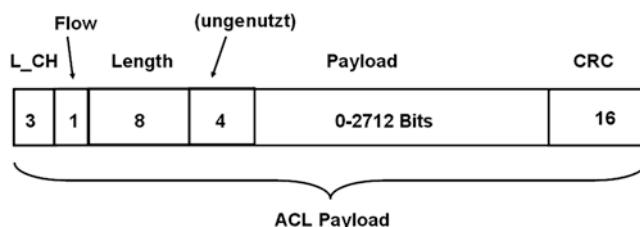


Abb. 5.6 Das ACL Payload-Feld mit Header

Paket-typ	Anzahl Slots	Linktyp	Payload (Bytes)	FEC	CRC
0101	1	HV1	10	1/3	Nein
0110	1	HV2	20	2/3	Nein
0111	1	HV3	30	Keine	Nein
1000	1	DV	10 (+ 0–9)	2/3	Ja

Die letzte Zeile der Tabelle zeigt einen Spezialpakettyp, der gleichzeitig SCO und ACL Daten enthält. Dieser Pakettyp wird verwendet, wenn neben den reinen Sprachdaten auch Steuerdaten zu übertragen sind. Wie später in Abschn. 5.6.4 im Zusammenhang mit dem Headset Profil gezeigt wird, werden zwischen einem Headset und einem Mobiltelefon nicht nur Sprachdaten, sondern auch in manchen Fällen Signalisierungsdaten (z. B. Lautstärkeregelung) übertragen. Die SCO Sprachdaten werden in einem solchen „DV“ Datenpaket dann in den ersten 10 Bytes übertragen, auf die 0–9 Bytes für den ACL Kanal folgen. Die in der Tabelle eingetragene Forward Error Correction und Checksumme wird nur für den ACL Teil verwendet. Der Standard schreibt die Verwendung eines DV Pakets nicht zwingend vor, falls Sprache und Daten gleichzeitig zwischen zwei Geräten zu übertragen sind. Eine weitere Möglichkeit ist, eigenständige ACL Pakete in den von der SCO Verbindung nicht verwendeten Slots zu senden. Dritte Möglichkeit ist, die Sprachdaten eines Slots zu verwerfen und statt des SCO Pakets ein ACL Paket zu schicken.

Da bei SCO Paketen nicht festgestellt werden kann, ob die Nutzdaten des Pakets korrekt übertragen wurden, werden bei schlechten Übertragungsbedingungen fehlerhafte Pakete an höhere Protokollsichten weitergegeben. Diese erzeugen bei der Wiedergabe der Sprache hörbare Knackgeräusche. Außerdem limitiert die maximale Geschwindigkeit eines SCO Kanals von 64 kbit/s die Anwendungsmöglichkeiten eines SCO Kanals, da z. B. Musikdaten beim Audiostreaming meist höhere Datenraten benötigen. Um diese Nachteile zu beseitigen, wurde mit Bluetooth Version 1.2 der Enhanced-SCO (eSCO) Pakettyp eingeführt. Dieser bietet folgende Vorteile:

Die Datenrate eines eSCO Kanals kann beim Aufbau der Verbindung festgelegt werden. Auf diese Weise sind konstante Datenraten bis zu 288 kbit/s in beide Richtungen möglich.

eSCO Pakete besitzen für den Nutzdatenteil eine Checksumme. Beim Auftreten eines Übertragungsfehlers kann das Paket erneut übertragen werden, falls noch genügend Zeit vor der Übertragung des nächsten regulären Pakets bleibt. Abb. 5.7 zeigt diese Situation. Bluetooth macht sich für dieses Verfahren den Umstand zunutze, dass z. B. bei einer 64 kbit/s eSCO Verbindung nur ein Bruchteil der gesamten Bandbreite des Kanals genutzt wird und somit genug Zeit für eine erneute Übertragung bleibt. Trotz der mehrfachen Übertragung eines Pakets bleibt dadurch die Datenrate konstant. Um der Gegenseite ein verlorenes oder fehlerhaftes Paket zu signalisieren, wird das von ACL Paketen bekannte Acknowledge Verfahren verwendet. Ist es bis zur Übertragung des nächsten

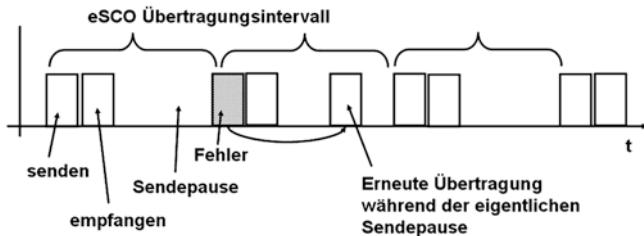


Abb. 5.7 Erneute Übertragung eines eSCO Pakets nach einem Übertragungsfehler

regulären Paketes nicht möglich ein Paket korrekt auszuliefern, wird es verworfen. Somit ist gewährleistet, dass der Datenstrom nicht ins Stocken gerät. Ab Bluetooth V2.1 kann auch ein nicht korrekt empfangenes Paket an höhere Schichten zusammen mit einer Fehlerindikation weitergegeben werden (Erroneous Data Reporting). Dies macht Sinn, wenn ein Codec kleine Übertragungsfehler selber ausgleichen kann.

Um die Übertragungsgeschwindigkeit von Bluetooth zu erhöhen, erschien 2004 die Bluetooth Version 2.0+Enhanced Data Rate (EDR). Kern von EDR ist die Verwendung von neuen Modulationsverfahren für den Nutzdatenteil eines ACL oder eSCO Paketes. Während Header und Nutzdatenteil der zuvor beschriebenen Pakete per GFSK moduliert werden, wird der Nutzdatenteil von EDR ACL oder eSCO Paketen per DQPSK oder 8DPSK moduliert. Diese Verfahren erlauben pro Übertragungsschritt die Übertragung von mehr als einem Bit. Auf diese Weise kann unter Beibehaltung der Kanalbandbreite von 1 MHz und der Slotzeit von 625 µs die Übertragungsgeschwindigkeit gesteigert werden. Um rückwärtskompatibel zu sein, wird der Header jedes Paketes weiterhin über GFSK moduliert. Somit kann der Header auch von einem Bluetooth Endgerät ohne EDR Funktionalität korrekt empfangen werden. Auch bei Wireless LAN wird dieses Verfahren verwendet, um die Kompatibilität zwischen 802.11b und den schnelleren 802.11 g und 11n Varianten zu gewährleisten. Die Beibehaltung der bisherigen Headermodulation sorgt außerdem dafür, dass auch nicht-EDR Geräte bei der Übertragung von Multislotpaketen zwischen dem Master und einem anderen Gerät weiterhin ihren Empfänger abschalten und somit Strom sparen können.

Die nachfolgende Tabelle gibt einen Überblick über alle möglichen ACL Pakettypen und die maximale Datenrate im asymmetrischen Betrieb. Asymmetrisch bedeutet, dass 5 Slot Pakete in Vorwärtsrichtung verwendet werden und 1 Slot Pakete in der Gegenrichtung. Im ersten Teil der Tabelle sind alle ACL Pakettypen aufgelistet, die von allen Bluetooth Endgeräten beherrscht werden. Im zweiten und dritten Teil der Tabelle sind dann die EDR ACL Pakettypen aufgelistet. 2-DH1, 3 und 5 werden mit DQPSK moduliert, 3-DH1, 3, 5 mit 8DPSK. Die Zahl 1, 3 oder 5 am Ende des Namens gibt die Anzahl der Slots an, die das Paket belegt.

Typ	Payload (Bytes)	Datenrate uplink (kbit/s)	Datenrate downlink (kbit/s)
DM1	0–17	108,8	108,8
DH1	0–27	172,8	172,8
DM3	0–121	387,2	54,4
DH3	0–183	585,6	86,4
DM5	0–224	477,8	36,3
DH5	0–339	723,2	57,6
2-DH1	0–54	345,6	345,6
2-DH3	0–367	1174,4	172,8
2-DH5	0–679	1448,5	115,2
3-DH1	0–83	531,2	531,2
3-DH3	0–552	1766,4	265,6
3-DH5	0–1021	2178,1	177,1

Durch die neuen Pakettypen ist es nicht mehr möglich, alle Pakettypen eindeutig über das 4 Bit lange Paket Type Feld zu identifizieren (vgl. Abb. 5.5). Die Bluetooth Spezifikation behilft sich deswegen mit folgendem Umweg: Im Grundzustand ist EDR deaktiviert. Erkennen zwei Bluetooth Endgeräte beim Einrichten einer Verbindung, dass sie beide EDR beherrschen, können die Link Manager der beiden Geräte (vgl. Abschn. 5.4.3) diese Funktionalität aktivieren und die Bitkombinationen des Paket Type Felds werden den 2-DHx und 3-DHx Typen zugeordnet.

Während EDR die DQPSK Modulation als verbindlich vorschreibt, bleibt die 8DPSK Modulation für die 3-DHx Pakete optional. Ob ein Endgerät also eine maximale Datenrate von 1448,5 oder 2178,1 Mbit/s unterstützt kann nicht von seiner EDR Fähigkeit abgeleitet werden.

Neben ACL, SCO und eSCO Paketen für die eigentliche Datenübertragung gibt es noch eine Anzahl weiterer Pakettypen, die nur für den Aufbau oder den Erhalt einer Verbindung verwendet werden:

ID Pakete werden vor dem Verbindungsaufbau von einem Gerät gesendet, um andere Geräte ausfindig zu machen. Da das Timing und die Hopping Sequenz der Gegenstelle zu diesem Zeitpunkt nicht bekannt sind, enthält ein solches Paket nur den Access Code.

Ein Frequency Hop Synchronization (FHS) Paket wird während eines Verbindungsaufbaus zwischen zwei Endgeräten in der Inquiry und Paging Phase gesendet. Inquiry und Paging werden im nächsten Unterkapitel genauer vorgestellt. Es enthält neben der 48 Bit Device Adresse des sendenden Geräts auch Timing Informationen, um die weitere Verbindungsaufnahme zu erleichtern.

NULL Pakete dienen der Empfangsbestätigung eines zuvor eingegangenen Pakets, enthalten aber keine Nutzdaten. NULL Pakete müssen nicht bestätigt werden. Somit bieten sie die Möglichkeit, den gegenseitigen Bestätigungszyklus zu unterbrechen, wenn keine Daten mehr im Sendepuffer anstehen.

Ein weiteres Spezialpaket ist das POLL Paket. Mit diesem kann überprüft werden, ob Slaves bei längerer Übertragungspause noch im Piconetz angesprochen werden können. Wie das NULL Paket enthält es keine Nutzdaten.

5.4.2 Der Link Controller

Auf dem Baseband Layer baut die Link Controller Schicht auf. Wie der Name schon andeutet, ist der Link Controller für den Aufbau, den Erhalt und den korrekten Abbau von Verbindungen zuständig. Für die Verwaltung der Verbindungen wird auf dieser Schicht ein Zustandsmodell verwendet. Für ein Gerät, das eine Verbindung zu einem anderen Gerät aufbauen möchte, gibt es folgende Zustände:

Möchte ein Endgerät bisher noch unbekannte Geräte in seiner Umgebung finden, wird der Link Controller von den höheren Protokollsichten angewiesen, in den Inquiry Zustand zu wechseln. In diesem Zustand sendet das Gerät in jedem Slot auf zwei unterschiedlichen Frequenzen ein ID Paket aus.

Alle Endgeräte, die eine Verbindungsaufnahme von unbekannten Geräten zulassen, müssen von Zeit zu Zeit in den Inquiry Scan Zustand wechseln und dort auf abwechselnden Frequenzen nach ID Paketen Ausschau halten. Die Empfangsfrequenz wird hier jedoch nur alle 1,28 s geändert. Um Strom zu sparen, oder die Verbindung mit anderen Endgeräten aufrecht zu erhalten, sucht ein Endgerät aber nicht im gesamten Intervall nach ID Paketen. Der Bluetooth Standard schlägt eine Scanzeit von 11,25 ms pro 1,28 s Intervall vor. Durch die Kombination aus schnellem Frequenzwechsel des suchenden Endgerätes und langsamem Frequenzwechsel des Ausschau haltenden Endgeräts, ergibt sich eine 90 % Wahrscheinlichkeit, dass sich die Geräte innerhalb von 10 s finden.

Um die Geschwindigkeit der Suche zu beschleunigen, wurde mit Bluetooth 1.2 der sogenannte Interlaced Inquiry Scan eingeführt. Mit dieser Methode wird statt auf einer Frequenz pro Periode auf zwei Frequenzen pro Periode nach ID Paketen gesucht. Außerdem ist es seit dieser Bluetooth Version möglich, eine Empfangsstärkemessung (RSSI, Received Signal Strength Indication) für gefundene Geräte an höhere Schichten weiterzugeben. Somit ist es möglich, die Liste der gefundenen Geräte nach der Empfangsstärke zu sortieren. Dies ist vor allem dann sinnvoll, wenn z. B. während einer Messe sehr viele Bluetooth Endgeräte in der Nähe sind und ein Nutzer seine elektronische Visitenkarte an ein Endgerät senden möchte, das sich in unmittelbarer Nähe befindet. Da dieses Gerät besser als weiter entfernte Geräte empfangen wird, erscheint es auf diese Weise ganz oben in der Liste.

Empfängt ein Endgerät ein ID Paket, sendet es ein Frequency Hop Synchronization (FHS) Paket zurück, das neben seiner Device-Adresse auch Frequency Hopping und Synchronisationsinformationen enthält.

Das suchende Endgerät hat nach Empfang des FHS Paketes die Möglichkeit, die Inquiry Prozedur fortzusetzen, um weitere Endgeräte zu finden. Alternativ kann die

Inquiry Prozedur auch beendet werden, um sofort über die nachfolgend beschriebene Paging Prozedur eine ACL Verbindung zu dem neu gefundenen Endgerät herzustellen.

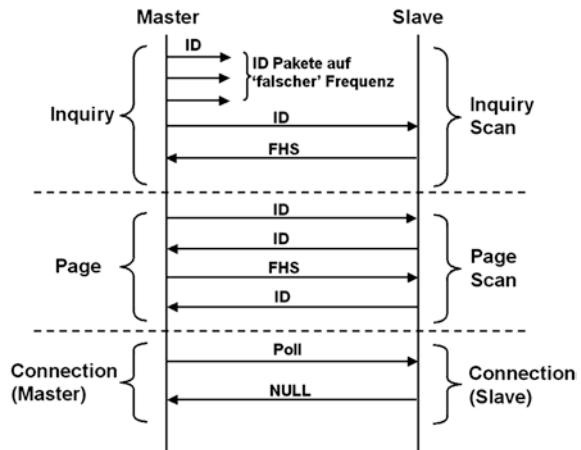
Auch Master Endgeräte, die sich schon in einer aktiven Verbindung befinden, können von Zeit zu Zeit in den Inquiry Scan Zustand wechseln. Somit sind sie auch während einer bestehenden Verbindung weiterhin für unbekannte Endgeräte sichtbar. Manche Endgeräte wie z. B. Mobiltelefone unterstützen diese optionale Funktionalität jedoch nicht.

Möchte ein Anwender gar keinen Kontakt von unbekannten Geräten zulassen, kann die Inquiry Scan Funktion abgeschaltet werden. Somit können nur noch Geräte mit der nachfolgend beschriebenen Paging Prozedur Kontakt aufnehmen, denen die Device Adresse des Endgeräts bekannt ist. Diese Einstellung ist sinnvoll, nachdem der Anwender seine Bluetooth Geräte untereinander bekannt gemacht hat (Pairing, siehe Abschn. 5.5.1) und fortan nur noch mit diesen kommunizieren will.

Um eine ACL Verbindung aufzubauen, müssen Endgeräte, denen die Device Adresse eines anderen Endgerätes schon bekannt ist, oder diese zuvor mit einer Inquiry Prozedur gefunden haben, eine Paging Prozedur durchführen. Das Paging funktioniert ähnlich dem Inquiry, ID Pakete werden in schneller Reihenfolge auf unterschiedlichen Frequenzen gesendet. Statt einer allgemeinen Adresse enthält das Paket jedoch die Geräteidentifikation der Gegenstelle, die zuvor über das FHS Paket ermittelt wurde, oder noch von der letzten Verbindung bekannt ist. Die Gegenstelle antwortet darauf ebenfalls mit einem ID Paket und gibt somit dem anfragenden Gerät die Möglichkeit, ein FHS Paket zurückzusenden, das seine Hopping Sequenz etc. enthält. Abb. 5.8 zeigt den Ablauf der Paging Prozedur und Übergang in den Connected Zustand.

Führt ein Endgerät Inquiry und Page Scans durch, und bestehen keine aktiven Verbindungen zu anderen Geräten, ist der Stromverbrauch eines Bluetooth Chips sehr niedrig. Typisch ist dann ein Energieverbrauch von weit unter einem Milliwatt. Bei

Abb. 5.8 Verbindungsauflaufbau zwischen zwei Bluetooth Geräten



Akkukapazitäten von Mobiltelefonen im Bereich von 4000–5000 mWs ist somit gewährleistet, dass die Bluetooth Funktionalität nur einen geringen Einfluss auf die Standby-Zeit des Geräts hat.

Nach erfolgreichem Paging befinden sich beide Endgeräte im Connection Active Zustand und der Datenaustausch über die neue ACL Verbindung kann beginnen.

Bei der Verbindungsaufnahme kann es vorkommen, dass der Slave der neuen Verbindung auch gleichzeitig Master einer anderen Verbindung ist, die schon vorher bestanden hat. In solchen Fällen wird von den oberen Bluetooth Protokollsichten schon beim eingehenden Paging die Verbindung nur mit der Bedingung zugelassen, sofort nach der Verbindungsaufnahme automatisch einen Master-Slave Rollentausch durchzuführen. Nur so ist es möglich, dass das Endgerät gleichzeitig mit zwei anderen Endgeräten Daten austauschen kann.

Der Stromverbrauch während einer aktiven Verbindung hängt im Wesentlichen von der Leistungsklasse des Endgeräts ab (vgl. Abschn. 5.2). Während einer aktiven Verbindung kann es jedoch auch vorkommen, dass für einige Zeit keine Daten zu übertragen sind. Gerade für Endgeräte wie Smartphones ist es in dieser Zeit sehr wichtig, möglichst wenig Strom zu verbrauchen und somit die Laufzeit des Gerätes zu erhöhen. Für solche Fälle definiert der Bluetooth Standard für den Connected Zustand drei Unterzustände:

Der erste Unterzustand ist der Connection-Hold Zustand. Um in diesen Zustand zu wechseln, einigen sich Master und Slave über die Dauer des Hold Zustandes. Danach können Sender und Empfänger für diese Zeitspanne komplett abgeschaltet werden. Nach Ende der Hold Periode wechseln Master und Slave wieder automatisch in den Connection-Active Zustand.

Wesentlich flexibler ist der Connection-Sniff Zustand. Dieser Stromsparmodus ist ideal für Verbindungen mit wenig oder zeitweise keiner Aktivität geeignet. Master und Slave einigen sich beim Aktivieren des Sniff-Modus darauf, in welchen Intervallen und für wie lange pro Intervall ein Slave den Übertragungskanal abhören soll. In der Praxis ist zu beobachten, dass der Connection-Sniff Mode für folgende Anwendungen genutzt wird:

- Bei allen Profilen bei längerer Inaktivität (z. B. 15 s): Üblich sind dann Sniff Intervalle von z. B. 2 s. Bei erneuter Aktivität wird der Sniff-Modus wieder abgeschaltet, um eine möglichst hohe Übertragungsgeschwindigkeit zu erreichen.
- Bei Human Interface Device (HID) Profilen für Tastaturen und Mäuse: Da hier die benötigte Bandbreite gering ist, können sich Verbindungen für diese Profile ständig im Sniff-Modus befinden.

Im Sniff-Modus reduziert sich der Stromverbrauch des kompletten Bluetooth Chips auf weit unter 1 mW.

Ab Bluetooth Version 2.1 gibt es zusätzlich den Sniff-Subrating Mode, um den Energieverbrauch vor allem für HID Geräte weiter zu verringern. Endgeräte im Sniff-Mode können mit diesem Mechanismus eine weitere Reduzierung des Sniff Intervalls

nach einem gewünschten Timeout aushandeln. Nach Ablauf des Timers fällt die Verbindung automatisch in den Sniff-Subrating Modus. Wird dann ein Paket empfangen, fällt die Verbindung in den normalen Sniff-Modus zurück und der Timer startet von neuem.

Um die Leistungsaufnahme noch weiter zu reduzieren, gibt es den Connection-Park Zustand. In diesem Zustand gibt der Slave seine Piconetadresse (LT_ADDR) auf und überprüft nur noch sehr selten, ob der Master die Verbindung reaktivieren möchte.

5.4.3 Der Link Manager

Die nächste Schicht des Protokoll Stacks (vgl. Abb. 5.4) ist die Link Manager Schicht. Während die zuvor besprochene Link Controller Schicht Datenpakete je nach Verbindungszustand sendet und empfängt, ist die Aufgabe des Link Managers die Einrichtung und Aufrechterhaltung von Verbindungen. Dies beinhaltet folgende Operationen:

- Aufbau einer ACL Verbindung zu einem Slave und Vergabe einer Linkadresse (LT_ADDR).
- Abbau von Verbindungen.
- Konfiguration einer Verbindung wie z. B. das Aushandeln der maximalen Anzahl von Slots von ACL oder eSCO Paketen.
- Einschalten der Enhanced Data Rate (EDR) Übertragung, falls beide Geräte diese Erweiterung unterstützen.
- Durchführung eines Master-Slave Rollentausches.
- Durchführen des in Abschn. 5.5.1 beschriebenen Pairings.
- Aktivierung und Kontrolle der Authentifizierung und Verschlüsselung, falls dies für die Verbindung von höheren Schichten gefordert wird.
- Kontrolle des mit Bluetooth 1.2 eingeführten Adaptive Frequency Hoppings (AFH).
- Management (Aktivierung/Deaktivierung) der Stromsparmodi Hold, Sniff und Park.
- Aufbau einer SCO oder eSCO Verbindung und Aushandeln der verwendeten Parameter wie z. B. die zu verwendenden Fehlerkorrekturmöglichkeiten, Datenübertragungsraten (nur eSCO), etc.

Der Link Manager führt diese Operation entweder auf Befehl von höheren Schichten aus (vgl. nächstes Kapitel), oder aufgrund von Anfragen des Link Managers der Gegenstelle. Link Manager zweier Bluetooth Endgeräte kommunizieren, wie in Abb. 5.9 gezeigt, über ACL Verbindungen mit dem Link Manager Protocol (LMP). Ob es sich bei einem eingehenden ACL Paket um Nutzdaten oder um eine LMP Nachricht handelt, erkennt der Link Manager, wie in Abb. 5.6 gezeigt, über das Logical Channel (L_CH) Feld des ACL Nutzdatenheaders.

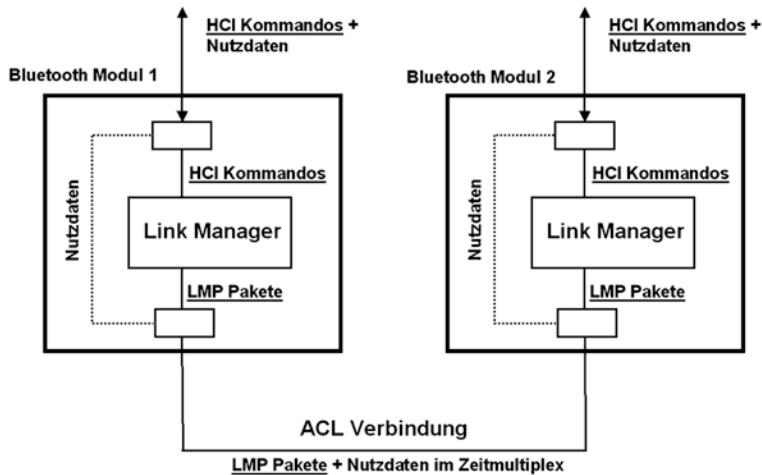


Abb. 5.9 Kommunikation zwischen zwei Link Managern per LMP

Damit eine Verbindung zu höheren Schichten nach erfolgreichem Aufbau einer ACL Verbindung hergestellt werden kann, muss zunächst der Link Manager des Geräts, das die ACL Verbindung veranlasst hat (Master), mit dem Link Manager der Gegenseite Kontakt aufnehmen. Dies geschieht mit einer LMP_Host_Connection_Request Nachricht. Danach können optionale Konfigurationsnachrichten ausgetauscht werden. Beendet wird die LMP Verbindungsphase durch gegenseitiges Senden einer LMP_Setup_Complete Nachricht. Nach diesem Schritt ist es dann möglich, Nutzdatenpakete transparent zwischen den zwei Endgeräten auszutauschen. Es können jedoch auch jederzeit innerhalb des Nutzdatenstroms weitere LMP Nachrichten eingeschoben werden, die für die am Anfang des Abschnitts beschriebenen Operationen notwendig sind.

5.4.4 Das HCI Interface

Die nächste Ebene im Bluetooth Protokollstack ist das Host Controller Interface (HCI). Bei den meisten Bluetooth Implementierungen wird dieses Interface verwendet, um das Endgerät und den Bluetooth Chip physikalisch voneinander zu trennen. Ausnahmen sind z. B. Headsets, die aufgrund ihrer physikalischen Größe und der Limitation auf Sprachübertragung alle Bluetooth Protokollsichten in einem Chip integrieren.

Über die HCI Schnittstelle können zwischen Endgerät (Host) und Bluetooth Chip (Controller) Daten und Kommandos für den Link Manager in definierten Kommandos und Nachrichtenpaketen übertragen werden. Der Bluetooth Standard sieht für das HCI Interface zwei Schnittstellentypen vor:

Für Endgeräte wie z. B. Notebooks eignet sich die USB (Universal Serial Bus) Schnittstelle am besten. Der Bluetooth Standard definiert für dieses Hardware Interface, wie HCI Kommandos und Datenpakete über USB zu übertragen sind.

Für kompakte Endgeräte, wie z. B. Smartphones kann auch ein serielles Interface verwendet werden, das UART (Universal Asynchronous Receiver and Transmitter) genannt wird. Von den verwendeten Spannungsspeichern abgesehen, ist dieses Interface mit der von PCs bekannten seriellen RS-232 Schnittstelle kompatibel. Während die RS-232 Schnittstelle jedoch auf eine Geschwindigkeit von 115 kbit/s beschränkt ist, können Daten über die UART Schnittstelle bei manchen Bluetooth Chips mit bis zu 1,5 Mbit/s übertragen werden. Dies ist auch notwendig, da die maximale Bluetooth Datenrate die Datenrate einer gewöhnlichen RS-232 Schnittstelle bei weitem übersteigt. Welche Geschwindigkeit auf der UART Schnittstelle verwendet wird, bleibt den Entwicklern des Host Endgerätes überlassen.

Auf der HCI Schnittstelle können eine Reihe unterschiedlicher Pakettypen übertragen werden. Dies sind:

- Kommandopakete (Commands), die vom Host an den Link Manager im Bluetooth Chip übertragen werden.
- Antwortpakete auf Kommandos, die der Bluetooth Controller an den Host zurück-schickt. Diese Pakete werden Events genannt. Events können auch ohne vorheriges Kommando an den Host geschickt werden, wenn z. B. ein anderes Bluetooth Gerät Kontakt aufnehmen möchte.
- Nutzdatenpakete von und zum Bluetooth Chip.

Auf der UART Schnittstelle werden die unterschiedlichen Pakettypen durch einen Header unterschieden. Das erste Byte eines Pakets gibt dabei an, um welchen Pakettyp es sich handelt. Wird USB als Übertragungsschnittstelle für das HCI Interface verwendet, werden die unterschiedlichen Pakettypen über unterschiedliche USB Endpunkte identifiziert. Eine USB Pollrate von einer Millisekunde sorgt dafür, dass Event Pakete und Nutzdatenpakete, die vom Bluetooth Chip an den Host zu übertragen sind, mit sehr kurzer Verzögerung erkannt und abgeholt werden.

Linux Distributionen für den PC unterstützen heute üblicherweise Bluetooth und bieten eine interessante Möglichkeit, über Shell Kommandos das HCI Interface zu tracen. Mit dem „hcitool con“ Kommando können beispielsweise alle Bluetooth Geräte angezeigt werden, die aktuell mit dem PC verbunden sind. Mit „hcitool info <Geräte-adresse>“ können weitere Details über ein Gerät ausgelesen werden. Das „hciconfig“ Kommando mit diversen Parametern bietet die Möglichkeit, die Konfiguration und weitere Informationen über die unterstützten Bluetoothfähigkeiten des PCs auszugeben.

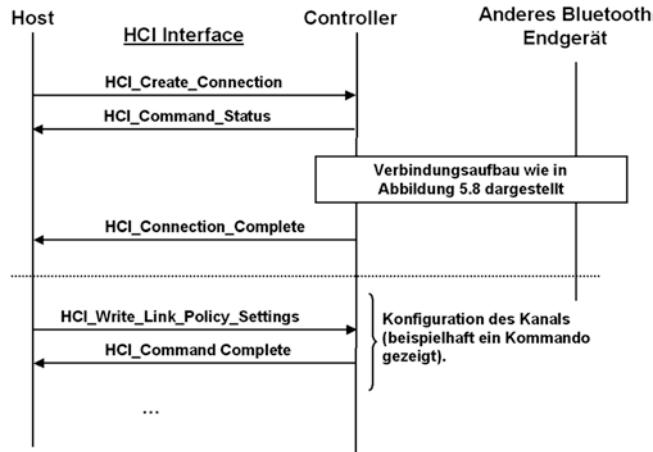


Abb. 5.10 Aufbau einer Verbindung per HCI Kommando

Das wohl interessanteste Kommando ist jedoch „hcidump -X“, mit dem der komplette Datenaustausch auf dem HCI Interface zwischen Betriebssystem und Bluetooth Chip visualisiert werden kann. Mit „hcidump -w dumpfilename“ können auch alle Pakete abgespeichert werden und dann z. B. mit Wireshark im Detail analysiert werden.

Abb. 5.10 zeigt, wie ein Bluetooth Modul über das HCI Interface veranlasst wird, eine Verbindung zu einem anderen Bluetooth Endgerät aufzubauen. Über das **HCI_Create_Connection** Kommando werden dem Bluetooth Controller alle benötigten Informationen für den Verbindungsaufbau übergeben. Der wichtigste Parameter ist die Device-Adresse des anderen Bluetooth Gerätes. Nach Erhalt des Kommandos quittiert der Controller dieses mit einer **HCI_Command_Status** Event Nachricht und startet als nächstes die Suche nach dem anderen Gerät. Der Ablauf dieser Suche ist in Abb. 5.8 zu sehen, wobei für diesen Fall jedoch die dort gezeigte Inquiry Phase entfällt, da die Bluetooth Device Adresse des anderen Gerätes schon bekannt ist. Konnte die Verbindung erfolgreich aufgebaut werden, sendet der Bluetooth Controller ein **HCI_Connection_Complete** Event zurück. Wichtigster Parameter ist ein Connection Handle, um Pakete von und zu unterschiedlichen Endgeräten unterscheiden zu können. Das Connection Handle steht über diese Zuweisung in direkter Beziehung zum L_CH Parameter eines ACL bzw. SCO Paketes.

Für die Kontrolle einer Verbindung und die Konfiguration des Bluetooth Controllers gibt es eine Vielzahl weiterer HCI Kommandos und Events. Nachfolgende Tabelle zeigt eine kleine Auswahl der Kommandos:

Kommando	Aufgabe
Setup_Synchronous_Connection	Für die Sprachübertragung (z. B. mit einem Headset) baut dieses Kommando einen SCO oder eSCO Sprachkanal auf
Accept_Connection_Request	Bei einer ankommenden Bluetooth Verbindung signalisiert der lokale Link Manager dies den höheren Schichten über ein Connection_Request Event. Möchte der Host die Verbindung zulassen, antwortet er dem Link Manager im Controller Chip mit diesem Kommando
Write_Link_Policy_Settings	Über dieses Kommando kann der Host die möglichen Verbindungszustände wie Hold, Park und Sniff erlauben oder sperren
Read_Remote_Supported_Features	Mit diesem Kommando kann ein Host den Bluetooth Controller anweisen, bei einer Gegenstelle eine Liste aller verfügbaren Bluetooth Funktionalitäten anzufordern. So kann der Host z. B. ermitteln, welche Multislot Pakettypen das andere Endgerät unterstützt, welche Stromsparmechanismen möglich sind, ob Adaptive Frequency Hopping verwendet werden kann, usw
Disconnect	Beenden einer Verbindung
Write_Scan_Enable	Mit diesem Kommando kann der Host kontrollieren, ob das Bluetooth Modul periodische Inquiry- und oder Page Scans durchführen soll. Wird beides abgeschaltet, können nur abgehende Verbindungen aufgebaut werden, das Gerät ist für andere Bluetooth Geräte unsichtbar
Write_Inquiry_Scan_Activity	Übergibt dem Bluetooth Controller Werte für die Konfiguration des Inquiry Scans wie z. B. die Größe des Inquiry Scan Zeitfensters
Write_Local_Name	Über dieses Kommando übergibt der Host einen „lesbaren“ Gerätenamen an das Bluetooth Modul. Dieser kann dann automatisch anderen Geräten übergeben werden, die nach Bluetooth Geräten suchen. So ist es möglich, dem Benutzer eine Liste mit Gerätenamen statt Bluetooth Device Adressen anzuzeigen

5.4.5 Der L2CAP Layer

Im nächsten Schritt der Verbindungsaufnahme wird über eine bestehende ACL Verbindung eine L2CAP (Logical Link Control and Adaptation Protocol) Verbindung aufgebaut. Diese Protokollsicht befindet sich über dem HCI Layer und kann mehrere logische Verbindungen zu einem Gerät über eine physikalische ACL Verbindung multiplexen. Somit kann z. B. während dem bestehen einer Bluetooth Dial-Up Verbindung zwischen einem PC und einem Mobiltelefon noch eine weitere zusätzliche logische Verbindung für die Übertragung eines Adressbucheintrages zwischen den Geräten aufgebaut

werden. Besteht zu einem Zeitpunkt noch weitere ACL Verbindungen zu anderen Geräten, kann die L2CAP Schicht auch Daten von und zu unterschiedlichen Geräte multiplexen. Ein solches Szenario ist in Abb. 5.11 dargestellt. Während einer Internet Dial-Up Verbindung über Slave 1 wird gleichzeitig noch eine Datei aus dem Speicher des Mobiltelefons zum Master übertragen, sowie ein MP-3 Datenstrom zwischen Master und Slave 2 übertragen.

Der Aufbau einer L2CAP Verbindung erfolgt über eine L2CAP_Connection_Request Nachricht. Wichtigster Parameter ist der Protocol Service Multiplexer (PSM). Dieser gibt an, an welche höhere Schicht Pakete nach erfolgreichem L2CAP Verbindungsauftakt weitergereicht werden sollen. Für die meisten Bluetooth Anwendungen wird der PSM 0x0003 verwendet, mit dem eine Verbindung zur RFCOMM Schicht hergestellt wird. Die RFCOMM Schicht stellt für Anwendungen eine virtuelle serielle Verbindung zu einem entfernten Bluetooth Endgerät her und wird in Abschn. 5.4.8 genauer beschrieben. Außerdem enthält die L2CAP_Connection_Request Nachricht eine Connection ID (CID), über die fortan alle L2CAP Pakete der Verbindung identifiziert werden. Die CID ist notwendig, da die RFCOMM Schicht von mehreren Diensten gleichzeitig verwendet werden kann und somit der PSM nur beim Verbindungsauftakt eindeutig ist. Nimmt die Gegenstelle die L2CAP Verbindung an, sendet sie ein L2CAP_Connection_Response zurück und teilt ihrerseits eine Connection ID zu, über die L2CAP Pakete in der Gegenrichtung identifiziert werden. Danach ist die Verbindung eingerichtet und kann verwendet werden. Optional gibt es jetzt die Möglichkeit, weitere Parameter für die Verbindung über das L2CAP_Configuration_Request Kommando zu übertragen. Dazu zählen, z. B. die Anzahl der erneuten Sendeversuche bei Paketverlust und die maximale Paketlänge, die von einem Gerät unterstützt wird.

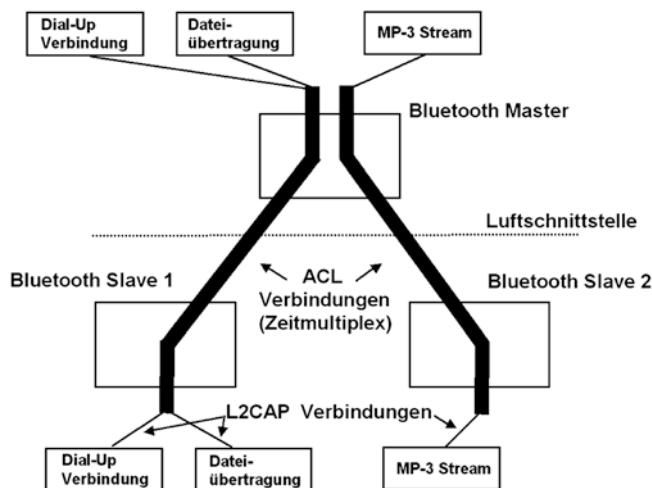


Abb. 5.11 Multiplexing verschiedener Datenströme

Eine weitere wichtige Aufgabe der L2CAP Schicht ist die Segmentierung von Datenpaketen aus höheren Schichten. Dies ist notwendig, wenn Pakete aus höheren Schichten größer als ein ACL Paket sind. Ein 5 Slot ACL Paket hat beispielsweise eine maximale Größe von 339 Bytes. Werden von der Anwendungsschicht größere Pakete angeliefert, werden diese in kleinere Stücke aufgeteilt und in mehreren ACL Paketen versandt. Im Header jedes ACL Paketes wird außerdem vermerkt, ob es den Anfang eines L2CAP Paketes darstellt, oder ein nachfolgendes Teilstück ist. Auf der Gegenseite kann dann die L2CAP Schicht mit dieser Information aus mehreren ACL Paketen wieder ein einziges Paket zusammensetzen, das an die Anwenderschicht weitergereicht wird.

5.4.6 Das Service Discovery Protocol

Theoretisch könnte nach dem Aufbau einer ACL und L2CAP Verbindung der Datentransfer zwischen zwei Endgeräten sofort aufgenommen werden. Bluetooth eignet sich jedoch für eine Vielzahl unterschiedlicher Dienste, und die meisten Endgeräte bieten mehrere Dienste gleichzeitig an. Ein Mobiltelefon beherrscht beispielsweise Dienste wie Internet Verbindung (Dial-Up Network), Dateitransfer, den Austausch von Adressen und Terminen und vieles mehr. Damit ein Bluetooth Gerät in Erfahrung bringen kann, welche Dienste andere Bluetooth Endgeräte bieten und wie diese angesprochen werden können, muss vor dem Verbindungsauflauf zum eigentlichen Dienst eine Service Datenbank befragt werden. Die Service Datenbank wird über L2CAP PSM 0x0001 angesprochen und das Protokoll zur Kommunikation wird Service Discovery Protocol (SDP) genannt. Dieser Schritt kann entfallen, wenn das Endgerät genau weiß, wie der Dienst angesprochen werden kann. Bluetooth ist jedoch sehr flexibel und erlaubt Diensten, ihre Verbindungsparameter zur Laufzeit zu ändern. Einer dieser Verbindungsparameter ist z. B. die zu verwendende RFCOMM-Kanalnummer. Mehr hierzu in Abschn. 5.4.8.

Auf Anwenderebene werden Dienste auch Profile genannt. Der Headset Dienst/das Headset Profil stellt sicher, dass ein Headset mit allen gängigen Bluetooth Telefonen zusammenarbeitet, die ebenfalls das Headset Profil unterstützen. Mehr zu Bluetooth Profilen in Abschn. 5.5.

Jeder Bluetooth Dienst hat seine eigene universelle Identifikationsnummer (Universally Unique ID, UUID), über die er in der SDP Datenbank gefunden werden kann. Der Dial-Up Server Dienst hat z. B. die UUID 0x1103. Damit sich der Bluetooth Stack eines PCs mit diesem Dienst z. B. auf einem Mobiltelefon verbinden kann, wird nach der ersten Verbindungsannahme zuerst die SDP Datenbank des Mobiltelefons nach den nötigen Einstellungen für diesen Dienst befragt. Dies geschieht über eine SDP_Service_Search_Attribute_Req Nachricht. Wichtigster Parameter, den der Client der SDP Datenbank des anderen Gerätes übergibt, ist die UUID des Dienstes. Die Datenbank liefert dann in einer SDP_Service_Search_Attribute_Response Nachricht die benötigten Parameter in Form von Records zurück. Im Falle des Dial-Up Server Dienstes liefert die

Datenbank die Information zurück, dass für diesen Dienst die L2CAP Schicht, sowie die im nächsten Unterkapitel vorgestellte RFCOMM Schicht zu verwenden sind (Abb. 5.12).

Die Service Datenbank eines Bluetooth Geräts bietet außerdem eine allgemeine Suchmöglichkeit. Diese wird von einem Endgerät verwendet, wenn es ein neues Bluetooth Gerät gefunden hat und der Benutzer wissen möchte, welche Dienste dieses Gerät anbietet. Die Nachricht für eine allgemeine Suche in der Datenbank lautet SDP_Service_Search_Request. Statt einer spezifischen UUID wie im Beispiel oben, wird die UUID der Public Browse Group (0x1002) übergeben. Die Datenbank liefert dann die UUIDs aller Dienste die es anbietet an das andere Endgerät. Die weiteren Parameter der einzelnen Dienste können nun mit SDP_Service_Search_Attribute_Request Anfragen an die Datenbank ausgelesen werden. Bei einer Anfrage liefert die Datenbank auch einen frei wählbaren Namen des angeforderten Dienstes im Klartext zurück. Auf diese Weise ist eine flexible länder- und sprachspezifische Anzeige eines Dienstnamens für den Anwender möglich. Der Name dient jedoch nur zur Benutzerinformation, der Bluetooth Stack selber identifiziert einen Dienst immer über die UUID und niemals über den Namen.

Oft werden die Informationen auch lokal auf der Anwenderschicht gespeichert, damit dem Anwender bei erneuter Nutzung eines Geräts die Liste der verfügbaren Dienste eines entfernten Geräts schneller angezeigt werden kann.

Um die Datenbankabfrage zu beenden, löst das abfragende Gerät die L2CAP Verbindung durch Senden einer L2CAP_Disconnection_Request Nachricht auf. Möchte das Gerät anschließend sofort eine Verbindung zu einem Dienst herstellen, bleibt die ACL Verbindung bestehen, und es wird sofort wieder ein L2CAP_Connection_Request Nachricht geschickt. Diese Nachricht enthält jedoch nicht die PSM ID 0x0001 für die Service Datenbank, sondern die PSM ID für die nächst höhere Schicht, die der gewünschte Dienst verwendet. Abgesehen von Sprachdiensten verwenden die meisten anderen Dienste den RFCOMM Layer, der eine virtuelle serielle Schnittstelle bietet. Dieser wird über den PSM 0x0003 angesprochen.

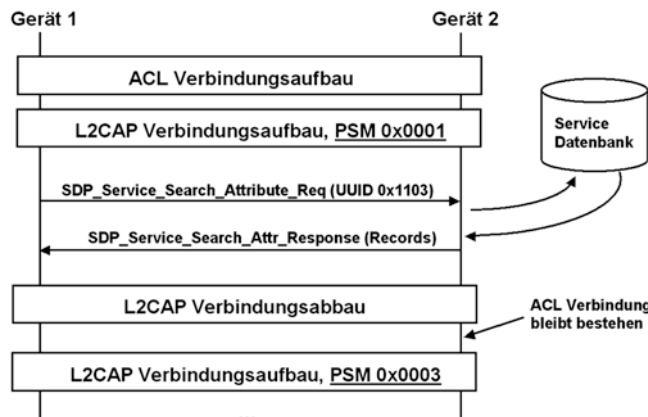


Abb. 5.12 Verbindungsauflauf zu einem Dienst mit vorheriger Datenbankabfrage

5.4.7 Der RFCOMM Layer

Wie in Abschn. 5.4.6 gezeigt, wird der L2CAP Layer verwendet, um mehrere Datenströme über eine physikalische Verbindung zu multiplexen. Die Service Datenbank ist z. B. eine Anwendung, die über den L2CAP Protocol Service Multiplexer (PSM) 0x0001 angesprochen wird. Andere Dienste könnten auf gleiche Weise über andere PSM angesprochen werden. In der Praxis verwenden jedoch einige Dienste noch einen weiteren gemeinsamen Layer, der RFCOMM genannt wird und über PSM 0x0003 angesprochen wird. RFCOMM stellt den Diensten virtuelle serielle Schnittstellen zur Verfügung und vereinfacht diesen dadurch die Datenübertragung.

Wie diese seriellen Schnittstellen verwendet werden, hängt von den übergeordneten Diensten ab. Mit dem „Serial Port“ Dienst beispielsweise wird über den RFCOMM Layer eine virtuelle serielle Schnittstelle für beliebige „nicht“ Bluetooth Anwendungen bereitgestellt. Diese unterscheidet sich aus Sicht einer Anwendung nicht von anderen seriellen Schnittstellen. Meist bekommen virtuelle serielle Bluetooth Schnittstellen vom Betriebssystem die COM-Port Nummern 3, 4, 5, 6, 7 usw. zugeteilt. Welche genau, entscheidet sich bei der Installation des Bluetooth Protokoll Stacks auf einem PC. Diese seriellen Schnittstellen wurden z. B. vor dem Aufkommen der Smartphone Wi-Fi Hotspot Funktionalität bei der Einrichtung eines neuen Modemtreibers für das DFÜ-Netzwerk verwendet. Sobald das DFÜ-Netzwerk für den Aufbau einer Internet Verbindung diesen COM-Port öffnet, wird automatisch eine Bluetooth Verbindung zur Gegenseite hergestellt. Damit diese automatische Verbindungsaufnahme funktioniert, muss zuvor über die Bluetoothsoftware diese COM-Port Nummer einmalig mit der gewünschten Gegenstelle verbunden werden.

Um Anwendungen eine komplette serielle Schnittstelle zu bieten, simuliert der RFCOMM Layer nicht nur die Sende- und Empfangsleitungen, sondern auch die Statusleitungen Request to Send (RTS), Clear to Send (CTS), Data Terminal Ready (DTR), Data Set Ready (DSR), Data Carrier Detect (CD), sowie die Ring Indicator (RI) Leitung. Bei einer physikalisch vorhandenen seriellen Schnittstelle werden diese Leitungen über einen UART (Universal Asynchronous Receiver and Transmitter) Baustein angesprochen. Aus diesem Grund simuliert die Bluetoothsoftware für den „Serial Port“ Dienst einen kompletten UART Baustein. Während ein UART Baustein die Befehle der Anwendungsschicht auf physikalische Leitungen umsetzt, sendet der virtuelle Bluetooth UART Baustein die erhaltenen Steuerkommandos und Daten in RFCOMM Paketen verpackt an den L2CAP Layer weiter.

Auch andere Dienste, wie z. B. der auch heute noch gebräuchliche Dateitransferdienst (OBEX), setzen die RFCOMM Schicht ein. Über unterschiedliche RFCOMM Kanalnummern ist es möglich, beim Verbindungsaufbau auszuwählen, welcher Dienst angesprochen werden soll. Die Kanalnummer ist Teil der Dienstbeschreibung in der Servicedatenbank. Fragt also ein anderes Gerät die Servicedatenbank eines Bluetooth Geräts nach dem OBEX Dienst, so erfährt es über die Antwort, dass dieser Dienst über

die L2CAP Schicht zu erreichen ist und als nächst höhere Schicht RFCOMM benutzt. Hieraus kann das Endgerät zunächst schließen, dass der L2CAP PSM 0x0003 zu verwenden ist, um die Verbindung zum RFCOMM Layer herzustellen (L2CAP nach RFCOMM). Außerdem entnimmt das Endgerät der OBEX Dienstbeschreibung, mit welcher RFCOMM-Kanalnummer dieser angesprochen werden kann (RFCOMM zu Anwendung). Da die RFCOMM-Kanalnummer dynamisch einem Dienst zugeordnet werden kann, ist vor der Verbindungsauftnahme deswegen immer die Service Datenbank zu befragen, um die korrekte Kanalnummer zu erhalten.

Abb. 5.13 zeigt, wie unterschiedliche Kanalschichten Datenströme multiplexen. Während der HCI Layer die Verbindung zu mehreren Geräten multiplext (Connection Handles), können über den L2CAP Layer unterschiedliche Dienste pro Gerät adressiert werden (PSM und CID). Dies wird in der Praxis verwendet, um zwischen der Service Datenbank (PSM 0x0001) und der RFCOMM-Schicht (PSM 0x0003) zu unterscheiden. Von der Service Datenbank abgesehen, verwenden die meisten Bluetooth Dienste die RFCOMM-Schicht und müssen deshalb noch zusätzlich durch unterschiedliche RFCOMM-Kanalnummern voneinander unterschieden werden.

Die RFCOMM Kanalnummer ermöglicht es außerdem, bis zu 30 RFCOMM Dienste zwischen zwei Geräten gleichzeitig zu verwenden. Somit ist es möglich, während einer Dial-Up Verbindung auch gleichzeitig Dateien mit dem Object Exchange Dienst (OBEX) zu übertragen. Da beide Dienste unterschiedliche RFCOMM Kanalnummern verwenden, können die RFCOMM Datenpakete der beiden Dienste im Zeitmultiplex übertragen werden und am Empfänger wieder dem richtigen Dienst zugestellt werden.

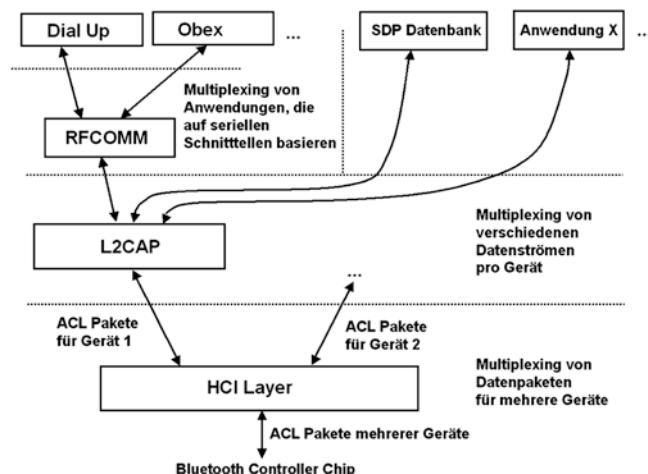


Abb. 5.13 Multiplexing auf den einzelnen Protokollsichten

5.4.8 Aufbau einer Verbindung im Überblick

Abb. 5.14 zeigt den Aufbau einer Bluetooth Verbindung durch die unterschiedlichen Schichten noch einmal im Überblick. Um Kontakt zu einer Anwendung auf einem entfernten Bluetooth Gerät aufzunehmen, baut ein Endgerät zunächst eine ACL Verbindung auf. Nach der Konfiguration des ACL Übertragungskanals wird dann über den Protocol Service Multiplexer (PSM) eine L2CAP Verbindung zur Bluetooth Service Datenbank aufgebaut, um den Service Record der Anwendung anzufordern. Dieser enthält alle Informationen für den weiteren Verbindungsauflauf, wie beispielsweise, welche Protokolle auf höheren Schichten zu verwenden sind und wie diese konfiguriert werden. Nach erfolgreicher Übertragung des Service Records wird die L2CAP Verbindung wieder abgebaut, die ACL Verbindung bleibt jedoch zwischen den zwei Geräten bestehen.

Über die ACL Verbindung wird jetzt Kontakt zur eigentlichen Anwendung aufgenommen. Dies geschieht im ersten Schritt durch Aufbau einer L2CAP Verbindung. Viele Anwendungen verwenden außerdem die RFCOMM Schicht, die serielle Schnittstellen bereitstellt. Aufgrund der beim RFCOMM Verbindungsauflauf übergebenen Kanalnummer kann der Bluetooth Stack schließlich die Verbindung zwischen dem RFCOMM Layer und der eigentlichen Anwendung, wie z. B. dem OBEX Dienst, herstellen. Wie die Anwendungsschichten der zwei Bluetooth Geräte miteinander kommunizieren, ist Sache der jeweiligen Anwendung und für alle bisher beschriebenen Schichten inklusive des RFCOMM Layers transparent. Um die Interoperabilität auch auf der Anwendungsschicht zu gewährleisten, definiert Bluetooth sogenannte Profile, die in Abschn. 5.6 beschrieben werden.

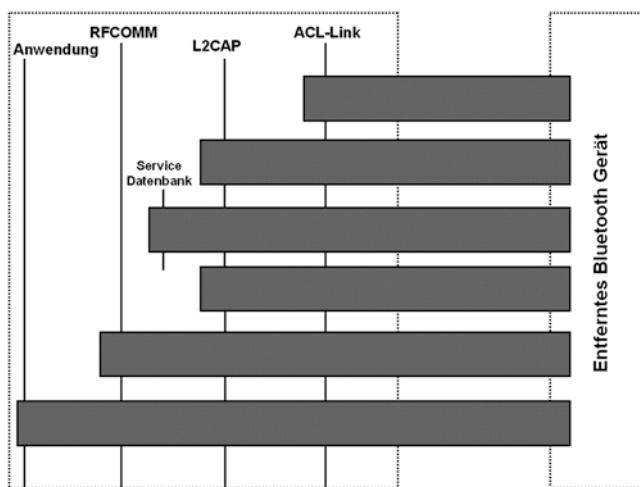


Abb. 5.14 Einzelne Stufen eines Bluetooth Verbindungsauflaufs

5.5 Bluetooth Sicherheit

Da Bluetooth Funkwellen nicht an der Wohnungstür hält machen, spezifiziert der Bluetooth Standard eine Reihe von Sicherheitsfunktionen. Alle Verfahren sind optional und müssen beim Verbindungsaufbau oder während einer laufenden Verbindung nicht unbedingt verwendet werden. Diese Entscheidung wurde bewusst getroffen, da manche Dienste keine Sicherheitsfunktionen benötigen. Welche Dienste dies sind, liegt im Ermessen des Herstellers und des Anwenders. So kann sich der Hersteller eines Mobiltelefons z. B. entscheiden, einen eingehenden Dateitransfer ohne Authentifizierung der Gegenstelle zuzulassen. Die eingehende Datei wird dann in einem Zwischenspeicher gehalten und der Benutzer kann dann auswählen, ob er die Datei speichern oder verworfen möchte. Bei anderen Diensten, wie z. B. beim früher verwendeten Modemdienst, ist es hingegen gerade umgekehrt. Hier sollte immer eine Authentifizierung beim Verbindungsaufbau erfolgen, da sonst ein fremdes Gerät z. B. eine Internetverbindung ohne Wissen des Gerätebesitzers aufbauen könnte.

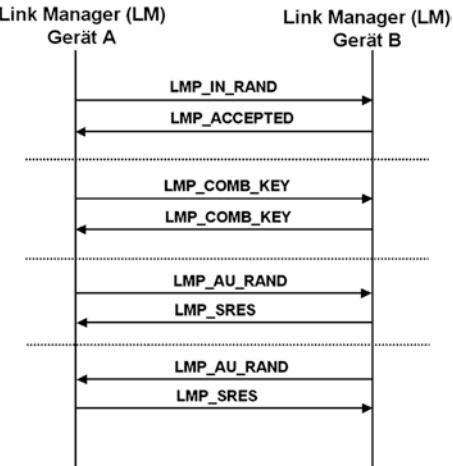
Die bei Bluetooth verwendeten SAFER+ (Secure And Fast Encryption Routine) Verschlüsselungsmechanismen wurden an der ETH Zürich entwickelt und sind öffentlich verfügbar. Bis heute wurden keine Methoden bekannt, diese zu kompromittieren. In der Praxis wurden jedoch zwischenzeitlich Schwachstellen beim einmaligen Aushandeln der Schlüssel gefunden. Diese erlauben es Angreifern, beim Abhören des gleich nachfolgend beschriebenen Pairing, die Schlüssel zu berechnen und Verbindungen dann zukünftig abzuhören. Aus diesem Grund wurden mit Bluetooth 2.1 neue Pairing Mechanismen eingeführt, die in Abschn. [5.5.2](#) beschrieben werden.

5.5.1 Pairing bis Bluetooth 2.0

Erster Schritt der Sicherheitsvorkehrungen, der einmalig durchgeführt werden muss, ist das sogenannte Pairing zweier Endgeräte. Aus Sicht des Anwenders bedeutet ein Pairing von zwei Endgeräten, dass auf beiden Endgeräten eine identische PIN Nummer eingegeben werden muss. Diese wird im Anschluss verwendet, um auf beiden Seiten einen Link Key zu generieren. Der Link Key wird in beiden Endgeräten gespeichert und kann in Zukunft für die Authentifizierung und Verschlüsselung verwendet werden. Das Pairing der zwei Endgeräte läuft, wie in Abb. [5.15](#) gezeigt, in folgenden Schritten ab:

Um das Pairing zu starten, sendet das auslösende Endgerät eine LMP_IN_RAND Nachricht über eine neue aufgebaute ACL Verbindung an das andere Endgerät. Der Inhalt der Nachricht ist eine Zufallszahl. Mit dieser wird zusammen mit der PIN und der Gerätedresse ein Initialisierungskey generiert, der K_{init} genannt wird. Da die PIN nicht zwischen den Geräten ausgetauscht wird, kann K_{init} nicht von einem dritten Gerät berechnet werden.

Abb. 5.15 Pairing zwischen zwei Bluetooth Geräten



Mithilfe von K_{init} , der auf beiden Seiten identisch ist, wird jetzt auf jeder Seite ein Teil eines Combination Keys erstellt. Dieser basiert auf K_{init} , der Geräteadresse eines der beiden Geräte und einer weiteren Zufallszahl, die aber nicht zwischen den zwei Geräten ausgetauscht wird. Im Anschluss werden die jeweils halben Combination Keys mit K_{init} noch XOR verknüpft und danach untereinander über LMP_COMB_KEY Nachrichten ausgetauscht. Die XOR Verknüpfung ist notwendig, um die zwei Combination Key Hälften nicht im Klartext über die Luftschnittstelle übertragen zu müssen.

Da K_{init} auf beiden Seiten bekannt ist, kann die XOR Verknüpfung wieder rückgängig gemacht werden und beide Seiten erhalten dann durch die Kombination der beiden Combination Key Hälften den endgültigen Link Key. Dieser ist zukünftig die Grundlage für die Authentifizierung und Verschlüsselung zwischen den zwei Geräten.

Da der mit dieser Methode generierte Link Key in beiden Endgeräten gespeichert wird, braucht das Pairing nur beim Aufbau der ersten Kommunikationsverbindung durchgeführt werden. Über die Endgeräteadresse der Gegenstelle kann bei der nächsten Verbindungsaufnahme der Link Key dann auf beiden Seiten aus der Link Key Datenbank entnommen werden. Die Authentifizierung erfolgt dann ohne zutun des Anwenders.

Um zu überprüfen, ob der Link Key auf beiden Seiten richtig erzeugt wurde, findet im Anschluss an das Pairing eine gegenseitige Authentifizierung statt. Wie diese Abläuft, wird im nächsten Unterkapitel beschrieben. Wie in Abb. 5.15 ebenfalls zu sehen ist, wird das komplette Pairing von der Link Manager Schicht in den Bluetooth Chips der beiden Endgeräte durchgeführt. Über das HCI Interface muss für die Pairing Prozedur lediglich die PIN Nummer übergeben werden.

5.5.2 Pairing ab Bluetooth 2.1 (Secure Simple Pairing)

In 2005 entdeckten Yaniv Shaked und Avishai Wool einige Schwachstellen die es ermöglichen, nach dem Abhören der Pairing Prozedur die PIN und die Link Keys zu berechnen. Dies war wohl ein wichtiger Grund, warum mit Bluetooth 2.1 der Pairing Mechanismus komplett geändert wurde. Der neue Mechanismus trägt den Namen Secure Simple Pairing und umfasst eine Reihe unterschiedlicher Pairing Protokolle für unterschiedliche Sicherheitsanforderungen:

Das Numeric Comparison Protocol: Der wichtigste Unterschied dieses Pairing Verfahrens zum bisherigen Verfahren ist, dass statt einer PIN ein Public/Private Key Verfahren zusammen mit dem Elliptic Curve Diffie-Hellmann Kryptoalgorithmus verwendet wird. Jedes Gerät hat dazu einen privaten und öffentlichen (public) Schlüssel. Beim Pairing schickten beide Endgeräte jeweils ihre öffentlichen Schlüssel zur Gegenstelle, die damit eine Zufallszahl verschlüsselt und zurückschickt. Nach Empfang der verschlüsselten Zufallszahl entschlüsseln die Endgeräte diese mit ihrem privaten Schlüssel und verwenden dann die Zufallszahlen um die Link Keys zu erzeugen. Die Ver- und Entschlüsselung funktioniert nur in eine Richtung, d. h. eine Nachricht, die mit dem öffentlichen Schlüssel chiffriert wurde, kann nur mit dem privaten Schlüssel wieder dechiffriert werden. Da die privaten Schlüssel niemals übertragen werden, kann somit kein anderes Gerät, welches das Pairing belauscht, die Nachrichten dekodieren und somit keine korrekten Link Keys erzeugen. Eine ähnliche Art der Authentifizierung findet sich auch bei Wireless LAN mit EAP-TLS im Enterprise Mode (vgl. Abschn. 1.3.7) sowie beim ersten Zugriff auf eine verschlüsselte Website mit Secure http (HTTPS, SSL/TLS).

Da sich die zwei Endgeräte bisher nicht kannten, könnte bei dieser Art des Pairing ein Angreifer ein Gerät zwischen A und B schalten und sich gegenüber A als B ausgeben und gegenüber B als A. Dies wird oft als Man in the Middle Attack (MITM) bezeichnet. Um diese Möglichkeit auszuschließen, geht das Numeric Comparison Protocol nach der Generierung der Link Keys noch einen Schritt weiter und beide Endgeräte errechnen eine 6-stellige Zahl, die dann dem Anwender gezeigt wird. Das Pairing ist erst dann abgeschlossen, wenn der Anwender auf beiden Endgeräten die Zahl bestätigt. Die Berechnungsvorschrift für die 6-stellige Zahl ist so gestaltet, dass bei einer MITM Attacke das zwischengeschaltete Endgerät diese Zahl nicht für beide Geräte berechnen kann. Die Bluetooth SIG gibt an, dass auf diese Weise die Chance eines erfolgreichen MITM Angriffs bei 1:1.000.000 liegt.

Das Just Works Protocol: Dieses Protokoll ist identisch zum Numeric Comparison Protokoll, es wird jedoch am Ende der Pairing Prozedur keine 6-stellige Zahl berechnet, die der Anwender auf beiden Endgeräten bestätigen muss. Dies bietet zwar keinen Schutz vor einem MITM Angriff, manche Endgeräte wie z. B. Headsets haben jedoch kein Display, um die 6-stellige Zahl darzustellen. Aus diesem Grund sollte ein Pairing für solche Geräte nur durchgeführt werden, wenn hinreichend sicher ist, dass kein Angreifer die Pairing Prozedur abhören und verändern kann. Da diese Schwachstelle

nur den Pairing Prozess betrifft, sind alle später aufgebauten und verschlüsselten Verbindungen trotzdem sicher, und das Just Works Protocol bietet somit für die meisten Anwendungen ausreichend Sicherheit beim Pairing. Sollte während des Pairings eine MITM Attacke erfolgreich gewesen sein, muss der Angreifer jedoch bei jeder zukünftigen Kommunikation dabei sein, da sonst der Verbindungsaufbau fehlschlägt.

Das Passkey Protokoll: Bei diesem Protokoll wird ein Passkey (PIN) für die Authentifizierung verwendet. Für den Anwender ist diese Art des Pairing identisch zum bisherigen Verfahren. Die PIN wird jedoch während des Pairings nicht wie in Abschn. 5.5.1 gezeigt verwendet, sondern es kommt wiederum zu einem Public/Private Key Austausch in Verbindung mit jeweils unabhängigen Zufallszahlen auf beiden Seiten. Für jedes einzelne Bit wird eine verschlüsselte Bestätigung, die Commitment genannt wird, auf beiden Seiten generiert. Eingangsparameter für den dazu verwendeten Algorithmus sind auf beiden Seiten beide öffentlichen Schlüssel, eine auf beiden Seiten unterschiedliche Zufallszahl und das aktuelle Bit der PIN. Im ersten Schritt tauschen beide Endgeräte das Commitment für ein Bit aus. Danach schickt Endgerät A die verwendete Zufallszahl, damit Endgerät B das Commitment über den Umkehralgorithmus überprüfen kann. War die Nachricht korrekt, schickt Endgerät B seine eigene Zufallszahl zurück, damit auch Gerät A überprüfen kann, ob das zuvor gesendete Commitment authentisch ist. Für das nächste Bit wird der Prozess in umgekehrter Richtung durchgeführt, d. h. Gerät B sendet als erstes sein Commitment. Ein Gerät in der Mitte kann bei diesem Prozess somit die Commitments nicht fälschen, da das PIN Bit erst aus dem Commitment zurückberechnet werden kann, nachdem im zweiten Schritt die Zufallszahlen ausgetauscht wurden. Da die Commitments alternierend sind, kann ein Angreifer also nur von jeder Seite ein Bit bekommen, bevor er selber zuerst ein Commitment schicken muss. Dies kann er jedoch nicht, da er nicht über das PIN Bit verfügt.

Das Out of Band Protokoll: Schließlich wurde mit Bluetooth 2.1 auch noch ein Verfahren spezifiziert, um die Authentifizierung nicht über den Bluetooth Funkkanal, sondern teilweise oder ganz über andere Übertragungswege durchzuführen. Beispielsweise kann diese Variante zusammen mit Near Field Communication (NFC) verwendet werden. Hierfür müssen sich die Geräte während des Pairings in unmittelbarer Nähe zueinander befinden, der Anwender hält die Geräte also in der Praxis zusammen. Dies schließt eine MITM Attacke aus, da ein eventueller Angreifer zwar potenziell den Nachrichtenaustausch abhören könnte, jedoch selber keine Möglichkeit hat, sich zwischen die zwei Teilnehmer zu schalten und Nachrichten zu fälschen. Der Bluetooth Standard unterstützt sowohl aktive NFC Chips, die senden und empfangen können, sowie passive NFC Chips, die nur senden können, wenn ihnen über die Antenne eine Spannung induziert wird. Dies ist notwendig, da manche Endgeräte wie z. B. Headsets keinen Platz für eine zusätzliche NFC Antenne haben. In solchen Fällen wird ein passiver NFC Chip z. B. auf dem Benutzerhandbuch oder der Verpackung angebracht. Während des Pairing Prozesses wird dann ein Bluetooth Endgerät mit aktivem NFC Chip, der sowohl senden als auch empfangen kann, an den passiven NFC Chip gehalten. Der passive NFC Chip überträgt

dann alle notwendigen Informationen um ein Pairing ohne weitere Benutzerinteraktion durchzuführen.

NFC eignet sich neben dem Pairing auch für Anwendungen, in denen bei Berührung von zwei Geräten eine Aktion durchgeführt werden soll. Ein praktisches Beispiel ist der automatische Ausdruck eines Fotos auf einem Fotodrucker, da mit einem Mobiltelefon oder einem anderen Gerät aufgenommen wurde. Der Nutzer wählt das Bild auf seinem Telefon aus und hält das Telefon dann an den Fotodrucker. Beide Geräte erkennen sich dann über ihre NFC Schnittstelle und beginnen automatisch mit der Übertragung des Bildes.

5.5.3 Authentifizierung

War das Pairing zweier Geräte erfolgreich, können sich diese fortan beim Verbindungsauftbau über den Link Key authentifizieren. Dieser Vorgang funktioniert nach dem allgemeinen Challenge/Response Verfahren, dass z. B. auch bei GSM, GRPS und UMTS verwendet wird. Für die Authentifizierung werden bei Bluetooth drei Parameter benötigt:

- Eine Zufallszahl
- Die Bluetooth Adresse des Geräts, das die Authentifizierung auslöst (BD_ADDR).
- Der 128 Bit Link Key, der beim Pairing der Geräte erzeugt wurde.

Wie in Abb. 5.16 gezeigt, schickt das auslösende Endgerät (Verifier) für die Authentifizierung die Zufallszahl an die Gegenstelle (Claimant). Der Link Manager des Claimant Endgeräts verwendet daraufhin die BD_ADDR des Verifier Endgeräts, um den Link Key für diese Verbindung über das HCI Interface vom Host anzufordern.

Mit der Zufallszahl, der BD_ADDR, sowie dem Link Key, berechnet der Link Manager des Claimant nun eine Antwort, die Signed Response* (SRES*) genannt wird. Die so berechnete SRES* schickt der Link Manager danach an das Verifier Endgerät zurück. Dieses hat die gleiche Operation ausgeführt und seine eigene SRES errechnet. Die beiden Ergebnisse können nur identisch sein, wenn der Link Key auf beiden Seiten identisch war. Da der Link Key niemals über die Luftschnittstelle übertragen wird, kann sich kein Gerät erfolgreich authentifizieren, mit dem zuvor kein Pairing durchgeführt wurde.

5.5.4 Verschlüsselung

Nach erfolgreicher Authentifizierung können beide Endgeräte jederzeit die Verschlüsselung aktivieren oder deaktivieren. Als Schlüssel wird jedoch nicht der beim Pairing erzeugte Link Key verwendet. Stattdessen wird ein auf beiden Seiten der Verbindung eigens bei der Aktivierung der Verschlüsselung generierter Ciphering Key

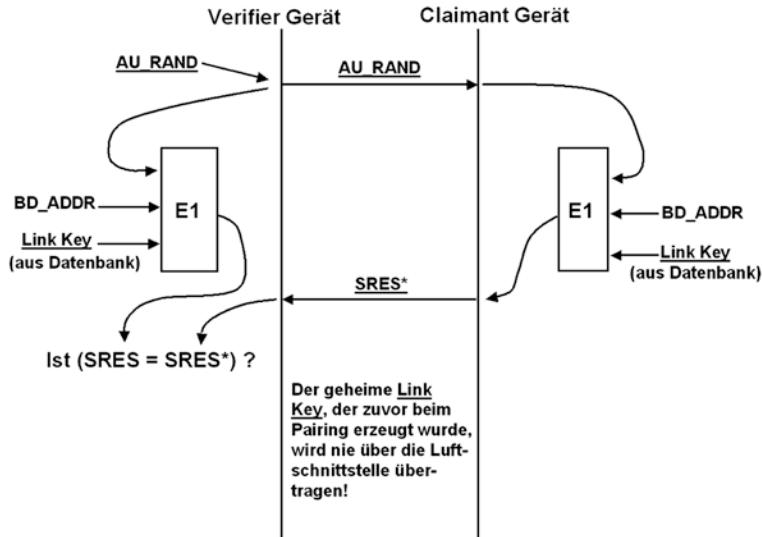


Abb. 5.16 Authentifizierung eines Endgeräts

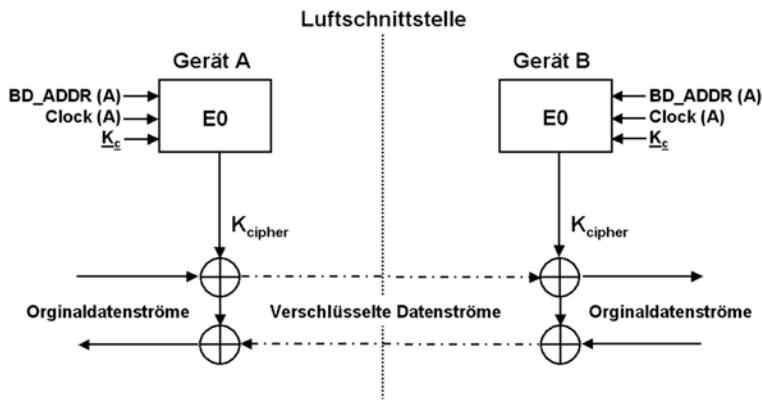


Abb. 5.17 Bluetooth Verschlüsselung mit einer Ciphersequenz

benutzt. Wichtigster Parameter für die Erzeugung des Ciphering Keys ist neben dem Link Key der Verbindung eine Zufallszahl, die beim Start der Verschlüsselung zwischen den Link Managern ausgetauscht wird. Auf diese Weise ist gewährleistet, dass bei jeder Aktivierung der Verschlüsselung ein neuer Ciphering Key verwendet wird (Abb. 5.17).

Der Ciphering Key hat üblicherweise eine Länge von 128 Bit. Es können jedoch auch kürzere Ciphering Keys verwendet werden, wenn Bluetooth Chips für ein Land hergestellt werden, für das es Exportrestriktionen für starke Verschlüsselungskeys gibt.

Zusammen mit der Geräteadresse des Masters und den 26 untersten Bits der Master Echtzeituhr (Master Real Time Clock) dient der Ciphering Key als Eingangswert für den SAFER+Algorithmus E0, der einen kontinuierlichen Bitstrom erzeugt. Da der aktuelle Wert der Master Real Time Clock auch dem Slave bekannt ist, kann auf beiden Seiten der Verbindung der gleiche Bitstrom generiert werden. Der Bitstrom wird dann über bitweise Modulo-2 Operationen mit dem zu verschlüsselnden Datenstrom kombiniert. Verschlüsselt wird der komplette Teil des ACL Nutzdatenpaketes inklusive der CRC Checksumme vor dem optionalen Hinzufügen einer Forward Error Correction (FEC).

5.5.5 Autorisierung

Ein weiteres wichtiges Konzept der Bluetooth Sicherheit ist die Autorisierung des Nutzers für einen Dienst. Dieser weitere Schritt ist nötig, um manche Dienste nicht allen, sondern nur bestimmten Endgeräten zugänglich zu machen. So könnte man auf einem PC einem Nutzer eines anderen Bluetooth Geräts das Recht einräumen, auf ein freigegebenes Verzeichnis Dateien abzulegen oder von dort abzuholen. Der Dateitransfer Dienst (OBEX) ist also für diesen Nutzer aktiviert.

Über die Autorisierung kann für jeden Dienst einzeln festgelegt werden, welche bekannten Bluetooth Endgeräte auf diesen zugreifen dürfen. Es bleibt dabei dem Hersteller eines Bluetooth Gerätes überlassen, wie diese Funktionalität genutzt wird. Manche Mobiltelefonhersteller beispielsweise erlauben jedem entfernten Endgerät, mit dem ein Pairing erfolgreich durchgeführt wurde, die Benutzung des Dial-Up Dienstes. Andere Mobiltelefonhersteller bauen jedoch noch eine zusätzliche Sicherung ein und fordern vom Nutzer des Mobiltelefons eine explizite Autorisierung des Verbindungs-wunsches. Dies geschieht über eine Nachricht auf dem Display des Mobiltelefons, die der Besitzer des Mobiltelefons bestätigen muss.

Bluetooth Stacks auf PCs bieten meist eine sehr flexible Autorisierungsfunktionalität an. Dienste können dort sehr flexibel konfiguriert werden:

- Dienst ohne Authentifizierung und Autorisierung nutzbar.
- Dienst darf von allen authentifizierten Geräten ohne weitere Autorisierung verwendet werden. Dies setzt ein einmaliges Pairing voraus.
- Dienst darf nach Authentifizierung und Autorisierung einmalig oder für eine bestimmte Zeitdauer verwendet werden.
- Dienst darf von einem bestimmten Endgerät nach Authentifizierung und einmaliger Autorisierung immer verwendet werden, eine nochmalige Autorisierung ist nicht erforderlich.

Zusätzlich bieten manche Bluetooth Stacks auf dem PC an, immer eine Information auf dem Bildschirm anzuzeigen, wenn ein Dienst von einem entfernten Gerät aufgerufen wird. Dies dient nur zur Information des Nutzers des PCs, der Zugriff wird automatisch gewährt.

5.5.6 Sicherheitsmodi

Zu welchen Zeitpunkten beim Verbindungsauftbau eine Authentifizierung, Verschlüsselung und Autorisierung durchgeführt werden, ist abhängig von der Implementation des Bluetooth Stacks und der Konfiguration durch den Anwender. Der Bluetooth Standard gibt dazu drei mögliche Konfigurationen vor:

Im Sicherheitsmodus 1 (Security Mode 1) findet keine Authentifizierung statt und die Verbindung wird nicht verschlüsselt. Dieser Sicherheitsmodus eignet sich z. B. für die Adress- oder Terminübertragung zwischen zwei Endgeräten. Oft kennen sich die Teilnehmer nicht und es wäre zu umständlich, mit den Geräten vor dem Austausch einer elektronischen Visitenkarte ein Pairing durchzuführen. Die elektronische Visitenkarte wird dann meist von den Geräten in ein extra Verzeichnis kopiert und erst in den Adresskalender aufgenommen, wenn der Benutzer dies bestätigt.

Im Sicherheitsmodus 2 bestimmt der Anwender, ob für eine Verbindung eine Authentifizierung, Verschlüsselung und Autorisierung nötig ist. Viele Bluetooth PC Benutzeroberflächen erlauben diese Konfiguration individuell für jeden einzelnen Dienst. Sicherheitsmodus 1 entspricht Sicherheitsmodus 2 eines Dienstes, der weder Authentifizierung noch Verschlüsselung aktiviert hat.

Im Sicherheitsmodus 3 wird beim Aufbau jeder Verbindung automatisch eine Authentifizierung und Verschlüsselung vom Bluetooth Chip hergestellt. Dies geschieht schon während der ersten Link Manager Kommunikation, also noch vor dem Aufbau einer L2CAP Verbindung. Bei einer eingehenden Kommunikation fordert deshalb der Bluetooth Controller über die HCI Schnittstelle den Link Key für eine neue Verbindung an. Wurde mit dem entfernten Gerät bisher kein Pairing durchgeführt, kann der Bluetooth Host dem Controller keinen Link Key zurückgeben. In diesem Fall schlägt der Verbindungsauftbau fehl. Sicherheitsmodus 3 ist also vor allem für Geräte gedacht, die nur mit Geräten kommunizieren, mit denen zuvor ein Pairing durchgeführt wurde. Für Mobiltelefone, die auch nicht authentifizierte Verbindungen z. B. für die Übertragung von Adressdaten erlauben, ist dieser Modus nicht geeignet.

Sicherheitsmodus 4 ist dem Service Level Enforced Security Mode 2 sehr ähnlich, wurde jedoch für die neuen Pairingmechanismen für Bluetooth 2.1 spezifiziert (vgl. Abschn. 5.5.2). In diesem Modus wählt ein Dienst aus, welche Security Kategorie er für das Pairing verlangt:

- Es wird ein gesicherter Link Key verlangt (Numeric Comparsion, Out of Band oder Passkey Protokoll sind notwendig)
- Es wird nur ein nicht gesicherter Link Key benötigt (Just Works Protokoll)
- Der Dienst benötigt keine Sicherheit

5.6 Bluetooth Profile

Wie in der Einleitung dieses Kapitels gezeigt, ist Bluetooth für eine Vielzahl sehr unterschiedlicher Anwendungen geeignet. Diese Anwendungen haben immer eine Server- und eine Client Seite. Ein Client nimmt durch Aufbau einer Bluetooth Verbindung Kontakt zum Master auf und die Datenübertragung beginnt. Bei den meisten Bluetooth Anwendungen sind die Aufgaben der Masterseite und der Clientseite unterschiedlich. Bei der Übertragung eines Adressbucheintrags beispielsweise, nimmt der Client Kontakt mit dem Server auf. Der Client überträgt einen Termin, ist also eine Sendekomponente, der Server empfängt ihn, ist also eine Empfangskomponente. Um zu gewährleisten, dass der Client auch mit einem Server kommuniziert, der von einem anderen Hersteller programmiert wurde, spezifiziert der Bluetooth Standard sogenannte Bluetooth Profile. Für jede Anwendung (Headset, Termin- und Dateiübertragung, Audiostreaming, etc.) gibt es ein Bluetooth Profil, das genau beschreibt, wie die Serverseite und die Clientseite miteinander kommunizieren. Unterstützen zwei Endgeräte das gleiche Bluetooth Profil, ist die Interoperabilität gewährleistet.

Anmerkung: Das Client/Server Prinzip der Bluetooth Profile darf nicht mit dem Master/Slave Konzept der unteren Bluetooth Protokollsichten verwechselt werden. Beim Master Slave Konzept geht es um die Kontrolle des Piconetzes, also wer zu welcher Zeit senden darf, während das Client/Server Prinzip einen Dienst und einen Nutzer des Dienstes beschreibt. Ob nun das Bluetooth Endgerät, auf dem der Server eines Dienstes läuft, der Master oder der Slave im Piconetz ist, spielt keine Rolle.

Nachfolgende Tabelle gibt einen Überblick über zahlreiche Bluetooth Profile für die verschiedensten Anwendungen. In der Praxis ist heute jedoch zu beobachten, dass sich die Bluetooth Nutzung auf nur noch wenige Profile beschränkt. Diese werden in den nachfolgenden Unterkapiteln genauer beschrieben.

Profilname	Anwendungsgebiet
Headset Profile	Kabellose Headsets für Mobiltelefone
<i>Hands-Free Profile</i>	Verbindung zwischen Freisprecheinrichtung und Mobiltelefon
<i>SIM-Access Profile</i>	Zugriff einer Freisprecheinrichtung auf die SIM-Karte eines Mobiltelefons
<i>Human Interface Device (HID) Profile</i>	Anbindung von Mäusen, Tastaturen und Joysticks an Endgeräte wie PCs, Notebooks und Smartphones
<i>File Transfer Profile</i>	Übertragung von Dateien zwischen Bluetooth Geräten
<i>Object Push Profile</i>	Einfache Übertragung von Dateien zwischen Bluetooth Geräten für Ad-Hoc Datenaustausch
<i>Advanced Audio Distribution Profile</i>	Profil für die Übertragung von Audio Streaming Dateien (z. B. MP-3)

Profilname	Anwendungsgebiet
<i>Dial Up Networking (DUN) Profile</i>	Bluetooth Verbindung zwischen einem Mobiltelefon und einem externen Gerät wie PC oder Notebook
<i>FAX Profile</i>	Profil für FAX Übertragung
<i>LAN Access Profile</i>	IP Verbindung zwischen Smartphone, PC oder Notebook zu einem Local Area Network (LAN) und dem Internet
<i>Personal Area Network (PAN) Profile</i>	Wie LAN Access Profile, es wird jedoch eine Ethernet Netzwerkkarte auf dem PAN Gerät simuliert
<i>Synchronization Profile</i>	Synchronisation von Personal Information Manager (PIM) Anwendungen für Adressen, Termine, Notizen, etc.
<i>Basic Imaging Profile</i>	Übertragung von Bildern, für den Einsatz mit Digitalkameras gedacht
<i>Hard Copy Cable Replacement Profile</i>	Kabelersatz zwischen Drucker und einem Endgerät (z. B. PC)
<i>Basic Printing Profile</i>	Drucken ohne Druckertreiber von mobilen Geräten wie Smartphones an beliebigen Druckern
<i>Unrestricted Digital Information Profile</i>	Übertragung von breitbandigen leitungsvermittelten Verbindungen zwischen einem Endgerät und einem Mobiltelefon

5.6.1 Grundlegende Profile: GAP, SDP und Serial Profile

Bluetooth spezifiziert zwei Profile, die keine eigentlichen Anwendungen aus Sicht des Benutzers darstellen. Das Generic Access Profile (GAP) legt fest, wie zwei Geräte in unterschiedlichen Situationen Kontakt miteinander aufnehmen und wie sie sich dabei verhalten sollen. Das Profil beschreibt unter anderem:

- Die Präsentation von Bluetooth spezifischen Parametern wie der Geräteadresse (BD_ADDR) oder der PIN für den Anwender.
- Sicherheitsaspekte (Security Mode 1–3)
- Verhalten im Idle Mode (z. B. Inquiry, Device Discovery)
- Verbindungsaufbau

Durch das GAP Profil kann somit sichergestellt werden, dass sich die Benutzeroberflächen für die Konfiguration des Bluetooth Stacks von verschiedenen Endgeräten in den wichtigsten Punkten sehr ähnlich sind. Außerdem wird durch das GAP Profil erreicht, dass bei der Verbindungsaufnahme genau spezifiziert ist, welche Aktionen und Nachrichten in welcher Reihenfolge durchgeführt werden.

Wie in Abschn. 5.4.7 gezeigt, besitzt ein Bluetooth Endgerät eine Service Datenbank, in der jeder Server-Dienst alle wichtigen Informationen für die Verbindungsaufnahme hinterlegen kann. Über das Service Discovery Profil (SDP) wird festgelegt, wie auf diese Datenbank zugegriffen werden kann, und wie und in welcher Struktur die nachfolgend vorgestellten Profile ihre Informationen in der Service Datenbank hinterlegen.

Das Serial Port Profile (SPP) ist ein grundlegendes Profil, auf dem zahlreiche nachfolgend vorgestellte Profile aufbauen. Wie der Name schon andeutet, stellt dieses Profil eine serielle Schnittstelle für beliebige Anwendungen zur Verfügung. Es verwendet dazu die in Abschn. 5.4.8 vorgestellte RFCOMM Schicht. Über das Serial Port Profile können beliebige Anwendungen, die Daten über eine serielle Schnittstelle übertragen, kommunizieren. Anpassungen der Anwendungen an Bluetooth sind nicht notwendig, da aus Ihrer Sicht auf eine ganz normale serielle Schnittstelle zugegriffen wird. Abb. 5.18 zeigt den Protokollstack, den das Serial Port Profile verwendet.

5.6.2 Object Exchange Profile: FTP, Object Push und Synchronize

Um strukturierte Objekte wie Dateien, Visitenkarten, Termine, Adressbucheinträge oder generell Objekte zu übertragen, wird das Object (OBEX) Exchange Profile verwendet (Abb. 5.19).

Die Verbindung zwischen zwei Geräten besteht bei diesen Profilen nur während der Übertragung eines oder mehrerer unmittelbar aufeinander folgender Objekte und wird danach sofort wieder abgebaut. Zu diesem Zweck definiert der Bluetooth Standard als Grundlage für weitere Profile das General Object Exchange (OBEX) Profile (GOEP), das auf den L2CAP und RFCOMM Schichten aufsetzt. Drei weitere Object Exchange (OBEX) Profile verwenden dann dieses Profil für spezifische Dienste.

Für die Übertragung einer oder mehrerer Dateien, oder sogar eines ganzen Verzeichnisbaumes, wurde das File Transfer Profile (FTP) entwickelt. Dieses sollte nicht mit dem File Transfer Protocol aus der TCP/IP Welt verwechselt werden, das ebenfalls mit FTP abgekürzt wird.

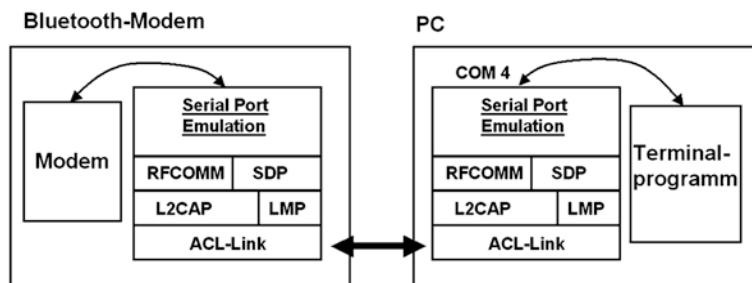


Abb. 5.18 Das SPP stellt eine serielle Schnittstelle zur Verfügung

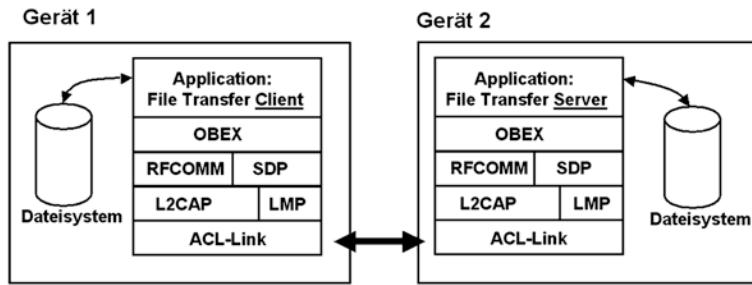


Abb. 5.19 OBEX mit File Transfer Profile als Anwendung

Eingesetzt wird das OBEX FTP Protokoll hauptsächlich, um zwischen PCs und Smartphones Dateien auszutauschen. Diese können sich an einem beliebigen Ort innerhalb eines Dateisystems befinden. Zu diesem Zweck definiert das allgemeine OBEX Profil (GOEP) die Kommandos CONNECT, DISCONNECT, PUT, GET, SETPATH und ABORT, die binär kodiert über eine aufgebaute RFCOMM Verbindung zur Gegenstelle übertragen werden. Manche PC Bluetooth Stacks klinken das Dateisystem einer Bluetooth Gegenstelle, ähnlich einer normalen Netzwerkverbindung, in den Verzeichnisbaum des lokalen Dateimanagers ein. Klickt der Benutzer das Bluetooth Gerät an, wird über das allgemeine OBEX GET Kommando das Root-Directory des entfernten Bluetooth Gerätes angefordert und dann im Dateimanager dargestellt. Der Anwender hat dann die Möglichkeit, eine oder mehrere Dateien auszuwählen und auf den lokalen PC zu übertragen. Auch diese Aktion wird in ein GOEP GET Kommando umgesetzt. Der Anwender kann auch eine Datei in ein Verzeichnis eines anderen Bluetooth Gerätes kopieren. Zu diesem Zweck wird das allgemeine OBEX PUT Kommando verwendet.

Wechselt der Anwender in ein Unterverzeichnis, wird in dieses über das OBEX SETPATH Kommando verzweigt und dessen Inhalt anschließend über das allgemeine OBEX GET Kommando angefordert. Wie das nachfolgende Beispiel in der Textbox zeigt, wird der Inhalt eines Verzeichnisses in lesbbarer Form als XML Beschreibung übertragen.

Im OBEX Protokoll Layer werden CONNECT, DISCONNECT, PUT, GET, SETPATH und ABORT Kommandos und die entsprechenden Antworten darauf als Pakete behandelt. Der Wert des ersten Byte des Pakets beschreibt die Art des Kommandos. Nach einem zwei Byte Längenfeld folgen dann die Parameter des Kommandos. Ein Parameter kann z. B. ein Verzeichnisname, eine Verzeichnisauflistung oder eine angeforderte Datei sein. Diese Parameter werden im Standard etwas verwirrend als Header bezeichnet. Um die Art der Parameter auseinander halten zu können, hat jeder Parameter im ersten Byte eine Typinformation. Der Typ eines Parameters kann z. B. „Dateiname“ oder „Body“ (also die eigentliche Datei) sein.

```

<xml version="1.0">
<!DOCTYPE folder-listing SYSTEM „obex-folder listing.dtd“>
<folder-listing-version="1.0">
    <folder name="Camera" modified="2004117T100840"
        user perm="RWD" group perm="W" />
    <folder name="other pics" modified="2004117T13321"
        user perm="RWD" group perm="W" />
</folder-listing>

```

Die maximale Paketgröße beträgt 64 kByte. Um größere Dateien (also Header vom Typ „Body“) zu übertragen, wird die Datei automatisch vom OBEX Layer in mehrere Pakete aufgeteilt.

Während FTP in der Praxis heute an Bedeutung verloren hat, findet die etwas einfachere Anwendung des General Object Exchange Profils, das Object Push Profile, in der Praxis durchaus noch Anwendung. Dieses wird z. B. verwendet, wenn der Benutzer eines Mobiltelefons einen Kalendereintrag, einen Adressbucheintrag oder eine Datei über Bluetooth zu einem anderen Gerät übertragen möchte. Die Funktionsweise dieses Profils ist identisch zum File Transfer Profil, es verwendet ebenfalls die allgemeinen OBEX Kommandos wie PUT und GET. Das Object Push Profile unterstützt jedoch keine Verzeichnisoperationen und Löschen von Dateien. Auf diese Weise wird erreicht, dass der Benutzer beim Senden der Informationen möglichst wenige Entscheidungen treffen muss und der Vorgang somit schnell durchgeführt werden kann (Abb. 5.20).

Viele Endgeräte erlauben einen eingehenden Object Push Transfer ohne vorherige Authentifizierung und Verschlüsselung. Das empfangene Objekt wird dann nach Erhalt zunächst in einen Zwischenpuffer gelegt und erst nach Bestätigung des Benutzers in den Terminkalender, in das Adressbuch, oder, im Falle einer Datei, in ein Verzeichnis kopiert.

Für die Übertragung von Kalender- und Adressbucheinträgen schreibt das Object Push Profile das vCalendar, bzw. das vCard Format vor (www.imc.org). Dies ist Voraussetzung, um Adressbuch- und Kalendereinträge zwischen beliebigen Programmen und

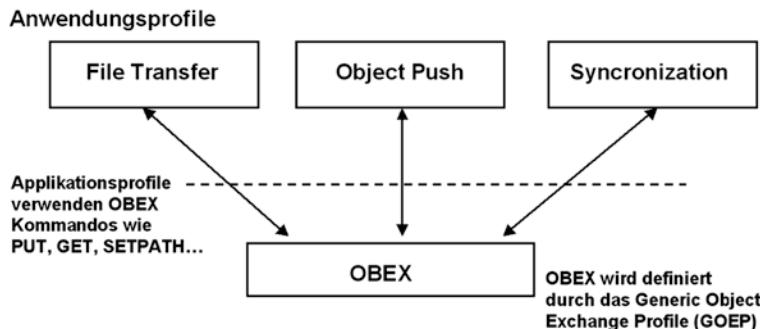


Abb. 5.20 Zusammenhang zwischen OBEX, GOEP, FTP, Object Push und Synchronization Profile

Endgeräten austauschen zu können. Bei anderen Objekten, wie z. B. Bildern, kann anhand der Endung des Dateinamens erkannt werden, um welche Art Datei es sich handelt.

Obwohl das Profil „Object Push“ heißt, spezifiziert es auch optional eine Business Card Pull Funktion. Mit dieser Funktion kann man eine zuvor hinterlegte Standard-visitenkarte von einem Gerät anfordern. Die Business Card Exchange Funktion ergänzt diese Funktion, in dem nicht nur eine Visitenkarte angefordert wird, sondern auch die bei sich hinterlegte Visitenkarte automatisch dem anderen Gerät geschickt wird.

Das dritte Profil, das auf GOEP aufsetzt, ist das Synchronization Profile. Wie das File Transfer Profil hat sich auch dieses Profil in der Praxis nicht weit verbreitet, soll aber vollständigkeitshalber trotzdem erwähnt werden. Das Synchronization Profil ermöglicht den automatischen Abgleich von Objekten wie Terminkalender- und Adressbucheinträgen, sowie Notizen zwischen zwei Geräten. Auch dafür werden wieder die allgemeinen OBEX Kommandos wie GET und PUT verwendet. Gegenüber dem Object Push Profil, über das vom Anwender nur ausgewählte Objekte, wie z. B. ein Adressbucheintrag, zu einem anderen Gerät übertragen werden können, spezifiziert das Synchronization Profile, wie der komplette Datenbestand einer Datenbank synchronisiert werden kann. Bei der ersten Synchronisation wird einmalig der komplette Datenbestand in beide Richtungen übertragen, bei allen folgenden Synchronisationen werden dann nur noch die geänderten Objekte übertragen. Zu diesem Zweck führen beide Geräte eine Protokolldatei über alle Änderungen. Damit Anwendungen unterschiedlicher Hersteller ihre Datenbankeinträge austauschen können, werden wie auch im Object Push Profil standardisierte Formate wie vCard oder vCalendar verwendet.

Der Bluetooth Standard definiert den Ablauf der Synchronisation nicht selbst, sondern verwendet dazu das Synchronisationssystem, das im IrMC Standard der Infrared Data Association (www.irda.org) definiert wurde.

5.6.3 Headset, Hands-Free und SIM-Access Profile

Drahtlose Headsets für Mobiltelefone waren die ersten Geräte, die mit Bluetooth Funktionalität auf den Markt kamen. Für die Sprachverbindung zwischen Mobiltelefon und Headset wird das Headset Profil verwendet. Dieses Profil ist eine Besonderheit, denn es verwendet als eines der wenigen Profile auch SCO oder eSCO Pakete (vgl. Abschn. 5.4.1). Mit diesen wird zwischen Mobiltelefon und Headset ein Sprachkanal mit 64 kbit/s aufgebaut. Sind Mobiltelefon und Headset kompatibel zu Bluetooth 1.2, werden automatisch eSCO Pakete verwendet, die Verbindung profitiert dann von automatischer Fehlerkorrektur und Adaptive Frequency Hopping (AFH). Diese in Bluetooth 1.2 eingeführten Funktionalitäten steigern die Sprachqualität vor allem dann wesentlich, wenn die Bluetooth Verbindung aufgrund eines großen Abstands, einer geringen Sendeleistung, oder durch Hindernisse nicht optimal ist.

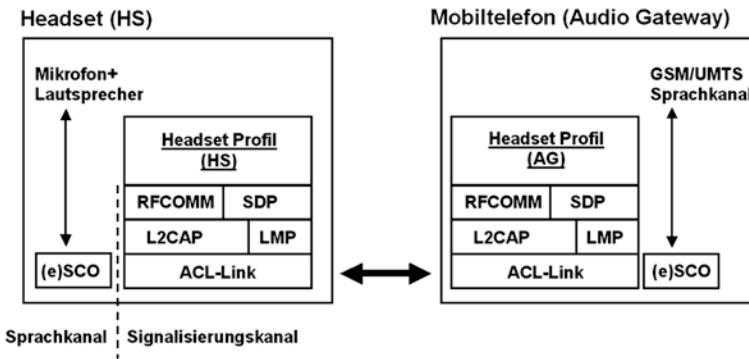


Abb. 5.21 Headset Protokollstack

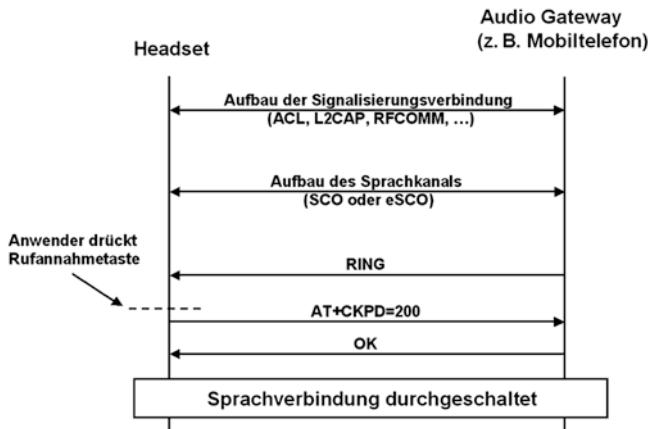


Abb. 5.22 Aufbau von Signalisierungs- und Sprachverbindung

Um ein Headset mit einem Mobiltelefon verwenden zu können, müssen die zwei Geräte einmalig miteinander ein Pairing durchführen. Danach versucht das Mobiltelefon bei jedem eingehenden Anruf automatisch, eine Verbindung zum Headset herzustellen. Für die Signalisierung zwischen Headset und Mobiltelefon, das im Headset Profil als Audio Gateway (AG) bezeichnet wird, wird eine ACL Verbindung verwendet. Wie in Abb. 5.21 zu sehen ist, wird für die Signalisierungsverbindung auf höheren Schichten L2CAP und RFCOMM verwendet.

Um Kommandos und die dazugehörigen Antworten zwischen Audio Gateway und Headset auszutauschen, wird das von Modems bekannte AT-Kommandoset verwendet. Das Headset Profil beschränkt sich jedoch auf nur wenige Kommandos. Wie in Abb. 5.22 zu sehen ist, baut das Audio Gateway bei einem eingehenden Anruf

zuerst eine Signalisierungsverbindung auf (ACL) und sendet über den Signalisierungs-kanal den String „RING“. Das Headset benachrichtigt daraufhin den Anwender über den eingehenden Anruf, in dem z. B. eine Melodie gespielt wird. Der Nutzer kann dann den Anruf durch Betätigen einer Taste am Headset annehmen. Das Betätigen der Taste bewirkt, dass das Headset das AT-Kommando `at+ckpd=200` an das Audio Gateway zurückschickt. Dieses nimmt daraufhin das Gespräch an und stellt es zum Headset durch.

Um ein abgehendes Gespräch zu führen, kann umgekehrt auch das Headset eine Verbindung zum Audio Gateway herstellen. Zusammen mit einer im Audio Gateway (also im Mobiltelefon) vorhandenen Sprachwahlfunktion lassen sich somit abgehende Gespräche über das Headset starten, ohne das Mobiltelefon in die Hand zu nehmen.

Da die Bedienmöglichkeiten durch die Größe des Headsets begrenzt sind, bietet das Headset Profil außer der Gesprächsfunktionalität nur noch die Steuerung der Lautstärke. Dies geschieht über die Befehle `+vgm` für die Lautstärke des Mikrofons und `+vgs` für die Lautstärke des Lautsprechers. Mit diesen Befehlen kann also vom Mobiltelefon aus die Lautstärke im Headset geändert werden.

Ein Headset kann auch mit einem PC gekoppelt werden, falls der Bluetooth Stack des PCs das Headset Profil unterstützt und die Rolle des Audio Gateways übernehmen kann. Auf diese Weise kann das Headset z. B. zusammen mit einer Voice over IP Software verwendet werden. Außerdem ist es durch die Umleitung der Soundkarten Ein- und Ausgänge auf das Headset theoretisch auch möglich, Musik, MP3 Streams, etc. über das Headset abzuspielen. Dies macht jedoch wenig Sinn, da der SCO Kanal auf 64 kbit/s begrenzt ist und nur für Sprachtelefonie ausgelegt ist. In der Praxis bedeutet dies, dass das Audiosignal nur mono übertragen wird und das Frequenzband auf 300–3400 Hz begrenzt ist.

Stark verwandt mit dem Headset Profil ist das Hands-Free Profil. Bei der Entwicklung dieses Profils standen jedoch nicht Headsets im Vordergrund, sondern KFZ-Freisprecheinrichtungen. Wichtigste Aufgabe des Hands-Free Profil ist das Ersetzen der Kabelverbindung zwischen Freisprecheinrichtung und Mobiltelefon. Auf diese Weise muss das Mobiltelefon bei Fahrtantritt nicht in einer Halterung festgemacht werden und kann sich während der Fahrt an einer beliebigen Stelle im Auto befinden. Diese Aufgabe könnte auch mit dem Headset Profil bewerkstelligt werden. Da Freisprecheinrichtungen aber heute weit mehr Funktionen bieten, als nur an- und abgehende Gespräche zu führen, wurde das Hands-Free Profil definiert.

Die grundsätzliche Funktionsweise des Hands-Free Profil ist mit dem Headset Profil identisch. Kommandos und entsprechende Antworten werden zwischen Freisprecheinrichtung (Hands-Free Unit) und dem Mobiltelefon (Audio Gateway) ebenfalls über AT-Kommandos ausgetauscht. Außerdem wird ebenso wie beim Headset Profil der Sprachkanal über eine SCO oder eSCO Verbindung geleitet. Zusätzlich zu den Funktionen des Headset Profils bietet das Hands-Free Profil auch folgende Möglichkeiten:

- Die Übertragung der Rufnummer des Anrufers an die Freisprecheinrichtung (CLIP Funktion).
- Abweisen von ankommenden Gesprächen von der Freisprecheinrichtung aus.
- Wählen einer Telefonnummer von der Freisprecheinrichtung.
- Gespräch halten sowie Dreierkonferenzsteuerung.
- Übertragung von Statusinformationen wie verbleibende Batteriekapazität und GSM/ UMTS Empfangsstärke des Mobiltelefons.
- Roaminganzeige.
- Deaktivieren der optionalen Echounterdrückung im Endgerät, falls dies vom Endgerät unterstützt wird. Dies ist sinnvoll, wenn die Freisprecheinrichtung eine eigene Echo- unterdrückung besitzt.

Seit Version 1.6 des Hands-Free Profile ist optional auch die Verwendung eines mono Wideband Sprachcodecs (mSBC) möglich und erlaubt somit eine bessere Sprachqualität, falls das Netzwerk AMR-Wideband unterstützt. Bei Headsets wird diese Erweiterung manchmal auch als „HD Voice“ kompatibel bezeichnet. Eine weitere Lösungsmöglichkeit für die gleichzeitige Verwendung eines Headsets und einer KFZ-Freisprecheinrichtung bietet das SIM-Access Profil. Im Unterschied zum Headset und Hands-Free Profil dient das Mobiltelefon beim SIM-Access Profil nicht als Audio Gateway, und somit als Brücke zum Mobilfunknetzwerk, sondern stellt nur die SIM Karte einem externen Gerät zur Verfügung. Abb. 5.22 zeigt dieses Szenario. Das externe Gerät, in den meisten Fällen also eine KFZ-Freisprecheinrichtung, enthält ein eigenes GSM/UMTS Mobiltelefon, jedoch ohne SIM Karte. Wird die Freisprecheinrichtung bei Fahrtantritt aktiviert, wird per Bluetooth Kontakt zum gekoppelten Mobiltelefon hergestellt. Durch die Aktivierung des SIM-Access Servers im Mobiltelefon wird automatisch der Mobilfunkteil deaktiviert. Dies ist notwendig, da die Mobiltelefoneinheit in der Freisprecheinrichtung fortan die Kommunikation mit dem Mobilfunknetzwerk übernimmt. Ein großer Vorteil dieser Methode ist weiterhin, dass die Freisprecheinrichtung auch an die KFZ-Spannungsversorgung und an eine Außenantenne angeschlossen ist. Dies können Headset und Hands-Free Profil nicht bieten.

Abb. 5.22 zeigt außerdem den für das SIM-Access Profil verwendeten Protokollstack. Auf der L2CAP Verbindung wird der RFCOMM Layer für eine serielle Übertragung zwischen Freisprecheinrichtung (SIM-Access Client) und Mobiltelefon (SIM-Access Server) verwendet. Neben SIM-Access Profil Kommandos für die Aktivierung, Deaktivierung und den Reset der SIM-Karte werden über den Bluetooth Kanal auch SIM-Karten Kommandos und Antwortnachrichten ausgetauscht. Kommandos und Antwortnachrichten werden als Application Protocol Data Units (APDUs) übertragen. Diese wurden bereits in Abschn. 6.10 beschrieben und in den Abb. 6.49 und 6.50 dargestellt. Statt die APDUs also zwischen Mobilfunkteil und SIM Karte des Mobiltelefons über das elektrische Interface auszutauschen, werden mit dem SIM-Access Profil die APDUs über die Bluetooth Schnittstelle ausgetauscht. Für die Software der Freisprecheinrichtung, die auf dem SIM-Access Profile aufsetzt, ist es also völlig transparent, dass die SIM Karte nicht fest eingebaut ist, sondern über Bluetooth angesprochen wird (Abb. 5.23).

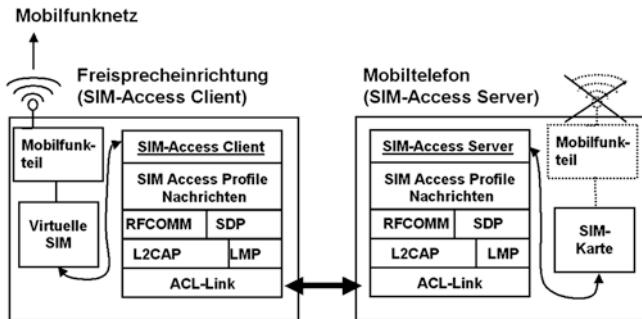


Abb. 5.23 Funktionsweise des SIM-Access Profils

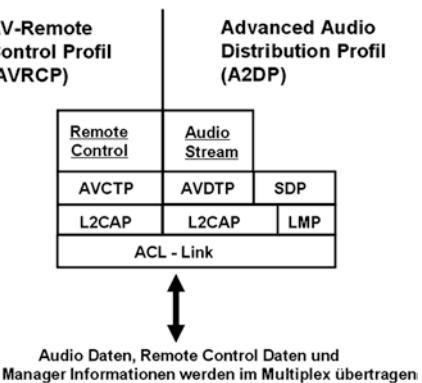
Durch die Verwendung von APDUs können nicht nur die Dateien auf der SIM Karte gelesen und geschrieben werden, sondern es kann auch der Authentifizierungsalgorithmus der SIM-Karte angesprochen werden, der zu einer Zufallszahl (RAND) eine Signed Response (SRES) erzeugt (vgl. Abschn. 6.6.4). Außerdem kann auch das SIM Application Toolkit Protokoll über die Bluetooth Verbindung genutzt werden. Auch diese Nachrichten werden, wie ebenfalls in Abschn. 6.10 gezeigt, in APDUs verpackt.

5.6.4 High Quality Audio Streaming

Sowohl das Handset- als auch das Handsfree Profil wurden ursprünglich entwickelt, um Sprache in Telefonqualität und in mono zu übertragen. Für Hifi Audiostreaming reicht diese Qualität jedoch bei weitem nicht aus. Für diese Anwendung wurde deshalb das Advanced Audio Distribution Profil (A2DP) entwickelt, das Audiodaten mit Bandbreiten von 127–345 kbit/s überträgt. Da solche Datenraten nicht über SCO Verbindungen transportiert werden können, verwendet dieses Profil ACL Links zum Datentransport. Erste Versionen des Profils gibt es schon seit 2003, es dauerte jedoch einige Jahre, bis erste Geräte etwa 2006/07 auf den Markt kamen. Zu Geräten, die das A2DP Profil unterstützen sind Mobiltelefone mit eingebautem MP-3 Player und Kopfhörer. Kopfhörer unterstützen üblicherweise sowohl das A2DP und die Handsfree und Headset Profile. Mit einem eingebauten Mikrofon können diese dann sowohl für Musik als auch zum Telefonieren verwendet werden.

Abb. 5.24 zeigt den A2DP Protokollstack. Das Profil basiert auf GAP und erlaubt somit anderen Endgeräten, die unterstützten Funktionalitäten in der SDP Datenbank abzufragen. Oberhalb des L2CAP Layer wurde das Audio Video Distribution Transfer Protocol (AVDTP) für die Datenübertragung spezifiziert. Wie der Name des Protokolls schon andeutet, kann es sowohl für die Übertragung von Audiodaten, als auch für die Übertragung von Videostreams verwendet werden. Das A2DP Profil verwendet das Protokoll jedoch lediglich für die Übertragung von Audiodaten. Neben der Übertragung

Abb. 5.24 Der A2DP Protokoll Stack inklusive Remote Control



des reinen Audiostreams werden auch Kontrollinformationen wie z. B. Codec Vereinbarungen und Austausch von Parametern wie der benötigten Bandbreite über das AVDTP Protokoll abgewickelt. Höhere Kontrollfunktionen wie z. B. das Springen zum nächsten Musikstück oder das Pausieren der Übertragung sind nicht Teil von AVDTP und werden über das Audio/Video Control Transport Protocol (AVCTP) übertragen, das nachfolgend beschrieben wird.

Der Bluetooth Standard erlaubt einem Endgerät, mehrere Verbindungen zu mehreren Geräten gleichzeitig geöffnet zu haben. Unterstützt ein Gerät dies, kann z. B. eine A2DP Verbindung zwischen einem Notebook und einem Kopfhörer aufgebaut sein, während gleichzeitig das Notebook Dateien, wie z. B. Bilder, mit einem weiteren Endgerät austauscht. Eine A2DP Übertragung benötigt jedoch für einen Audiostream in guter Qualität schon einen großen Teil der über Bluetooth möglichen Bandbreite, sodass die Dateübertragung entsprechend langsam erscheint. Unterstützen alle Endgeräte im Piconet den Bluetooth 2.0+EDR Standard, wird dies sicher weniger auffallen, da die Bandbreite dann etwa 2 Mbit/s beträgt. Dies ist deutlich mehr als bei Version 1.2 mit einem Limit von 723 kbit/s, von denen dann mit dem besten Audio Codec 345 kbit/s für die Audioübertragung verwendet werden.

Das A2DP Profil spezifiziert zwei Rollen für eine Verbindung. Die Audio Source Rolle wird von Geräten wie MP-3 Playern, Mobiltelefonen oder einem Mikrofon übernommen. Die andere Seite der Verbindung ist die Audio Senke (Audio Sink) Rolle, die üblicherweise von einem Headset oder einem Bluetooth Lautsprecherset übernommen wird.

Um mindestens einen gemeinsamen Audio Codec für eine A2DP Übertragung zwischen zwei Geräten zu gewährleisten, enthält die A2DP Spezifikation ein proprietäres Audioformat, das Sub-band Codec (SBC) genannt wird. Dieses muss von allen A2DP kompatiblen Endgeräten unterstützt werden und wird nachfolgend kurz beschrieben. Außerdem definiert der Standard die Übertragung anderer Codecs wie MPEG 1–2 Audio, MPEG-2,4, AAC und ATRAC über das Audio/Video Distribution Protocol (AVDTP). Diese Codecs sind optional. Der Standard erlaubt auch die Übertragung von weiteren Codecs über AVDTP. Um eine Interoperabilität zwischen Geräten unterschiedlicher Hersteller zu gewährleisten, muss ein Gerät jedoch immer in der Lage sein, einen

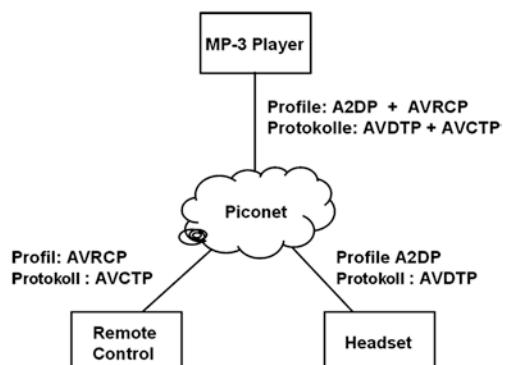
Audiostream in SBC zu konvertieren, wenn ein anderes Gerät kein anderes optionales Format unterstützt.

Grundsätzlich ist der SBC Codec wie folgt aufgebaut: Als Eingangssignal erwartet der SBC Codec ein PCM kodiertes Audiosignal mit einer Abtastfrequenz von 44,1 oder 48 kHz. Der Codec teilt im ersten Schritt dann das Frequenzband des Eingangssignals in mehrere Teilbereiche auf, die auch als Unterbänder (Sub-Bands) bezeichnet werden. Der Standard rät, das Signal entweder in vier oder in acht Unterbänder aufzuteilen. Danach wird ein Skalierungsfaktor für jeden Unterkanal berechnet, der die Lautstärke des Signals in diesem Unterband beschreibt. Die Skalierungsfaktoren werden dann miteinander verglichen, um dem Unterband mit der meisten Signalinformation auch die meisten Bits für die Kodierung zuzuordnen. Der Standard schlägt vor, mindestens 19 Bit für eine mittlere Audioqualität und einen Monokanal zu verwenden und bis zu 55 Bits für Stereokanäle mit hoher Qualität. Nach der Kodierung der Audioinformation der Unterkanäle werden die Datenströme komprimiert. Der Kompressionsfaktor ist variabel, und es besteht somit hier nochmals die Möglichkeit, eine Abwägung zwischen Datenrate und Audioqualität zu treffen. Wird für die Kompression der niedrigste Faktor verwendet und Stereokanäle mit höchster Qualität kodiert, erzeugt dies einen Datenstrom mit einer Geschwindigkeit von 345 kbit/s.

Um Benutzeranweisungen vom Audio Sink Device (z. B. einem Kopfhörer) wie Lautstärkeregelung, nächster Track, Pause, etc., zurück zum Audio Source Gerät (z. B. einem MP-3 Player) zu übertragen, wird das Audio/Video Remote Control Profile verwendet. Wie in Abb. 5.25 gezeigt, wird dazu das Audio/Video Control Transport Protocol verwendet. Auch diese Nachrichten sind standardisiert um sicherzustellen, dass Endgeräte verschiedener Hersteller zusammenarbeiten. Das Profil unterscheidet Controller und Target (Ziel) Geräte und gruppiert diese in folgende Kategorien:

- Kategorie 1: Abspiel- und Aufnahmegeräte
- Kategorie 2: Monitor/Verstärker
- Kategorie 3: Audio/Video Empfänger (z. B. Radio)
- Kategorie 4: Menu

Abb. 5.25 Gleichzeitige Übertragung eines Audio Streams und Kontrollkommandos zwischen verschiedenen Geräten



Des Weiteren definiert der Standard eine Vielzahl von Kontrollkommandos (Operation IDs) und legt fest, welche Geräte in welchen Gerätetypen diese Kommandos jeweils unterstützen müssen und welche optional sind. Standardisierte Kontrollkommandos sind z. B.: „select“, „up“, „right“, „root menu“, „setup menu“, „channel up“, „channel down“, „volume up“, „volume down“, „play“, „stop“, „pause“, „eject“, „forward“ und „backward“. Endgerätehersteller können auch selber Kommandos definieren, diese können jedoch nicht zwischen Geräten unterschiedlicher Hersteller verwendet werden.

Zwischen der Audio Streaming Session mit dem A2DP Profil und der Kontrollsitzung mit dem Remote Control Profil gibt es keine direkte Verbindung. Somit ist es möglich, in einem Piconet einen MP-3 Player für die Übertragung von Musik zu einem Kopfhörer zu verwenden, während Lautstärkekommandos und andere Anweisungen von einem dritten Gerät, wie z. B. einer Fernbedienung, an den MP-3 Player gesendet werden können. Dieses Szenario ist in Abb. 5.25 dargestellt.

5.6.5 Das Human Interface Device (HID) Profile

Eine Anwendung, die in den letzten Jahren an Bedeutung gewonnen hat, ist die Anbindung von Eingabegeräten wie Tastaturen und Mäusen an Notebooks und Tablets. Im Notebookbereich verwenden drahtlose Mäuse oft ein proprietäres Protokoll und einen proprietären USB Empfänger. Dies ist bei Tablets nicht möglich, da kein proprietärer USB Empfänger an ein solches Gerät angeschlossen werden kann. Da Tablets aber üblicherweise mit Bluetooth ausgestattet sind, können solche Eingabegeräte über das Human Interface Device (HID) Profil verwendet werden.

Das HID-Protokoll verwendet zwei L2CAP Verbindungen. Die erste Verbindung wird als ein Kontrollkanal verwendet und Daten werden synchron übertragen, d. h. auf jede Anfrage folgt immer eine Antwort, bevor die nächste Anfrage gesendet wird. Die zweite L2CAP Verbindung wird für einen HID Interrupt Kanal verwendet, über den asynchrone Informationen ausgetauscht werden, z. B. wenn der Anwender eine Taste drückt oder diese wieder loslässt. Da HID ein generisches Profil ist, werden Informationen in der SDP Datenbank verwendet, um herauszufinden, welche Ein- und Ausgabenachrichten ein Gerät unterstützt. Eingabenachrichten können z. B. Tastaturbenachrichtigungen oder Mausbewegungen sein. Ausgabenachrichten werden in umgekehrter Richtung gesendet, z. B. an Force Feedback Joysticks.

Da HID-Geräte üblicherweise batteriebetrieben sind, spielt der Stromverbrauch eine große Rolle. Auf der Bluetooth Seite aktivieren deshalb der Host und das HID Device den Bluetooth Sniff Modus, nachdem die L2CAP Verbindung und der Interrupt Kanal aufgebaut wurden. Eine typische Sniff Rate in der Praxis sind 40 ms. Zusätzlich kann Sniff Subrating verwendet werden, um den Stromverbrauch in Nutzungspausen der Tastatur und der Maus weiter zu senken. Abb. 5.26 zeigt Ausschnitte einer HID Input Nachricht, die von einer Tastatur zu einem Notebook gesendet wurde. Hier ist zu sehen, dass die Nachrichtenlänge nur 19 Bytes beträgt, trotz Nutzung der ACL, L2CAP und

```
Frame 176: 19 bytes on wire (152 bits), 19 bytes captured (152 bits)
Encapsulation type: Bluetooth H4 with linux header (99)
[...]
[Protocols in frame: hci_h4:bthci_acl:bt12cap:bthid]
Point-to-Point Direction: Received (1)

Bluetooth HCI H4
[Direction: Rcvd (0x01)]
HCI Packet Type: ACL Data (0x02)

Bluetooth HCI ACL Packet
.... 0000 0010 0011 = Connection Handle: 0x0023
..10 .... .... .... = PB Flag: First Automatically Flushable Packet (2)
00.. .... .... .... = BC Flag: Point-To-Point (0)
Data Total Length: 14

Bluetooth L2CAP Protocol
Length: 10
CID: Dynamically Allocated Channel (0x0041)
[PSM: HID-Interrupt (0x0013)]

Bluetooth HID Profile
1010 .... = Transaction Type: DATA (0x0a)
.... 00.. = Parameter reserved: 0x00
.... ..01 = Report Type: Input (0x01)
Protocol Code: Keyboard (0x01)
0.... .... = Modifier: RIGHT GUI: False
.0.... .... = Modifier: RIGHT ALT: False
..0.... .... = Modifier: RIGHT SHIFT: False
[...]
Reserved: 0x00
Keycode 1: a (0x04)
Keycode 2: <ACTION KEY UP> (0x00)
[...]

0000 02 23 20 0e 00 0a 00 41 00 a1 01 00 00 04 00 00
0010 00 00 00
```

Abb. 5.26 HID Input Nachricht einer Tastatur

HID Protokolle, die ineinander verschachtelt sind. Außerdem kann man in der Nachricht sehen, dass PSM 13 verwendet wurde, um den HID Interruptkanal aufzubauen. Die Nutzdaten in der Nachricht beschränken sich auf nur ein Byte (0x04h), welches das kleine „a“ repräsentiert, auf das der Nutzer auf der Tastatur gedrückt hat.

5.7 Fragen und Aufgaben

1. Welche maximale Geschwindigkeit bietet Bluetooth und von welchen Faktoren hängt diese ab?
2. Was bedeutet der Begriff Frequency Hopping Spread Spectrum (FHSS) und welche erweiterten Möglichkeiten bietet der Bluetooth 1.2 Standard?
3. Was ist der Unterschied zwischen Inquiry und Paging?
4. Welche Stromsparmodi gibt es bei Bluetooth?
5. Welche Aufgaben hat der Link Manager?
6. Wie können über das L2CAP Protokoll unterschiedliche Datenströme für unterschiedliche Anwendungen im Zeitmultiplex übertragen werden?
7. Welche Aufgaben hat die Service Discovery Datenbank?
8. Wie können mehrere Dienste gleichzeitig die RFCOMM Schicht verwenden?
9. Was ist der Unterschied zwischen der Bluetooth Authentifizierung und Autorisierung?
10. Warum gibt es eine Vielzahl unterschiedlicher Bluetooth Profile?
11. Welche Profile gibt es für die einfache und schnelle Übertragung von Dateien und Objekten zwischen zwei Bluetooth Endgeräten?
12. Wie unterscheidet sich das Hands-Free Profil vom SIM-Access Profil?

Lösungen sind auf der Website zum Buch unter <http://www.cm-networks.de> zu finden.



Mit GSM, dem Global System for Mobile Communication begann Anfang der 90er-Jahre ein beispielloser Wandel in der mobilen Kommunikation. Hatte das Vorläufersystem C-Netz in seiner Glanzzeit in Deutschland knapp eine Million Anschlüsse, wurden GSM-Netze in ihren Hochzeiten fast von der ganzen Bevölkerung genutzt. Dies war vor allem einer stetigen Weiterentwicklung in allen Bereichen der Telekommunikation und dem anhaltenden Preisverfall der digitalen Technik sowie der Endgeräte zu verdanken. Auch wenn heute GSM an Bedeutung verloren hat, sind in den meisten Ländern in Europa noch immer GSM Netze in Betrieb. Damit wird sichergestellt, dass auch Menschen mit alten Endgeräten weiterhin mobil telefonieren können. Außerdem wird GSM und die paketdatenorientierte Erweiterung GPRS im „Machine to Machine“ Umfeld noch sehr häufig genutzt. Darüber hinaus spielt GSM auch noch eine wichtige Rolle beim internationalen Roaming, da Voice over LTE sich hier noch nicht durchsetzen konnte. Und schließlich gibt es weiterhin Orte, an denen nur GSM Netzabdeckung vorhanden ist. Aus diesen Gründen ist es weiterhin wichtig, ein Verständnis von GSM und der paketdatenorientierten Erweiterung GPRS zu haben.

6.1 Leitungsvermittelnde Datenübertragung

GSM wurde ursprünglich als leitungsvermittelndes System konzipiert, das eine direkte und exklusive Verbindung zwischen zwei Teilnehmern auf jeder Schnittstelle zwischen zwei Netzwerkknoten herstellt. Ein erster Überblick über diese klassische Leitungsvermittlung folgt im nächsten Unterkapitel. Im Laufe der Zeit wurde jedoch diese physische Leitungsvermittlung virtualisiert und GSM Netzwerkknoten sind heute über breitbandige IP-Schnittstellen miteinander verbunden. Eine Sprachverbindung zwischen zwei Teilnehmern kann somit nicht mehr direkt einer Leitung oder einem Zeitschlitz auf einer

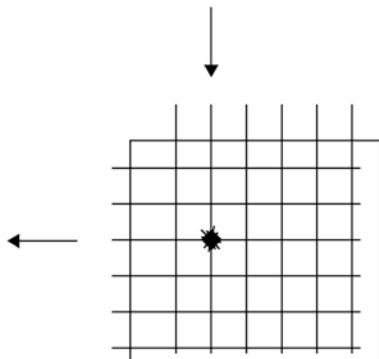
Leitung zugeordnet werden. Die Gründe dafür und weitere Details beschreibt dann das nächste Unterkapitel.

6.1.1 Klassische Leitungsvermittlung

GSM-Mobilfunknetze zählten ursprünglich genauso wie drahtgebundene Fernsprechnetze, auch Festnetze genannt, zu den leitungsvermittelnden Kommunikationsnetzen (Circuit Switched Networks). Beim Beginn eines Gespräches wurde dabei vom Netzwerk eine Leitung direkt von Teilnehmer zu Teilnehmer geschaltet, die diese dann exklusiv für sich verwenden konnten. In der Vermittlungsstelle (Switching Center) befand sich dafür, wie in Abb. 6.1 gezeigt, eine Verbindungsmautrix (Switching Matrix), die einen beliebigen Eingang mit einem beliebigen Ausgang verbinden konnte. Nachdem die Verbindung aufgebaut wurde, wurden alle Signale transparent über die Verbindungsmautrix zwischen den Teilnehmern ausgetauscht. Erst wenn einer der beiden Teilnehmer die Verbindung beendete, wurde die Vermittlungsstelle wieder aktiv und baute die Verbindung in der Verbindungsmautrix wieder ab. Diese Vorgehensweise war in einem Festnetz und einem Mobilfunknetz identisch.

Drahtgebundene Fernsprechnetze wurden anfangs nur für die Sprachdatenübertragung konzipiert, und es wurde ein analoger Kanal zwischen den Teilnehmern aufgebaut. Mitte der 80er-Jahre wurden diese Netze in Deutschland digitalisiert. Dies bedeutete, dass die Sprache nicht mehr analog von Ende zu Ende übertragen wurde, sondern in der Vermittlungsstelle digitalisiert und danach digital weiter übertragen wurde. Am anderen Ende wurden die digitalen Sprachdaten wieder in ein analoges Signal umgewandelt und über die Telefonleitung zum Endteilnehmer geschickt. Mit der Einführung von ISDN-Anschlüssen wurde dann diese Umwandlung von analog nach digital und zurück ins Endgerät (z. B. Telefon) verlegt, und die Sprache wurde erstmals Ende zu Ende digital übertragen.

Abb. 6.1 Verbindungsmautrix in einer Vermittlungsstelle



Für GSM wurde das Rad nicht neu erfunden. Statt ein komplett neues System zu entwickeln, wurde auf die bereits vorhandene Festnetztechnik in Form von Vermittlungsstellen und Weitverkehrsübertragungstechnik zurückgegriffen. Neu entwickelt werden musste jedoch die Technik für den eigentlichen Anschluss der Teilnehmer. Im Festnetz ist der Teilnehmeranschluss sehr einfach, für jeden Teilnehmer werden lediglich zwei Kabel benötigt. In einem Mobilfunknetzwerk jedoch kann der Teilnehmer seinen Standort frei wählen. Somit war es nicht mehr möglich, ein Gespräch immer über den gleichen Anschluss der Verbindungsmautrix zu einem Teilnehmer durchzuschalten.

Da ein Mobilfunknetzwerk wie ein Festnetz viele Vermittlungsstellen besaß, die jeweils ein begrenztes geografisches Gebiet versorgten, war in einem Mobilfunknetzwerk nicht einmal gewährleistet, dass ein Teilnehmer immer über die gleiche Vermittlungsstelle zu erreichen war. Somit konnte auch die im Festnetz verwendete Software für die Teilnehmerverwaltung und Gesprächsvermittlung für ein Mobilfunknetzwerk nicht weiterverwendet werden. Statt einer statischen 1:1-Zuweisung von Teilnehmer und Leitung wurde die Software in der Vermittlungsstelle um eine Mobilitätsmanagementkomponente erweitert. Diese verwaltete alle Teilnehmer und kannte den aktuellen Aufenthaltsort jedes erreichbaren Teilnehmers.

Da ein Teilnehmer auch während eines Gespräches den Aufenthaltsort ändern konnte und somit eventuell das Gespräch auf eine andere Leitung geschaltet werden musste, wurde auch die Gesprächsverwaltung neu entwickelt (Abb. 6.2).

Weiterverwendet wurden im Mobilfunknetzwerk ursprünglich jedoch fast die komplette Hardware einer Festnetzvermittlungsstelle, sowie die unteren Softwareschichten, die für das Schalten der Verbindungsmautrix und die Signalisierung zuständig waren. Somit war es auch nicht weiter verwunderlich, dass alle großen Netzwerkhersteller wie z. B. Ericsson, Nokia und viele weitere Hersteller, die heute jedoch vom Markt verschwunden sind, ihre Hardwareplattform für Vermittlungsstellentechnik sowohl für Festnetze, als auch für Mobilfunknetze anboten. Einzig die Software entschied darüber, für welchen Zweck die Vermittlungsstelle eingesetzt wurde.

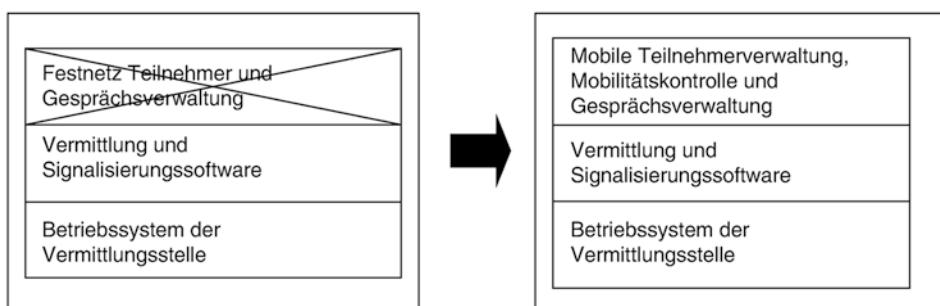


Abb. 6.2 Softwareänderungen von Festnetz- zu Mobilfunkvermittlung

6.1.2 Virtuelle Leitungsvermittlung über IP

War noch zu Anfang der GSM-Ära in den 1990er- Jahren Sprachkommunikation die dominierende Kommunikationsform, hat sich dies inzwischen mit dem Aufkommen des Internets grundlegend geändert. Zwar ist die Sprachkommunikation weiterhin sehr wichtig, es spielen jedoch nun auch Kommunikationsformen wie eMail, Instant Messaging (IM), soziale Netzwerke (z. B. Facebook), Blogs, Wikis und viele andere eine große Rolle. All diese Dienste haben gemein, dass sie das Internet Protocol (IP) verwenden und über das Internet Menschen global verbinden.

Während bei der Leitungsvermittlung ein exklusiver Kanal zwischen zwei Teilnehmern aufgebaut wird, basiert das Internet auf der Übertragung von Datenpaketen. Eine Leitung mit einer hohen Übertragungsbandbreite wird hier verwendet, um Datenpakete vieler Anwender zu übertragen. Jeder Netzwerknoten entscheidet anhand der Zieladresse eines Paketes, über welche Leitung es als nächstes weiter übertragen wird. Diese Entwicklung führte dazu, dass Netzbetreiber über viele Jahre zwei Netzwerke nebeneinander betrieben: Ein leitungsorientiertes Netzwerk für Sprachübertragung und ein paketorientiertes Netzwerk für internetbasierte Dienste.

Da der gleichzeitige Betrieb von zwei unterschiedlichen Netzwerkarten sehr ineffizient und teuer ist, haben Netzbetreiber in den 2010er Jahren die leitungsvermittelnde Verbindungsmatrix durch ein sogenanntes Media Gateway ersetzt. Damit wurde erreicht, dass Sprachverbindungen im Kernnetzwerk nicht mehr leitungsvermittelnd übertragen werden, sondern in IP-Paketen. Eine physisch exklusive Verbindung zwischen zwei Teilnehmern wird somit nun virtuell über einen IP-Paketdatenstrom übertragen.

Das physische Vorhandensein eines leitungsvermittelnden Netzwerkes ist somit nicht länger nötig und der Netzbetreiber kann sich auf den Betrieb eines einzigen, IP-basierten Netzwerkes beschränken. Dieser Ansatz ist unter dem Namen Bearer Independent Core Network (BICN) bekannt. An der grundsätzlichen Funktionsweise des GSM-Netzwerkes änderte diese Virtualisierung jedoch nichts. Unterschiede ergaben sich hauptsächlich für die unteren Protokollsichten der Gesprächssignalisierung und den unteren Protokollsichten bei der Sprachübertragung. Beides wird nachfolgend noch genauer erläutert.

Auch im GSM-Radionetzwerk wurden leitungsvermittelnde Übertragungswege durch breitbandige IP Verbindungen ersetzt. GSM nutzt dort nun die gleichen Leitungen wie die LTE und 5G Netzwerkkomponenten am gleichen Senderstandort. Die Luftschnittstelle zwischen Endgerät und Netzwerk ist von der Umstellung auf IP nicht berührt. Für Mobiltelefone und andere mobile Geräte war diese Migration somit völlig transparent.

6.2 Standards

Da sich im weltweiten Markt für Telekommunikationsnetzwerke auch schon in den 1990er Jahren viele Firmen um Aufträge der Netzwerkbetreiber bemühten, war eine Standardisierung der Schnittstellen und technischen Vorgänge notwendig. Ohne diese Standards, die unter anderem für leitungsvermittelnde Sprachnetzwerke von der International Telecommunication Union (ITU) definiert wurden, wäre eine länderübergreifende Telefonie nicht möglich gewesen, und Netzwerkbetreiber wären fest an einen Netzwerklieferanten gebunden gewesen.

Einer der wichtigsten ITU Standards ist das in Abschn. 6.4 vorgestellte Signalisierungssystem SS-7 für die Gesprächsvermittlung. Viele ITU Standards repräsentieren jedoch nur den kleinsten gemeinsamen internationalen Nenner. Jedes Land behielt sich vor, nationale Erweiterungen vorzunehmen. Dies verursachte in der Praxis enorme Kosten bei der Softwareentwicklung, da für jedes Land spezielle Erweiterungen nötig waren. Auch der Übergang zwischen Netzen unterschiedlicher Länder wurde dadurch sehr erschwert.

Mit GSM wurde zum ersten Mal ein einheitlicher Standard in Europa für die mobile Kommunikation geschaffen, der später auch weltweit übernommen wurde. Diesem Umstand ist es zu verdanken, dass Teilnehmer heute weltweit in allen GSM-Netzen, die ein sogenanntes Roamingabkommen mit seinem Heimatnetz abgeschlossen haben, telefonieren und mobil Daten übertragen können. Auch wurde es so möglich, die Entwicklungskosten wesentlich zu reduzieren, da die Systeme ohne große Modifikationen in alle Welt verkauft werden können. Dem European Telecommunication Standards Institute (ETSI), das neben GSM auch noch viele weitere Telekommunikationsstandards für Europa spezifiziert hat, kam dabei eine wesentliche Rolle bei der Erarbeitung dieser Standards zu. Die ETSI GSM Standards umfassen dabei eine Vielzahl von unterschiedlichen Standarddokumenten, auch Technical Specifications (TS) genannt, die jeweils einen Teil des Systems beschreiben. Da GSM zunehmend auch international verwendet wurde und es zu Beginn der UMTS-Standardisierung in den 1990er Jahren absehbar war, dass auch dieser über Europa hinaus große Bedeutung erlangen würde, gründete ETSI zusammen mit weiteren internationalen Standardisierungsgremien aus aller Welt das 3rd Generation Partnership Project (3GPP). Dieses Gremium kümmert sich seither um die Standardisierung von GSM, UMTS, LTE und 5G. An dieser Stelle sei jedoch angemerkt, dass 3GPP sich heute hauptsächlich auf die Weiterentwicklung von 5G und dessen Nachfolgetechnologien fokussiert.

6.3 Übertragungsgeschwindigkeiten

Die kleinste Geschwindigkeitseinheit in einem klassischen leitungsvermittelnden Telekommunikationsnetzwerk war der Digital Signal 0 (DS0)-Kanal. Dieser hatte eine feste Übertragungsgeschwindigkeit von 64 kbit/s. Über einen solchen Kanal konnten Sprache oder auch Daten übertragen werden. Aus diesem Grund wurde üblicherweise nicht von einem Sprachkanal, sondern allgemein von einem Nutzdatenkanal gesprochen.

Die Referenzeinheit in einem klassischen Telekommunikationsnetzwerk war die E-1-Verbindung, die zumeist über Twisted Pair oder Koaxialkabel geführt wurde. Die Bruttodatenrate einer E-1-Verbindung betrug 2.048 Mbit/s. Diese Bruttodatenrate wurde in 32 Zeitschlitzte (Timeslots) à 64 kbit/s aufgeteilt, in denen jeweils unabhängige Datenströme (DS0s) übertragen wurden.

Ein Zeitschlitz pro E-1 wurde für die Synchronisation benötigt und konnte somit keinen DS0 übertragen. Somit standen pro E-1-Verbindung 31 Zeitschlitzte zur Verfügung. Davon konnten beispielsweise 29 oder 30 Zeitschlitzte für die Nutzdatenübertragung verwendet werden und ein oder zwei für die nötigen Signalisierungsdaten. Mehr zu Signalisierungsdaten in Abschn. 6.4 über das SS-7-Protokoll (Abb. 6.3).

Eine einzelne E-1-Verbindung mit 31 DS0s reichte für Verbindungen zwischen Vermittlungsstellen nicht aus. Eine Alternative war z. B. eine E-3 Verbindung, ebenfalls über Twisted Pair oder Koaxialkabel mit einer Geschwindigkeit von 34.368 Mbit/s. Dies entsprach 512 DS0s.

Für höhere Übertragungsgeschwindigkeiten und für große Übertragungsdistanzen wurden optische Systeme verwendet, die nach dem Synchronous Transfer Mode (STM) Standard arbeiteten. Die nachfolgende Tabelle zeigt einige Übertragungsraten und die Anzahl der Nutzdatenkanäle à 64 kbit/s (DS0s), die pro Glasfaserpaar übertragen werden konnten.

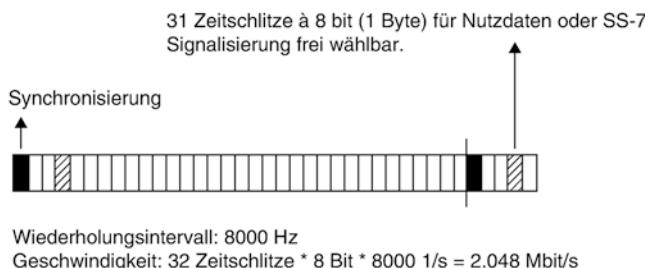


Abb. 6.3 Zeitschlitzarchitektur einer E-1-Verbindung

Typ	Geschwindigkeit (Mbit/s)	Anzahl 64 kbit/s Verbindungen (ca.)
STM-1	155,52	2300
STM-4	622,08	9500
STM-16	2488,32	37.000
STM-64	9953,28	148.279

Die hier vorgestellten Übertragungssysteme und Übertragungsgeschwindigkeiten wurden in den meisten Ländern dieser Welt verwendet. Lediglich Nordamerika und Japan bildeten eine Ausnahme und verwendeten eigene Übertragungsstandards.

Bei der virtuellen Leitungsvermittlung über das IP Protokoll kommen heute im Nahbereich zwischen unterschiedlichen Netzwerkknoten am selben Standort optische Ethernet Verbindungen mit einer Geschwindigkeit im dreistelligen GBit/s Bereich zum Einsatz.

6.4 Das Signalisierungssystem Nr. 7

Für den Aufbau, den Erhalt und den Abbau einer Verbindung müssen zwischen den Geräten Signalisierungsinformationen ausgetauscht werden. Bei Teilnehmern mit analogem Festnetztelefon fand diese Signalisierung durch Abnehmen oder Auflegen des Handapparats statt, die gewünschte Rufnummer wurde dem Netzwerk dann früher per Pulswahl und später mit der schnelleren Dual Tone Multi Frequency (DTMF) Tonwahl übermittelt. Bei GSM-Mobiltelefonen erfolgt diese Signalisierung getrennt vom Sprachkanal über einen eigenen Signalisierungskanal. Die Informationen wie zum Beispiel die Telefonnummer werden digital in Nachrichtenpaketen übertragen.

Waren mehrere Netzwerkkomponenten wie z. B. mehrere Vermittlungsstellen am Verbindungsaufbau beteiligt, mussten zwischen diesen ebenfalls Signalisierungsinformationen ausgetauscht werden. Für diese Signalisierung wurde in digitalen Fernsprechnetzwerken das Signalisierungssystem Nr. 7 (SS-7) verwendet. Auch der GSM-Mobilfunkstandard verwendet SS-7, wobei jedoch zusätzliche SS-7-Protokolle bei ETSI standardisiert wurden, die für die zusätzlichen Aufgaben eines Mobilfunknetzwerkes notwendig waren.

Grundsätzlich gibt es bei SS-7 drei unterschiedliche Netzwerkknoten:

Service Switching Points: SSPs sind Vermittlungsstellen, also Netzwerkelemente, über die Daten- und Sprachverbindungen aufgebaut, zugestellt oder weitergeleitet werden können.

Service Control Points: SCPs sind Datenbanken mit dazugehörender Software, die den Aufbau einer Verbindung beeinflussen können. Bei GSM werden SCPs z. B. für die Speicherung des aktuellen Aufenthaltsorts jedes Teilnehmers verwendet. Bei einem Verbindungsaufbau zu einem mobilen Teilnehmer müssen dann die Vermittlungsstellen zuerst dort nachfragen, wo sich der Teilnehmer befindet. Mehr hierzu im Abschn. [6.6.3](#) über das Home Location Register.

Signaling Transfer Points: STPs sind für das Weiterleiten von Signalisierungsnachrichten zwischen SSPs und SCPs notwendig, da nicht jeder Netzwerkknoten eine dedizierte Verbindung zu jedem anderen Knoten unterhalten kann. Von der prinzipiellen Funktionsweise kann man diese Knoten mit IP-Routern im Internet vergleichen, die ebenfalls Pakete in unterschiedliche Netze an unterschiedliche Geräte weiterleiten. Im Gegensatz zu diesen befördern STPs aber keine Nutzdaten wie Datenrufe oder Telefongespräche, sondern nur die zum Aufbau, Abbau oder Aufrechterhaltung einer Verbindung notwendigen Signalisierungsinformationen.

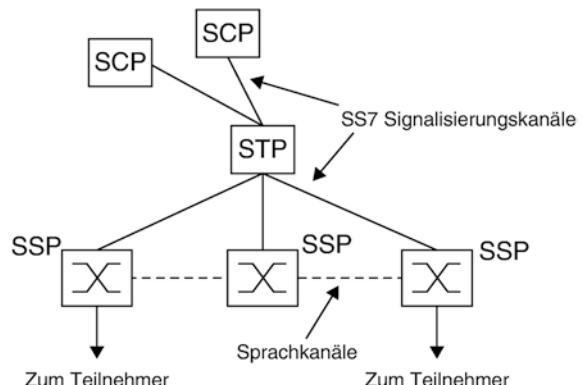
Auch bei der virtuellen Leitungsvermittlung über ein IP-Netzwerk wird das SS-7-Protokoll für die Übertragung von Signalisierungsnachrichten zwischen den Netzwerkknoten verwendet. Die Nachrichten werden jedoch nicht in speziell dafür konfigurierten Zeitschlitzten eines E-1 basierten Netzwerkes übertragen, sondern in IP-Paketen. Die nachfolgenden Unterkapitel beschreiben nun zuerst den klassischen, heute nicht mehr verwendeten SS-7-Protokollstack und danach, wie das SS-7-Protokoll über IP übertragen wird (Abb. 6.4).

6.4.1 Klassischer SS-7-Protokollstack

Das klassische Signalisierungssystem Nr. 7 (SS-7) basierte auf einer Anzahl von Protokollen, die schichtweise aufeinander aufgebaut sind. Das bekannteste und meistverwendete Modell zur Erklärung der Protokolle auf den unterschiedlichen Schichten ist dabei das OSI 7-Schichten-Modell.

Bei der klassischen Leitungsvermittlung wurden SS-7 Signalisierungsnachrichten in einem oder mehreren Zeitschlitzten einer E-1-Verbindung, wie in Abb. 6.3 gezeigt, übertragen. Das dafür nötige Protokoll auf den unteren drei Netzwerkschichten hieß Message Transfer Part (MTP) und bestand aus drei Teilen:

Abb. 6.4 Ein SS-7-Netzwerk mit einem STP, zwei SCP Datenbanken und 3 Vermittlungsstellen



Das Message Transfer Part – 1 (MTP-1) Protokoll beschrieb auf Schicht 1 des OSI-Modells die Eigenschaften des Übertragungsmediums. Diese Schicht wird auch Physical Layer genannt. Dazu gehört unter anderem die Definition der möglichen Kabelarten, die zu verwendenden Signalpegel, mögliche Übertragungsgeschwindigkeiten, etc.

Auf Schicht 2, dem Data Link Layer, wurden Nachrichten in Pakete eingepackt und mit einer Start- und Endekennung versehen.

Der Network Layer auf Schicht 3 war für die Weiterleitung von Datenpaketen zuständig. Jedes Paket wurde dazu mit einer Quell- und Zieladresse versehen. Auf diese Weise konnten Netzwerknoten Datenpakte weiterleiten (routen), die nicht für sie selber bestimmt sind. Im SS-7-Protokollstapel war das MTP-3 Protokoll hierfür zuständig. Für Leser, die bereits Kenntnisse in der TCP/IP-Welt haben, sei an dieser Stelle erwähnt, dass das MTP-3 Protokoll sehr gut mit dem IP Protokoll verglichen werden kann. Statt einer IP-Adresse verwendete das MTP-3 Protokoll aber so genannte Point Codes, um Quelle und Ziel einer Nachricht eindeutig zu identifizieren (Abb. 6.5).

Auf Layer 4–7 kamen nun je nach Bedarf unterschiedliche Protokolle zum Einsatz. Diente die Signalisierungsrichtung zum Auf- oder Abbau eines Übertragungskanals in der klassischen Leitungsvermittlung, wurde das ISDN User Part (ISUP) Protokoll verwendet.

Abb. 6.6 zeigt, wie ein Gespräch zwischen zwei Teilnehmern aufgebaut wurde. Teilnehmer A ist dabei ein Mobilfunkteilnehmer und B ein Festnetzteilnehmer. Während A über eine Mobilfunkvermittlungsstelle verbunden ist, die auch Mobile Switching Center (MSC) genannt wird, ist B ein Festnetzteilnehmer.

Um Teilnehmer B zu erreichen, übermittelte A seiner MSC die Telefonnummer von B. Anhand der Vorwahl von B erkannte die MSC, dass B ein Festnetzteilnehmer war. Für die Sprachübertragung gab es in Abb. 6.6 dorthin eine direkte Verbindung. Dies konnte auch durchaus in der Praxis vorkommen, wenn zum Beispiel von einem Mobiltelefon in München ein Festnetztelefon ebenfalls in München angerufen wurde.

Da es sich bei B um einen Festnetzteilnehmer handelte, musste die MSC nun einen Nutzdatenkanal für die Sprachübertragung zur Festnetzvermittlungsstelle aufbauen. Dies geschah über das ISUP-Protokoll mit einer Initial Address Message (IAM). Diese Nach-

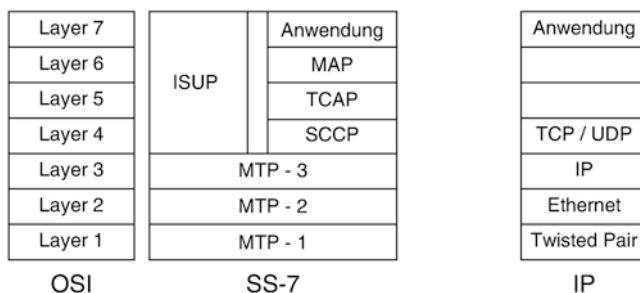


Abb. 6.5 SS-7-Protokollstack im Vergleich zum IP-Protokollstack

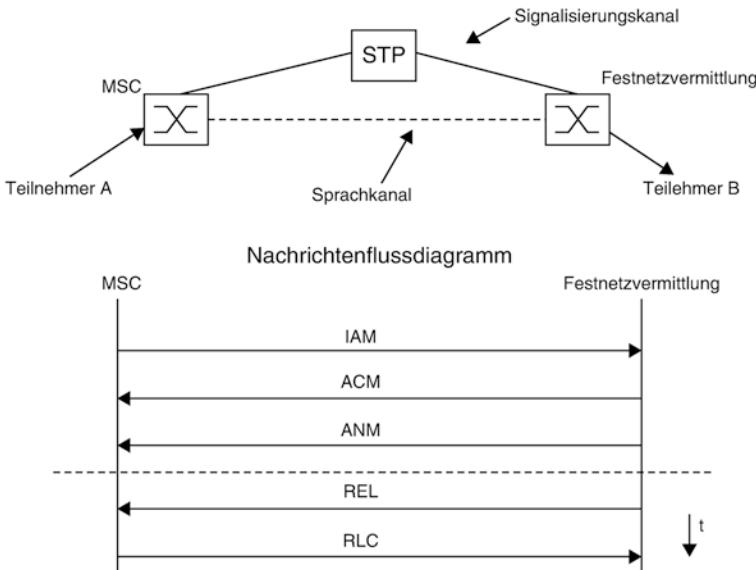


Abb. 6.6 Aufbau einer Verbindung zwischen Vermittlungsstellen

richt enthielt unter anderem die Telefonnummer von B, sowie die Information, welcher Nutzdatenkanal zwischen den zwei Vermittlungsstellen für das Gespräch verwendet werden sollte. Die IAM wurde dabei nicht direkt zwischen den Vermittlungsstellen ausgetauscht, sondern lief über einen STP.

Die Festnetzvermittlungsstelle empfing diese Nachricht, analysierte die darin enthaltene Rufnummer und stellte die Verbindung zu Teilnehmer B her. Sobald dessen Telefon klingelte, wurde eine Address Complete Message (ACM) an die MSC zurückgeschickt. Die MSC wusste somit, dass die Rufnummer korrekt war und Teilnehmer B gerufen wird.

Beantwortete Teilnehmer B den Anruf durch Abnehmen des Telefons, schickte die Festnetzvermittlungsstelle eine Answer Message (ANM) an die MSC zurück, und das Telefongespräch begann.

Legte Teilnehmer B am Ende des Gespräches auf, schickte die Festnetzvermittlungsstelle eine Release Message (REL) an die MSC. Diese schickte daraufhin eine Release Complete Message (RLC) als Quittung zurück. Beendete Teilnehmer A das Gespräch, liefen diese Nachrichten in die jeweils andere Richtung.

Für die Kommunikation zwischen Vermittlungsstellen (SSPs) und Datenbanken (SCPs) kam auf Schicht 4 das Signalling Connection and Control Part (SCCP) zum Einsatz. Seine Funktionsweise war in weiten Teilen sehr ähnlich zum TCP und UDP-Protokoll in der IP-Welt. Über Protokolle der Schicht 4 konnten unterschiedliche Anwendungen auf einem System unterschieden werden. In TCP und UDP gibt es dazu

sogenannte Ports. Wird ein PC z. B. als Web Server und gleichzeitig als FTP Server verwendet, sind diese Server zwar über die gleiche IP-Adresse erreichbar, verwenden aber unterschiedliche Port-Nummern. Anhand dieser Port Nummer kann dann der Protokollstapel entscheiden, an welche Applikation das Datenpaket weitergegeben wird. In der SS-7-Welt wurde diese Aufgabe von SCCP erledigt. Statt Port Nummern wurden hier jedoch Subsystem-Nummern (SSNs) an unterschiedliche Applikationen vergeben.

Für den Zugriff auf Datenbanken wurde für SS-7 das Transaction Capability Application Part (TCAP) entwickelt. Dies stellte für SCP-Datenbankabfragen eine Anzahl von unterschiedlichen Nachrichtenbausteinen bereit, um Abfragen möglichst einheitlich zu gestalten.

6.4.2 Spezielle SS-7-Protokolle für GSM

Neben den bereits genannten SS-7-Protokollen, die sowohl in einem Festnetz, wie auch im GSM Mobilfunknetz verwendet wurden, waren für ein GSM-Mobilfunknetz eine Reihe weiterer Protokolle notwendig, um den zusätzlichen Aufgaben eines Mobilfunknetzwerkes Rechnung zu tragen.

Das Mobile Application Part (MAP)-Protokoll: Dieses Protokoll wurde in ETSI TS 09.02 spezifiziert und wird auch heute noch für die Kommunikation zwischen einer MSC und dem Home Location Register (HLR) verwendet, das Teilnehmerinformationen verwaltet. Das HLR wird zum Beispiel gefragt, wenn eine MSC eine Verbindung zu einem mobilen Benutzer herstellen soll. Das HLR liefert in einem solchen Fall der MSC die Information zurück, wo sich der gewünschte Teilnehmer gerade aufhält. Mit dieser Information konnte dann die MSC das Gespräch zur aktuellen Vermittlungsstelle dieses Teilnehmers mit den in Abb. 6.6 beschriebenen ISUP Nachrichten herstellen.

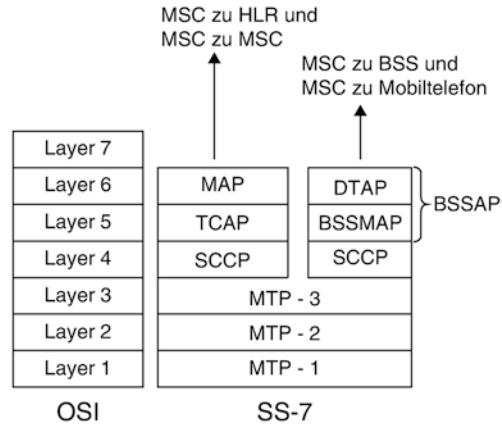
MAP wird außerdem zwischen MSCs verwendet, wenn sich ein Teilnehmer während eines Gesprächs in das Versorgungsgebiet einer anderen MSC bewegt und die Verbindung dorthin weitergeleitet werden muss.

Wie in Abb. 6.7 dargestellt ist, setzte das MAP-Protokoll zunächst auf die bereits beschriebenen TCAP, SCCP und MTP-Protokolle auf.

Das Base Station Subsystem Mobile Application Part (BSSMAP): Dieses Protokoll dient auch heute noch der Kommunikation zwischen MSC und dem Radionetzwerk. Es wird zum Beispiel verwendet, um dem Radionetzwerk die Anweisung zu geben, einen dedizierten Funkkanal für eine neue Verbindung zu einem Mobilfunkteilnehmer herzustellen. Da es sich hier nicht um Datenbankabfragen wie beim MAP-Protokoll handelt, setzt BSSMAP nicht auf TCAP, sondern direkt auf SCCP auf.

Direct Transfer Application Part (DTAP): Über dieses Protokoll kann ein Endgerät, im englischen auch Mobile Station (MS) genannt, direkt mit einer MSC Nachrichten austauschen. Um eine Verbindung zu einem anderen Teilnehmer aufzubauen, wird beispielsweise die SETUP-Nachricht verwendet. Diese enthält unter anderem die Telefon-

Abb. 6.7 Erweiterungen des SS-7-Protokollstacks für GSM



nummer des Gesprächspartners. Alle Netzwerkelemente zwischen Endgerät und MSC leiten diese Nachrichten transparent weiter.

6.4.3 IP-basierter SS-7-Protokollstack

Bei der heute üblichen Verwendung eines IP-Netzwerkes für die Übertragung von SS-7-Signalisierungsnachrichten werden die Protokolle MTP-1 und MTP-2 durch den IP-Protokollstack ersetzt. Abb. 6.8 zeigt die beiden Stacks im Vergleich.

Im Internet folgt auf das IP Protokoll auf Schicht 4 normalerweise UDP (Universal Datagram Protocol) oder TCP (Transmission Control Protocol). Für die Übertragung von SS-7-Nachrichten wurde jedoch ein neues Protokoll spezifiziert, das Stream Control Transmission Protocol. Dieses bietet gegenüber TCP und UDP Vorteile bei vielen gleichzeitigen virtuellen Signalisierungsverbindungen zwischen zwei Netzwerknoten.

Auf SCTP folgt das M3UA (MTP-3 User Adaptation Layer) Protokoll. Wie der Name andeutet, werden auf dieser Schicht Informationen übertragen, die im klassischen SS-7-Netzwerk über MTP-3 übertragen wurden. Für höhere SS-7-Protokollsichten, wie z. B. SCCP, simuliert M3UA alle Funktionalitäten von MTP-3 und macht somit für alle höheren Schichten die Übertragung über IP transparent.

Oft wird der IP-basierte SS-7-Protokollstack oder die IP-basierte Übertragung von SS-7-Nachrichten als SIGTRAN bezeichnet. Diese Bezeichnung stammt vom Namen der IETF (Internet Engineering Task Force) Arbeitsgruppe, die für die Definition der Protokolle gegründet wurde.

Wie in Abschn. 6.1.1 beschrieben, dient das ISUP-Protokoll für den Aufbau von Sprachverbindungen zwischen Vermittlungsstellen. Da in einem IP-basierten Netzwerk jedoch statt eines leitungsvermittelten Nutzdatenkanal für eine Sprachverbindung ein IP-Datenstrom verwendet wird, musste auch das ISUP-Protokoll entsprechend angepasst

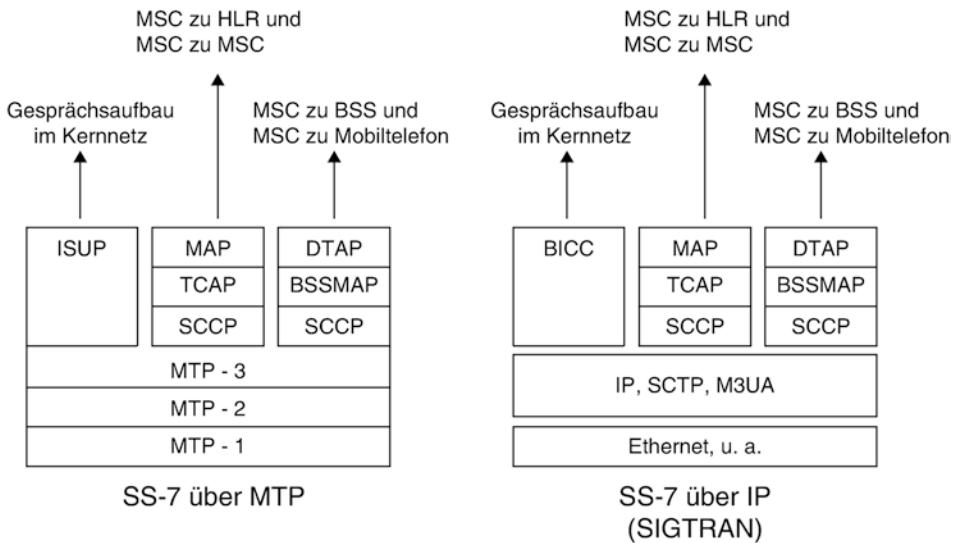


Abb. 6.8 Vergleich von klassischem und IP-basiertem SS-7-Protokollstack

werden. Das Resultat war das Bearer Independent Call Control (BICC)-Protokoll, das ISUP sehr ähnlich ist.

In wenigen Fällen mag es in der Praxis heute noch vorkommen, dass SS-7 über IP noch nicht an den Übergängen zwischen Netzen unterschiedlicher Betreiber verwendet wird. In diesen Fällen werden an den Schnittstellen zwischen klassischer SS-7-Kommunikation und IP-basierter SS-7-Übertragung Signaling Gateways (SGs) verwendet. Diese übersetzen die unteren Ebenen des Stacks und machen somit den Übergang für beide Seiten transparent.

Media Gateways, die sowohl E-1-basierte leitungsvermittelte Kanäle und breitbandige IP-Verbindungen für die Weiterleitung von Sprachverbindungen haben, werden für die Umsetzung der Sprachkanäle zwischen den unterschiedlichen Technologien verwendet.

6.5 Die GSM Subsysteme

Ein GSM Netzwerk wird in 3 unterschiedliche Subsysteme eingeteilt:

Das Basestation Subsystem (BSS), auch Radio-Netzwerk genannt, enthält alle Elemente und Funktionen, die für die Verbindung zwischen Netzwerk und mobilen Teilnehmern über die Funkschnittstelle, die auch Luftschnittstelle genannt wird, notwendig sind.

Das Network Subsystem (NSS), auch Core Network oder Kernnetzwerk genannt, enthält alle Komponenten für die Vermittlung von Gesprächen, für die Teilnehmerverwaltung und das Mobilitätsmanagement.

Das Intelligent Network Subsystem (IN), besteht aus SCP-Datenbanken, die zusätzliche Dienste zur Verfügung stellen. Einer der wichtigsten IN Dienste in einem Mobilfunknetzwerk ist beispielsweise der Prepaid Service, der das Abtelefonieren eines zuvor eingezahlten Guthabens in Echtzeit erlaubt.

6.6 Das Network Subsystem

Die wichtigste Aufgabe des NSS sind Verbindungsauflaufbau, Verbindungskontrolle und Vermittlung von Verbindungen zwischen unterschiedlichen mobilen Vermittlungsstellen (MSC) und anderen Netzwerken. Andere Netzwerke können z. B. das nationale Festnetz, das im englischen auch Public Standard Telephone Network (PSTN) genannt wird, internationale Festnetze sowie andere nationale und internationale Mobilfunknetze sein. Außerdem umfasst das NSS die Teilnehmerverwaltung. Die dazu notwendigen Komponenten und Prozesse werden in den nächsten Abschnitten beschrieben und sind schematisch in Abb. 6.9 für eine klassische Netzwerkarchitektur dargestellt, sowie in Abb. 6.11 für ein auf IP basierendes Kernnetzwerk.

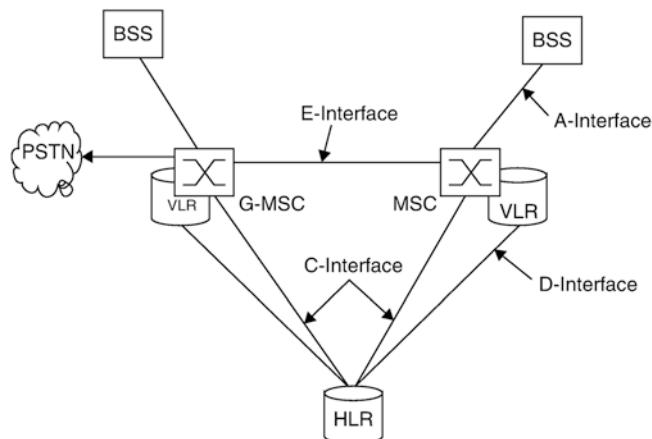


Abb. 6.9 Schnittstellen und Komponenten im NSS

6.6.1 Die Mobile Vermittlungsstelle (MSC), Server und Gateway

Die Mobile Vermittlungsstelle, auch Mobile Switching Center (MSC) genannt, ist das zentrale Element eines GSM Mobilfunknetzwerkes, das auch Public Land Mobile Network (PLMN) genannt wird.

Alle Verbindungen zwischen Teilnehmern, auch wenn diese sich in der gleichen Funkzelle befinden, werden immer über eine MSC geleitet und kontrolliert. Diese Aufgabe wird Call Control (CC) genannt und umfasst folgende Aufgaben:

- Registrieren des Teilnehmers (Registration): Beim Einschalten des Endgeräts registriert sich dieses im Netzwerk und ist anschließend für alle Teilnehmer erreichbar.
- Der Verbindungsaufbau (Call Routing) zwischen zwei Teilnehmern.
- Weiterleiten von Kurznachrichten (SMS).

Da sich Teilnehmer im Mobilfunknetzwerk frei bewegen können, ist die MSC auch für die Mobilitätskontrolle (Mobility Management) zuständig. Man unterscheidet zwei Zustände:

- Authentifizieren des Teilnehmers bei Verbindungsaufnahme (Authentication): Dies ist notwendig, da ein Teilnehmer nicht mehr wie im Festnetz anhand der verwendeten Leitung identifiziert werden kann. Weitere Information über die Teilnehmerauthentifizierung im Zusammenhang mit dem Authentication Center sind in Abschn. 6.6.4 zu finden.
- Besteht keine aktive Verbindung zwischen Netzwerk und Endgerät, muss das Endgerät eine Änderung seiner Position dem Netzwerk mitteilen, um für den Fall eines eingehenden Anrufs oder einer Kurzmitteilung (SMS) auffindbar zu sein. Dieser Vorgang wird Location Update genannt und in Abschn. 6.8.1 näher beschrieben.
- Bewegt sich ein Teilnehmer während einer bestehenden Verbindung, sorgt die MSC dafür, dass die Verbindung nicht abbricht und in die jeweils geeigneten Zellen weitergegeben wird. Dieser Vorgang wird Handover genannt und in Abschn. 6.8.3 näher beschrieben

Um mit anderen MSCs und Netzwerkkomponenten zu kommunizieren, ist die MSC mit diesen über standardisierte Schnittstellen verbunden. Dies ermöglicht, dass die Netzwerkkomponenten von unterschiedlichen Netzwerkherstellern stammen können. Die nachfolgend vorgestellten Schnittstellen wurden früher über Zeitschlüsse in E-1-Leitungen übertragen und heute über ein IP-basiertes Netzwerk. Wie anfangs beschrieben, sind dabei nur die unteren Protokollsichten unterschiedlich, auf Applikationsebene verhalten sich beide Varianten gleich.

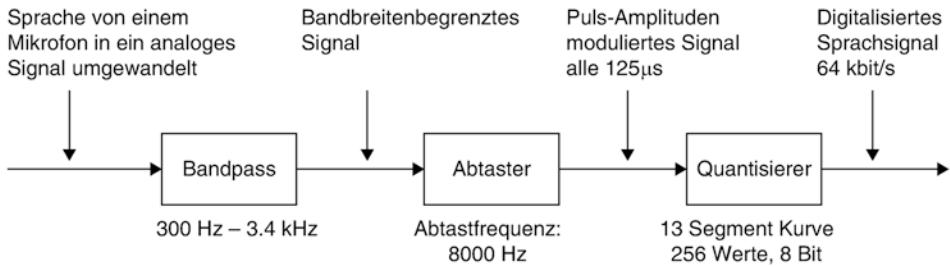
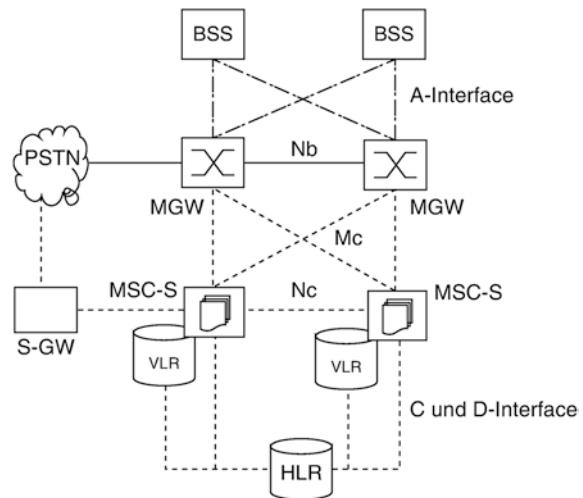


Abb. 6.10 Sprachdigitalisierung

Abb. 6.11 Release 4 Bearer Independent Core Network (BICN) Architektur



Das BSS, über das alle Teilnehmer mit dem Mobilfunknetzwerk kommunizieren, wurde im klassischen SS-7 System über eine Anzahl von 2 MBit/s E-1-Leitungen mit einer MSC verbunden. Diese Schnittstelle wird A-Interface genannt und heute typischerweise über SS-7 over IP basierte optische Verbindungen realisiert. Wie in Abschn. 6.4 bereits gezeigt, werden auf dem A-Interface das BSSMAP und DTAP-Protokoll verwendet.

Da eine MSC nur eine begrenzte Vermittlungsleistung und Rechenkapazität besitzt, bestand ein großes Mobilfunknetzwerk früher aus dutzenden voneinander unabhängiger MSCs. Jede MSC versorgte dabei einen eigenen geografischen Bereich. Auch zwischen den MSCs wurden E-1-Verbindungen verwendet, die wieder optisch gebündelt und weitergeleitet wurden. Dank steigender Rechenleistungen kommen in heutigen Mobilfunknetzwerken weitaus weniger MSCs zum Einsatz, deren Komponenten aus Redundanzgründen auch physikalisch getrennt an mehreren Orten aufgestellt sein können. Da sich ein Teilnehmer während eines Gesprächs auch über die geografische Versorgungsgrenze einer MSC hinaus bewegen kann, muss das Gespräch entsprechend

an die für dieses Gebiet zuständige MSC weitergeben werden können. Diese Gesprächsweitergabe wird auch Handover genannt. Die dafür notwendigen Signalisierungs- und Sprachverbindungen werden E-Interface genannt. Als Protokoll zwischen den MSCs kommt ISUP für die Verbindungskontrolle und MAP für die Signalisierung des Handovers zum Einsatz. Näheres hierzu in Abschn. 6.8.3.

Über das C-Interface ist die MSC mit der Teilnehmerdatenbank, dem Home Location Register (HLR) des Mobilfunknetzwerkes verbunden. Während zum A-Interface und dem E-Interface immer auch zwingend Sprachkanäle gehören, ist das C-Interface eine reine Signalisierungsverbindung.

Wie im Abschn. 6.3 beschrieben, wurde in klassischen digitalen leitungsvermittelnden Festnetz- und Mobilfunksystemen ein Sprachkanal im Kernnetz in einem 64 kbit/s E-1-Zeitschlitz übertragen. Ein analoges Sprachsignal musste dazu aber zuerst digitalisiert werden (Abb. 6.10).

Bei einem analogen Festnetzanschluss erfolgte die Digitalisierung in der Vermittlungsstelle, bei einem ISDN-Anschluss und bei einem GSM-Teilnehmer bereits im Endgerät. Auf IP-basierten Schnittstellen wird ein Sprachkanal heute ebenfalls mit einer Nutzdatengeschwindigkeit von 64 kbit/s übertragen, jedoch nicht in einem Timeslot, sondern in einem IP Datenstrom, der gleichzeitig mit vielen anderen Sprachverbindungen auf der gleichen breitbandigen Leitung übertragen wird.

Ein analoges Sprachsignal wird auch heute noch in 3 Schritten digitalisiert: Im ersten Schritt wird die Bandbreite des analogen Signals auf 300 Hz–3400 Hz begrenzt, damit dies später auch in einem 64 kbit/s Timeslot übertragen werden kann. Danach wird das analoge Signal 8000 mal pro Sekunde abgetastet und der Wert einem Quantisierer übergeben. Der Quantisierer wandelt nun den analog abgetasteten Wert in einen 8 Bit digitalen Wert von 0–255 um.

Je höher die Amplitude des abgetasteten Wertes, also je lauter das Sprachsignal, desto größer der digitale Wert. Um auch leise Töne möglichst gut zu übertragen, erfolgt die Quantisierung nicht linear im gesamten Bereich, sondern nur abschnittsweise. Für kleine Amplituden, also leise Sprache, werden dabei wesentlich mehr digitale Werte verwendet, als für laute Töne.

Das so digitalisierte Signal wird Pulse Code Modulated (PCM)-Signal genannt. Für welche Lautstärke welcher digitale Wert zugeordnet ist, beschreibt in Europa der a-Law Standard, in Nordamerika der μ -Law Standard. Die Verwendung unterschiedlicher Standards erschwert natürlich die Sprachübertragung zwischen Netzen, die jeweils den anderen Standard verwenden. Zwischen Deutschland und Nordamerika muss das Sprachsignal deshalb an den Netzübergängen entsprechend umkodiert werden.

Da die MSC alle Verbindungen kontrolliert, ist sie auch für die spätere Abrechnung (Billing) zuständig. Zu diesem Zweck erstellt die MSC für jedes Gespräch einen sogenannten Billing Record, der nach dem Gespräch gespeichert und zum Abrechnungssystem übertragen wird. Der Billing Record enthält dabei unter anderem die Informationen über die Nummer des Anrufers, Nummer des Angerufenen, die ID der

Funkzelle bei Gesprächsbeginn, Zeitpunkt des Gesprächsbeginns, Dauer des Gesprächs und vieles mehr.

Verbindungen von Prepaid-Teilnehmern werden hingegen schon während der laufenden Verbindung von einem Billing-Dienst abgerechnet, der sich auf einem IN-System und nicht in der MSC befindet. Mehr hierzu in Abschn. 6.11.

In modernen Sprachnetzwerken wurden, wie bereits zuvor beschrieben, alle leitungsvermittelnden Komponenten und Verbindungen gegen IP Komponenten ausgetauscht. Die Vermittlungsstelle wurde im Zuge dessen in einen MSC-Server (MSC-S) und ein Media Gateway (MGW) aufgeteilt. Dies wird in Abb. 6.11 gezeigt und ist in 3GPP TS 23.205 spezifiziert. Die MSC-Server Komponenten sind in dieser Architektur für Call-Control und Mobility Management zuständig, während die Media Gateways für die virtuellen Sprachverbindungen (Nutzdaten) zuständig sind.

Um Sprachverbindungen auf- und abzubauen, kommunizieren MSC-Server über die Mc Schnittstelle mit den Media Gateways. Diese Schnittstelle gibt es im klassischen Model nicht, da die MSC hier noch nicht in zwei Komponenten aufgeteilt war. Für die Verbindungssteuerung wird auf dieser IP basierten Schnittstelle laut 3GPP TS 29.232 das H.248/MEGACO (Media Gateway Control) Protokoll verwendet. Mit diesem Protokoll ist es den MSC-Servern z. B. möglich, Sprachkanäle zu zwei Teilnehmern herzustellen, diese im Media Gateway miteinander zu verbinden, Ansagen einzuspielen (z. B. „Der Teilnehmer ist zur Zeit nicht erreichbar“) und Konferenzgespräche auf- und abzubauen. Aus Redundanzgründen und zur Lastverteilung kann in diesem Model ein MSC-Server mit mehreren Media Gateways verbunden werden. Außerdem kann ein Media Gateway auch mit mehreren MSC-Servern kommunizieren. Fällt ein MSC-Server aus, kann ein MGW dadurch trotzdem in Betrieb bleiben. Außerdem entfällt der Bezug eines MSC-Servers zu einer bestimmten Region und ein MSC-Server kann nun Verbindungen im ganzen Netzwerk kontrollieren.

Das Radionetzwerk wird weiterhin über das A-Interface angebunden, heute typischerweise über eine IP-Verbindung. Eine Radionetzwerkkomponente kann dabei mit mehreren MGWs kommunizieren. Dies stellt sicher, dass beim Ausfall eines MGWs trotzdem weiterhin Gespräche in allen Teilen des Radionetzwerks vermittelt werden können.

Für die Weiterleitung von Sprachverbindungen innerhalb des Kernnetzes, z. B. zu anderen Mobilfunknetzwerken oder dem Festnetz wird das Nc Interface zur Signalisierung verwendet. Das Bearer Independent Call Control (BICC) Protokoll wird auf dieser Schnittstelle verwendet, das dem zuvor besprochenen klassischen ISUP Protokoll sehr ähnlich ist und in ITU Q.1901 und 3GPP TS 29.205 spezifiziert ist. Über ein Signaling Gateway (S-GW), links in Abb. 6.11 gezeigt, können mit einer Umwandlung in ISUP auch Gespräche in andere Kernnetzwerke aufgebaut werden, die noch eine klassische Leitungsvermittlung nutzen.

Virtuelle Sprachkanäle, die über das Nc Interface verhandelt wurden, werden zwischen den Media Gateways über das Nb Interface gleitet. Das Nb und Nc Interface zusammengenommen realisieren also das aus der klassischen Architektur bekannte

E-Interface. Je nach Radionetzwerktyp, Konfiguration des Netzwerkes und Fähigkeiten des Endgeräts werden die Sprachdaten in IP Paketen entweder als PCM/G.711, Narrowband-AMR oder Wideband-AMR übertragen. An den Grenzen des Kernnetzwerkes zu Netzwerken, die noch leitungsvermittelte Verbindungen verwenden, wird der Datenstrom vom Media Gateway dann wieder umgewandelt, z. B. von Narrowband-AMR über IP zu G.711/PCM über E-1. Dazu werden natürlich zusätzlich zu Ethernet Ports auch E-1 Ports benötigt.

Wie im klassischen Kernnetz werden auch bei BICN das C und D Interface verwendet, um mit dem HLR zu kommunizieren. Es fand jedoch auch hier ein Umstieg von E-1 basierten Verbindungen zu breitbandigen und auf IP basierten Schnittstellen statt.

6.6.2 Das Visitor Location Register (VLR)

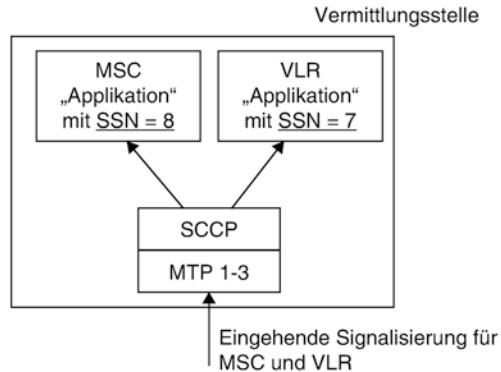
Jeder MSC ist eine Visitor Location Register (VLR)-Datenbank zugeordnet, die Informationen über alle aktuellen Teilnehmer in deren Versorgungsbereich verwaltet. Diese Daten sind jedoch nur eine temporäre Kopie der Originaldaten, die sich im Home Location Register (HLR) befinden, das im nächsten Abschnitt behandelt wird. Das VLR wird hauptsächlich verwendet, um die Signalisierung zwischen MSC und HLR zu reduzieren. Bewegt sich ein Teilnehmer in den Bereich einer MSC, werden die Daten einmalig aus dem HLR in das VLR kopiert und stehen somit lokal bei jeder Verbindungsaufnahme von oder zu Teilnehmern für eine Überprüfung zur Verfügung. Die Überprüfung der Teilnehmerdaten bei jedem Verbindungsaufbau ist notwendig, da jedem Teilnehmer individuell Dienste aktiviert oder gesperrt werden können. So ist es zum Beispiel möglich, ausgehende Anrufe eines Teilnehmers zu sperren oder Missbrauch zu unterbinden.

Während es die 3GPP Standards ermöglichen, das VLR als eine eigenständige Hardwarekomponente zu implementieren, haben alle Hersteller diese jedoch als Softwarekomponente in die MSC integriert. Dies ist möglich, da MSC und VLR über unterschiedliche SCCP-Subsystemnummern (vgl. Abschn. 6.4.1) angesprochen werden (Abb. 6.12).

Bewegt sich ein Teilnehmer aus dem Versorgungsbereich einer MSC, werden die Daten des Teilnehmers aus dem HLR in das VLR der neuen MSC kopiert und danach aus dem alten VLR gelöscht.

Für die Kommunikation mit dem HLR wurde in den GSM-Standards das D-Interface spezifiziert, das zusammen mit den Schnittstellen der MSC in Abb. 6.9 und 6.11 im Überblick dargestellt ist.

Abb. 6.12 Vermittlungsstelle mit integriertem VLR



6.6.3 Das Home Location Register (HLR)

Das Home Location Register (HLR) ist die Teilnehmerdatenbank eines GSM-Mobilfunknetzwerkes. Es enthält für jeden Teilnehmer Informationen, welche Dienste des Mobilfunknetzwerkes diesem zur Verfügung stehen.

Die International Mobile Subscriber Identity, kurz IMSI genannt, ist eine weltweit eindeutige Nummer, die einen Teilnehmer identifiziert und bei fast allen teilnehmerbezogenen Signalisierungsvorgängen im GSM-Netzwerk verwendet wird. Neben der SIM-Karte wird die IMSI auch im HLR gespeichert und ist dort der Schlüssel zu allen Informationen eines Teilnehmers. Weitere Details sind im Kapitel über LTE zu finden.

Aus Kundensicht ist der wichtigste im HLR gespeicherte Parameter die Telefonnummer eines Teilnehmers, die auch Mobile Subscriber ISDN Number (MSISDN) genannt wird. Diese darf maximal 15 Stellen lang sein und besteht aus:

- dem Country Code, also der internationalen Vorwahl des Landes, wie z. B. (+)49 für Deutschland
- dem National Destination Code (NDC), der nationalen Vorwahl des Netzbetreibers, normalerweise 3 Stellen lang
- einer eindeutigen Nummer innerhalb eines Mobilfunknetzwerks

Zwischen der IMSI und der MSISDN besteht ein 1:1 oder 1:N Zusammenhang, der im HLR festgelegt wird. Normalerweise bekommt ein Mobilfunkkunde nur eine Telefonnummer für seinen Mobilfunkanschluss. Da jedoch die IMSI und nicht die MSISDN einen Teilnehmer eindeutig identifiziert, ist es auch möglich, mehrere Telefonnummern pro Teilnehmer zu vergeben.

Ein weiterer Vorteil der IMSI als Schlüssel für alle Teilnehmerinformation ist, dass die Telefonnummer eines Teilnehmers jederzeit geändert werden kann, ohne dass die SIM-Karte getauscht werden muss. Hierfür muss lediglich im HLR eine neue MSISDN für den Benutzer eingetragen werden, die IMSI bleibt unverändert. Da auf der SIM-Karte

nur die IMSI, nicht jedoch die MSISDN gespeichert ist, sind hier keine Änderungen notwendig. Dies bedeutet auch, dass das Endgerät seine eigene Telefonnummer nicht kennt. Dies ist auch nicht notwendig, da diese bei einem abgehenden Telefonanruf von der MSC automatisch in die Nachrichten für den Verbindungsauftakt eingefügt wird, damit sie beim angerufenen Teilnehmer angezeigt werden kann.

Seit der Einführung der Mobile Number Portability (MNP) in Deutschland kann über die nationale Vorwahl (NDC) nicht mehr ermittelt werden, zu welchem Netzbetreiber ein Teilnehmer gehört. Dies hat zwar den großen Vorteil für den Kunden, seine Rufnummer bei einem Wechsel zu einem anderen Netzbetreiber mitnehmen zu können, verursacht aber einen Mehraufwand bei Signalisierung, Routing und Billing. Statt das Gespräch über den NDC (Vorwahl) zum richtigen Mobilfunknetzwerk weiterzuleiten, muss jetzt zuvor eine Mobile Number Portability-Datenbank befragt werden.

Neben der IMSI und MSISDN enthält das HLR für jeden Teilnehmer eine Menge weiterer Informationen über Dienste, die dieser verwenden darf. In der nachfolgenden Tabelle sind einige grundsätzliche Dienste (Basic Services) aufgeführt, die für einen Teilnehmer aktiviert werden können:

Basic Service	Aufgabe
Telefonie	Gibt an, ob ein Teilnehmer für die Sprachtelefonie freigeschaltet ist
Short Message Service (SMS)	Gibt an, ob ein Teilnehmer für den Kurznachrichtendienst SMS freigeschaltet ist
Datendienste	Gibt an, welche leitungsvermittelnden Datendienste (z. B. 2,4 kbit/s, 4,8 kbit/s, 9,6 kbit/s und 14,4 kbit/s) der Teilnehmer verwenden darf
FAX	Aktiviert oder sperrt FAX Übertragungen für einen Teilnehmer

Neben diesen grundsätzlichen Diensten bietet ein GSM-Netzwerk seinen Teilnehmern eine Menge weiterer Dienste an, die ebenfalls einzeln freigeschaltet oder gesperrt werden können. Da dies zusätzliche Dienste sind, werden diese auch Supplementary Services genannt:

Supplementary Service	Zweck
Call Forward Unconditional (CFU)	Erlaubt einem Benutzer das Setzen und Löschen einer sofortigen Gesprächsweiterleitung. Ist diese konfiguriert, wird der Ruf automatisch weitergeleitet, ohne dass das Telefon klingelt
Call Forward Busy (CFB)	Gibt dem Benutzer die Möglichkeit, ein Gespräch an eine andere Telefonnummer weiterzuleiten, wenn während eines laufenden Gesprächs ein weiterer Anruf eingeht

Supplementary Service	Zweck
Call Forward No Reply (CFNRY)	Leitet ein Gespräch weiter, wenn der Teilnehmer das Gespräch nach einer bestimmten Zeit nicht angenommen hat. Das Intervall kann vom Benutzer vorgegeben werden (z. B. 25 s)
Call Forward Not Reachable (CFNR)	Leitet ein Gespräch weiter, wenn das Mobiltelefon ausgeschaltet ist, oder keinen Netzempfang hat
Barring of All Outgoing Calls (BAOC)	Sperren aller abgehenden Anrufe. Kann auch vom Netzwerkbetreiber gesetzt werden, wenn der Teilnehmer seine Rechnung nicht bezahlt hat
Barring of All Incoming Calls (BAIC)	Ankommende Anrufe werden zum Teilnehmer nicht durchgestellt
Call Waiting (CW)	Das Anklopfen. Ermöglicht die Signalisierung eines weiteren ankommenden Gesprächs. Das erste Gespräch kann dann auf Halten (HOLD) gelegt werden, um das Zweite anzunehmen. Kann vom Netzbetreiber erlaubt oder gesperrt sein und vom Teilnehmer an- oder abgeschaltet werden
Call Hold (HOLD)	Zum Halten eines Gesprächs, um ein zweites eingehendes Gespräch anzunehmen oder um ein zweites Gespräch zu beginnen
Calling Line Identification Presentation (CLIP)	Anzeige der Rufnummer des Anrufers
Calling Line Identification Restriction (CLIR)	Mit CLIR kann ein Anrufer die Anzeige seiner Rufnummer beim Gesprächspartner unterdrücken
Connected Line Presentation (COLP)	Zeigt dem Anrufer, auf welche Telefonnummer sein Anruf umgeleitet wird, wenn eine Anruferleiterung aktiviert ist
Connected Line Presentation Restriction (COLR)	Unterdrückung des COLP Service
Multiparty (MPTY)	Erlaubt dem Teilnehmer, Konferenzen mit mehreren anderen Teilnehmern zu führen. Üblich sind Konferenzbrücken mit 3 oder 6 Teilnehmern

Die meisten Supplementary Services sind vom Netzwerkbetreiber an- und abschaltbar und ermöglichen ihm somit, für einzelne Dienste eine zusätzliche Gebühr zu verlangen. Während davon in den Anfangsjahren des Mobilfunks zum Teil gebrauch gemacht wurde, sind diese Dienste heute meistens ohne zusätzlichen Kosten nutzbar.

Die meisten dieser Dienste können vom Benutzer über das Mobiltelefon konfiguriert werden, wenn diese vom Netzbetreiber freigeschaltet sind. Meist bieten Endgeräte

dafür eine Menüstruktur an. Hinter diesen Menüs, die den Umgang mit diesen Diensten wesentlich vereinfachen, verbergen sich jedoch Zahlencodes, die mit einem „*“ Zeichen beginnen und zwischen Endgerät und Netzwerk ausgetauscht werden. Diese Codes sind im GSM-Standard 22.030 festgelegt und somit in allen Netzwerken und in allen Endgeräten gleich. Diese Codes kann ein Benutzer auch selber über die Tastatur eingeben. Nach Drücken der Ruftaste wird der eingegebene Zahlencode dann über die MSC zum HLR übertragen, wo der gewünschte Dienst aktiviert oder deaktiviert wird. Um zum Beispiel eine Anrufweiterleitung bei besetzt (CFB) auf die Nr. 0170992333 zu setzen, muss der Code **67*0170992333#+Ruftaste eingegeben werden.

6.6.4 Das Authentication Center (AC)

Ein weiterer wichtiger Bestandteil des HLR ist das Authentication Center. In ihm ist für jeden Teilnehmer ein geheimer Schlüssel Ki abgelegt, von dem nur eine weitere Kopie auf der SIM-Karte des Teilnehmers existiert. Dieser ist im Authentication Center und besonders auf SIM-Karte so gespeichert, dass er nicht ausgelesen werden kann.

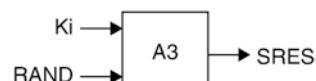
Bei vielen Vorgängen im Netzwerk, wie z. B. beim Beginn eines Gesprächs, wird der Teilnehmer mithilfe dieses Schlüssels authentifiziert. Abb. 6.13 und 6.14 zeigen diesen Vorgang.

Bei einer Verbindungsaufnahme zwischen Netzwerk und einem Teilnehmer fordert die MSC beim HLR/Authentication Center sogenannte Authentication Triplets an. Teil dieser Anforderung ist die IMSI des Teilnehmers. Das Authentication Center sucht anhand der IMSI den Ki des Teilnehmers und den zu verwendenden Authentifizierungsalgorithmus, der A3 genannt wird. Mit Ki wird dann das Authentication Triplet gebildet, das aus folgenden drei Werten besteht:

- RAND: Eine 128 Bit Zufallszahl.
- SRES: Die Signed Response SRES wird aus Ki und RAND mit dem Authentifizierungsalgorithmus A3 erzeugt und hat eine Länge von 32 Bit.
- Kc: Auch der Ciphering Key Kc wird aus Ki und RAND erzeugt. Er wird für die Verschlüsselung des Datenverkehrs nach erfolgreicher Authentifizierung verwendet. Mehr dazu in Abschn. 6.7.7.

RAND, SRES (und Kc) werden anschließend der MSC übergeben, die die eigentliche Authentifizierung des Teilnehmers vornimmt. Wichtig ist hierbei, dass der geheime Schlüssel Ki das Authentication Center nicht verlässt.

Abb. 6.13 Erzeugen der Signed Response (SRES)



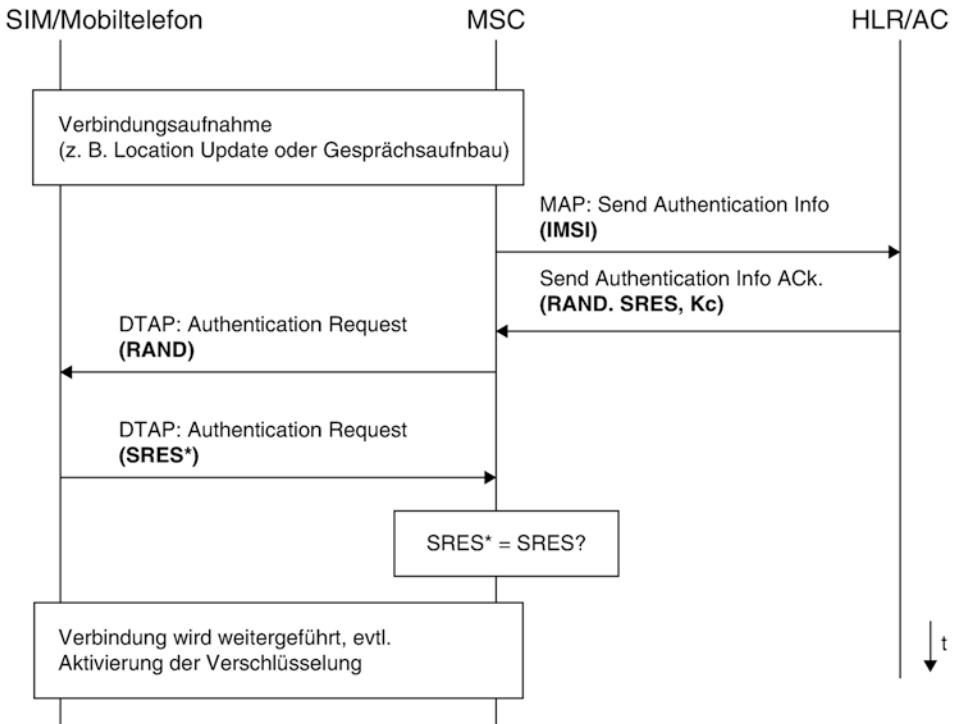


Abb. 6.14 Nachrichtenfluss während einer Authentifizierung

Um nachfolgende Verbindungsaunahmen zu beschleunigen, schickt das Authentication Center normalerweise gleich mehrere Authentication Triples in einer Nachricht zur MSC zurück. Diese werden dann in der MSC/VLR für die nächsten Verbindungsaunahmen zwischengespeichert.

Im nächsten Schritt sendet die MSC dem Endgerät die Zufallszahl (RAND) in einer Authentication Request-Nachricht. Das Endgerät übergibt die Zufallszahl der SIM-Karte, die dann mit der Kopie von Ki und dem Authentifizierungsalgorithmus A3 die Antwort, also die Signed Response (SRES*) berechnet. Diese wird dann dem Endgerät zurückgegeben und von diesem in einer Authentication Response-Nachricht zur MSC zurückgeschickt. Stimmen SRES und SRES* überein, ist der Teilnehmer erfolgreich authentifiziert und hat somit die Berechtigung, das Netzwerk zu verwenden.

Da der geheime Schlüssel Ki zu keiner Zeit im potenziell abhörgefährdeten Netzwerk oder per Funk übertragen wird, ist es einer dritten Person nicht möglich, SRES zu berechnen. Da bei der nächsten Authentifizierung eine neue Zufallszahl verwendet wird, ist auch das Abhören der zuvor gesendeten SRES nutzlos.

Abb. 6.15 zeigt Ausschnitte aus einer Authentication Request und einer Authentication Response-Nachricht. Neben den Formaten von RAND und SRES ist auch

Ausschnitt aus einer dekodierten Authentication Request Nachricht

```
SCCP MSG: Data Form 1
DEST. REF ID: 0B 02 00
DTAP MSG      LENGTH: 19
PROTOCOL DISC. : Mobility Management
DTAP MM MSG: Auth. Request
Ciph. Key Seq. : 0
RAND in hex: 12 27 33 49 11 00 98 45
87 49 12 51 22 89 18 81 (16 Byte = 128 Bit)
```

Ausschnitt aus einer dekodierten Authentication Response Nachricht

```
SCCP MSG: Data Form 1
DEST. REF ID: 00 25 FE
DTAP MSG      LENGTH: 6
PROTOCOL DISC. : Mobility Management
DTAP MM MSG: Auth. Response
SRES in hex: 37 21 77 61 (4 Byte = 32 Bit)
```

Abb. 6.15 Authentifizierung zwischen Netzwerk und Endgerät

sehr interessant, welche Protokolle des SS-7-Stacks zum Einsatz kommen (vgl. hierzu auch Abschn. 6.4.2).

6.6.5 Das Short Message Service Center (SMSC)

Ein weiteres wichtiges Netzwerkelement ist das Short Message Service Center (SMSC), das für die Weiterleitung und Speicherung von Kurznachrichten (SMS) zuständig ist. Erst etwa 4 Jahre nach dem Start der ersten GSM-Netze wurde dieser Dienst in Betrieb genommen. Binnen kurzer Zeit jedoch erfreute er sich so enormer Popularität, dass Netzbetreiber zeitweilig einen zweistelligen Prozentsatz ihres Umsatzes mit diesem Dienst erwirtschaften. Zwar ist die Bedeutung und der Umsatz des SMS Dienstes durch das Aufkommen von Internet basierten Alternativen heute wesentlich geringer als früher, der Dienst bleibt aber weiterhin lukrativ für die Netzbetreiber und eine wichtige Kommunikationsform.

Der SMS Dienst ermöglicht sowohl den direkten Austausch von Kurznachrichten zwischen Teilnehmern, als auch automatisch generierte SMS-Nachrichten, z. B. als Reaktion auf weitergeleitete Gespräche zur Sprachbox (Voice Mail System). Das Prinzip der Übertragung einer SMS ist jedoch in beiden Fällen identisch:

Der Sender erstellt eine SMS und überträgt diese zur MSC über einen Signalisierungskanal. Eine SMS ist somit nichts anderes als eine DTAP SS-7-Nachricht,

wie z. B. eine Location Update-Nachricht oder eine Setup-Nachricht zum Aufbau eines Gesprächs. Inhalt der SMS ist der Nachrichtentext selber, sowie die Telefonnummer (MSISDN) des Zielteilnehmers. Die MSC leitet die SMS ohne weitere Bearbeitung direkt an das Short Message Service Center (SMSC) weiter. Das SMSC bestätigt dem Sender daraufhin den korrekten Empfang der SMS. Dies wird dann auch auf dem Display des Teilnehmers angezeigt (Abb. 6.16).

Für die Zustellung einer SMS analysiert das SMSC die MSISDN des Empfängers und befragt das entsprechende HLR nach dessen aktuellem Aufenthaltsort (MSC). Danach wird die SMS an diese MSC geschickt. Ist der Teilnehmer in dieser MSC als aktiv angemeldet (attached), versucht die MSC Kontakt mit ihm aufzunehmen und die SMS zuzustellen. Die korrekte Zustellung wird dem SMSC quittiert, und die SMS kann daraufhin im SMSC gelöscht werden.

Ist der Teilnehmer nicht erreichbar (z. B. Akku leer, keine Netzbdeckung, Endgerät ausgeschaltet, etc.) kann die SMS nicht sofort zugestellt werden. Daraufhin wird im VLR-Eintrag des Empfängers das Message Waiting Flag gesetzt, und die SMS wird im SMSC zwischengespeichert. Sobald sich der Empfänger wieder meldet, sieht die MSC dieses Flag und kann das SMSC davon unterrichten. Daraufhin versucht das SMSC erneut, die SMS zuzustellen.

Da auch im HLR ein Message Waiting Flag gesetzt wird, erreicht die SMS einen Empfänger auch dann noch, wenn dieser sein Mobiltelefon z. B. in Frankfurt ausgeschaltet hat und sich während der SMS-Zustellung gerade im Flugzeug nach Paris befindet. Beim Einschalten des Mobiltelefons in Paris meldet die dortige MSC dem Heimat-HLR des Teilnehmers dessen neue Position (Location Update). Das HLR schickt daraufhin dem neuen MSC/VLR eine Kopie der Teilnehmerdaten inklusive des Message Waiting Flags, und die SMS kann wiederum korrekt zugestellt werden.

Üblicherweise wird dem Nutzer eines Endgeräts nur der erfolgreiche Versand einer SMS Nachricht bis zum SMSC signalisiert, nicht jedoch, ob und wann die Nachricht auch an den Empfänger ausgeliefert wurde. Falls vom Endgerät unterstützt, ist es jedoch

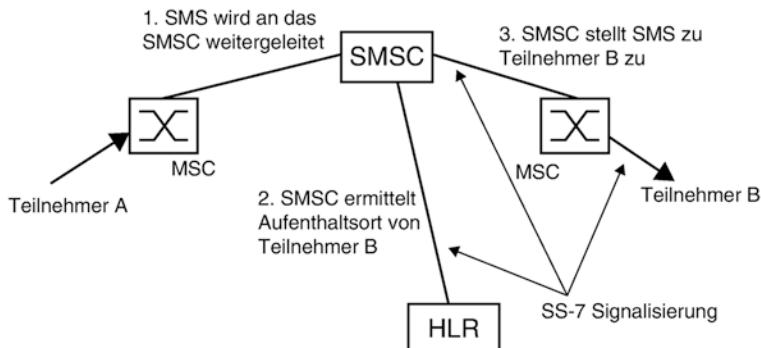


Abb. 6.16 Zustellungsprinzip einer SMS

auch möglich, für SMS Nachrichten eine Ende-zu-Ende Empfangsbestätigung vom Netzwerk anzufordern. In der Praxis gibt es unterschiedliche Ansätze, dies in Endgeräten umzusetzen. In den SMS Einstellungen mancher Betriebssysteme kann eine Bestätigung nach dem Versand an den Empfänger für alle ausgehenden SMS Nachrichten aktiviert werden. Diese wird dann z. B. als Häkchen neben einer Nachricht angezeigt oder in einer separaten Liste.

6.7 Das Base Station Subsystem (BSS) und Sprachcodierung

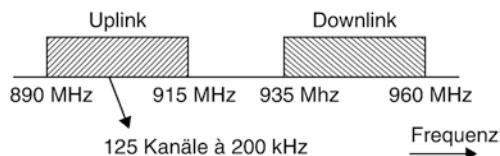
Während ein Großteil der zusätzlichen Funktionalität für den Mobilfunk im NSS in den 1990er Jahren durch neue Software implementiert wurde, musste im Radionetzwerk ein Großteil der Hard- und Software neu entwickelt werden. Dies wurde schon alleine deswegen nötig, da alle Vorgängertechnologien noch auf analoger Technik für die Funkübertragung basierten.

6.7.1 Frequenzbereiche

In Europa wurde GSM zunächst im 900 MHz Frequenzband von 890–915 MHz im Uplink und von 935–960 MHz im Downlink spezifiziert. Uplink ist dabei die Senderichtung von Mobiltelefon zu Netzwerk, Downlink die Senderichtung von Netzwerk zu Mobiltelefon. Die Bandbreite von 25 MHz ist dabei in 125 Kanäle mit einer Bandbreite von jeweils 200 kHz aufgeteilt. Diese Kanäle teilten sich in Deutschland ursprünglich die Mobilfunkbetreiber Deutsche Telekom/T-Mobile (vormals D1) und Vodafone (vormals D2) (Abb. 6.17).

Schon bald war abzusehen, dass diese Kanalanzahl für den schnell wachsenden Mobilfunkverkehr in vielen europäischen Ländern nicht ausreichend sein würde. Deshalb wurde in einem zweiten Schritt ein Frequenzband im Frequenzbereich von 1710–1785 MHz im Uplink und 1805–1880 im Downlink für GSM in Europa geöffnet. Statt einer Bandbreite von 25 MHz wie im 900 MHz-Bereich steht hier eine Bandbreite von 75 MHz zur Verfügung. Dies entspricht 375 zusätzlichen Kanälen. Ein Teil dieser Kanäle wurde in Deutschland zunächst von E-Plus (heute Teil von Telefonica O₂) verwendet, ein weiterer Teil dann von Viag Interkom (heute Telefonica O₂). Da dieser Frequenzbereich eine geringere Reichweite als das 900 MHz Band zulässt, entschloss sich die

Abb. 6.17 Uplink und Downlink im 900 MHz Frequenzband



Regulierungsbehörde später, auch O₂ und E-Plus Spektrum im 900 MHz Band zuzu-teilen und im Gegenzug T-Mobile und Vodafone zusätzlich Spektrum im 1800 MHz Band zugänglich zu machen. Mit dem Auslaufen der Frequenznutzungsrechte für das ursprüngliche GSM 900 MHz Band im Jahre 2016 und durch die anhaltende Weiter-entwicklung des deutschen Mobilfunkmarktes sind heute drei Netzbetreiber in Deutsch-land mit GSM im 900 MHz Band vertreten und haben sich aus dem 1800 MHz Band zurückgezogen, da dies nun hauptsächlich für LTE verwendet wird. Die Funktionsweise von GSM ist auf beiden Frequenzbändern identisch, sie unterscheiden sich lediglich durch andere Kanalnummern, die Absolute Radio Frequency Channel Number (ARFCN) genannt werden.

Während in Nordamerika zunächst die alten analogen Mobilfunknetze weiter betrieben wurden, etablierte sich GSM neben anderen digitalen Techniken auch dort. Da sowohl das 900 MHz, als auch das 1800 MHz Band schon von anderen Funkdiensten genutzt wurden, musste man hier auf Frequenzen im 1900 MHz Band ausweichen. Dies hatte den gravierenden Nachteil, dass zunächst viele Mobiltelefone aus den USA und Kanada in Europa nicht funktionierten und umgekehrt. Da auch im 1900 MHz Band die Frequenzen knapp wurden, wurde später ein weiteres Band im 850 MHz-Bereich für den nordamerikanischen Markt geöffnet. Auch dieses ist zum 900 MHz Band, das in den meisten anderen Ländern verwendet wird, inkompatibel. Um weltweit in GSM-Netzen erreichbar zu sein, kamen nach einigen Jahren dann Quad-Band Mobiltelefone auf den Markt, die alle GSM Bänder unterstützten.

Name	ARFCN	Uplink (MHz)	Downlink (MHz)
GSM 900 (Primary)	0–124	890–915	935–960
GSM 900 (Extended)	975–1023, 0–124	880–915	925–960
GSM 1800	512–885	1710–1785	1805–1880
GSM 1900 (Nordamerika)	512–810	1850–1910	1930–1990
GSM 850 (Nordamerika)	128–251	824–849	869–894
GSM-R	0–124, 955–1023	876–915	921–960

Neben öffentlichen GSM-Netzen etablierte sich für die europäischen Eisenbahnen in den 2000er Jahren eine digitale Zugfunkgeneration, die auf dem GSM-Standard basiert und auch heute noch verwendet wird. Zusätzlich zu den GSM-Funktionalitäten wurden spezielle für Eisenbahnen benötigte Dienste wie z. B. Gruppenrufe entwickelt. Dieser Standard wurde GSM for Railways, kurz GSM-R genannt. Da es sich hier nicht um öffentliche, sondern um private Netzwerke handelt, wurde den GSM-R Netzen auch ein eigenes Frequenzband unmittelbar unterhalb des öffentlichen 900 MHz GSM-Bands zugeteilt. Um GSM-R zu nutzen, sind Mobiltelefone mit leichten Hardware-modifikationen notwendig, um in diesem Frequenzbereich senden und empfangen zu können. Um eisenbahnspezifische Dienste wie z. B. Gruppenrufe verwenden zu können, wurde zusätzlich die Mobiltelefonsoftware erweitert. In Deutschland sind heute alle

wesentlichen Bahnstrecken mit GSM-R ausgerüstet. Mehr zum Thema GSM-R ist unter <http://www.uic.org/gsm-r> zu finden.

6.7.2 Base Transceiver Station (BTS)

Basisstationen, auch Base Transceiver Station (BTS) genannt, sind durch ihre Antennen die wohl sichtbarsten Netzwerkelemente eines GSM-Mobilfunksystems. Diese ersetzen im Vergleich zum Festnetz die kabelgebundene Verbindung mit dem Benutzer durch eine Funkverbindung, die auch Luftschnittstelle oder Air Interface genannt wird. Laut Presseberichten hat jeder Netzwerkbetreiber in Deutschland einige zehntausend dieser Basisstationen (Abb. 6.18).

Theoretisch kann eine BTS eine Fläche mit einem Radius von bis zu 35 km abdecken. Dieses Gebiet wird auch Zelle genannt. Da eine BTS aber nur mit einer begrenzten Anzahl an Nutzern gleichzeitig kommunizieren kann, sind Zellen vor allem in städtischen Bereichen wesentlich kleiner. Sie reichen dort von 2–3 km Radius in Wohngebieten bis zu wenigen 100 m in Innenstädten. Aber auch auf dem Land sind Zellen mit einem Radius von mehr als 15 km nur sehr selten anzutreffen. Hier ist die maximale Sendeleistung der Endgeräte von 1–2 W der begrenzende Faktor.

Grundsätzlich gilt, dass die von einer Basisstation verwendeten Sende Frequenzen nicht von Nachbarstationen verwendet werden dürfen, da diese sich sonst gegenseitig stören (Interferenz). Da eine Basisstation wie in Abb. 6.19 normalerweise mehrere Nachbarstationen besitzt, können nur eine sehr begrenzte Anzahl an Frequenzen pro Basisstation verwendet werden.

Um die Kapazität einer BTS zu steigern, wird das abgedeckte Gebiet oft in zwei oder drei Sektoren eingeteilt, die jeweils von einer eigenen Sende- und Empfangshardware der BTS auf unterschiedlichen Frequenzen abgedeckt werden. Somit können die Frequenzen im zweidimensionalen Raum gesehen öfters wieder verwendet werden. Jeder Sektor ist dabei eine eigenständige Zelle (Abb. 6.20).

Abb. 6.18 Eine typische Antenne einer GSM-Basisstation. Die zusätzliche optionale Richtfunkantenne (runde Antenne unten) verbindet die Basisstation mit dem GSM-Netzwerk



Abb. 6.19 Zelle mit Nachbarzellen

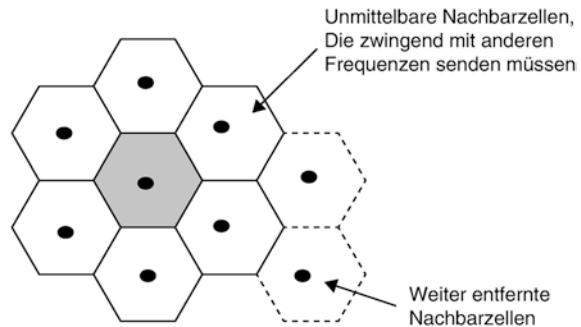


Abb. 6.20 Sektorisierte Zellkonfigurationen

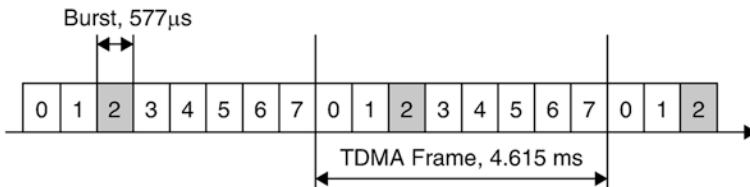
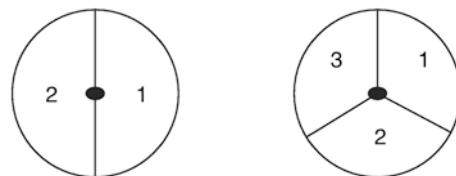


Abb. 6.21 Ein GSM-TDMA-Frame

6.7.3 Die GSM-Luftschnittstelle

Der Übertragungsweg zwischen BTS und Mobilfunkteilnehmer wird bei GSM als Luftschnittstelle, Air Interface oder Um-Interface bezeichnet.

Damit eine BTS mit mehreren Teilnehmern gleichzeitig kommunizieren kann, werden bei GSM zwei Verfahren angewandt. Das erste Verfahren ist der Frequenzmultiplex (Frequency Division Multiple Access, FDMA), also die gleichzeitige Nutzung mehrerer Frequenzen pro Zelle.

Das zweite Verfahren ist der Zeitmultiplex, auch Time Division Multiple Access (TDMA) genannt. Bei GSM können pro Trägerfrequenz mit 200 kHz Bandbreite bis zu 8 Teilnehmer gleichzeitig kommunizieren (Abb. 6.21).

Dazu werden auf dem Träger 4,615 ms lange Frames übertragen. Jeder Frame enthält 8 voneinander unabhängige physische Zeitschlitzte (Timeslots) für die Kommunikation mit unterschiedlichen Teilnehmern. Das Zeitintervall eines Timeslots wird Burst genannt und beträgt 577 μs. Bekommt ein Endgerät beispielsweise Timeslot Nr. 2 eines Frames

für ein Telefongespräch zugeteilt, darf es in jedem Frame in diesem Timeslot senden und empfangen. Danach muss es den restlichen Frame abwarten, bevor es erneut an der Reihe ist.

Nachdem die grundsätzlichen Mehrfachzugriffsverfahren nun bekannt sind, kann in grober Nähnung die Gesamtkapazität einer BTS ermittelt werden. Für nachfolgendes Beispiel wird eine BTS mit 3 sektorisierten Zellen betrachtet, die jeweils über 3 Frequenzen verfügen. Pro Sektor stehen somit $3 \times 8 = 24$ Timeslots zur Verfügung. Von diesen müssen 2 Timeslots für Signalisierungsaufgaben abgezogen werden. Somit bleiben 22 Timeslots pro Sektor. Von diesen werden meist 4 oder mehr Timeslots für den paketorientierten Datendienst GPRS verwendet, der im nächsten Kapitel beschrieben wird. Somit bleiben pro Sektor 18, pro BTS somit 54 Kanäle für die Sprachübertragung. Das bedeutet also, dass in der Praxis 54 Teilnehmer gleichzeitig pro BTS kommunizieren können.

Eine BTS versorgt jedoch wesentlich mehr Teilnehmer eines Netzwerkes, da nicht alle Teilnehmer gleichzeitig telefonieren. Mobilfunknetzbetreiber gehen davon aus, dass im Durchschnitt ein Teilnehmer pro Stunde 3 min telefoniert. Somit versorgt eine BTS in grober Nähnung etwa 20 mal mehr passive als aktive Teilnehmer. In diesem Beispiel versorgt die BTS also etwa 1080 Teilnehmer.

Teilt man die gesamte Nutzerzahl eines Netzwerkes, im Falle von O₂ in Deutschland 2014 etwa 20 Mio., durch diesen Wert, so kommt man auf etwa 20.000 Basisstationen, die für diese Anzahl Teilnehmer im gesamten Bundesgebiet benötigt werden. Diese Zahl ist im Bereich der von den Netzwerkbetreibern veröffentlichten Werte und vermittelt einen ersten Eindruck über die Dimensionen eines großen Netzwerkes. Da GSM Kanäle jedoch auch für die GPRS Datenübertragung verwendet wird und heute Sprachtelefonie auch im UMTS und LTE Netz abgewickelt wird, ist diese Rechnung jedoch nur eine sehr grobe Näherung.

Jeder Burst eines TDMA Frames ist wie in Abb. 6.22 gezeigt in unterschiedliche Bereiche aufgeteilt:

Während einer durch die Guard Time festgelegten Zeit am Ende jedes Bursts werden keine Daten übertragen. Dies ist notwendig, da sich Teilnehmer auch während der Dauer einer Verbindung bewegen und sich der Abstand zur BTS ständig ändern kann. Da sich

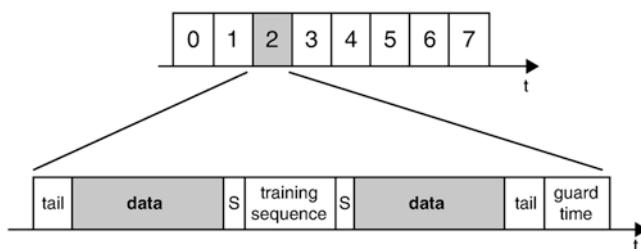


Abb. 6.22 Ein GSM-Burst nach 3GPP TS 45.001

die Funkwellen ‚nur‘ mit Lichtgeschwindigkeit ausbreiten, treffen die Daten eines weiter entfernten Teilnehmers erst später als die Daten eines Teilnehmers ein, der sich näher an der Basisstation befindet. Um Überlappungen zu vermeiden, sind diese Pausenzeiten nötig. Die Guard Time ist jedoch sehr kurz, da durch eine aktive Sendezeitregelung, die Timing Advance genannt wird, diese Unterschiede weitestgehend ausgeglichen werden. Mehr zum Timing Advance im Laufe dieses Kapitels.

In der Mitte des Bursts befindet sich die Training Sequence mit einem immer gleichen Bitmuster. Diese ist notwendig, da sich das Signal bei der Funkübertragung durch verschiedene Phänomene wie Reflexion, Absorption und Mehrfachausbreitung verändert. Diese Effekte müssen auf der Empfängerseite wieder ausgeglichen werden. Der Empfänger vergleicht dazu das ihm bekannte Bitmuster mit dem empfangenen Signal und kann daraus schließen, wie aus dem empfangenen Signal die Originaldaten wieder rekonstruiert werden können.

Am Anfang und Ende des Bursts wird ein bekanntes Bitmuster gesendet, damit der Empfänger den Beginn und das Ende des Bursts korrekt erkennen kann. Diese Felder werden Tail genannt.

Die eigentlichen Nutzdaten des Bursts, also z. B. digitalisierte Sprache, werden in zwei Nutzdatenfelder (data) mit jeweils 57 Bit Länge übertragen. Somit werden pro 577 µs Burst genau 114 Bit Nutzdaten übertragen.

Schließlich gibt es vor und nach der Training Sequence noch jeweils 2 Bits, die Stealing Flags genannt werden. Sind sie gesetzt, befinden sich in den Datenfeldern keine Nutzdaten, sondern dringende Signalisierungsinformationen. Werden Signalisierungsdaten in diesen Feldern übertragen, gehen die Nutzdaten verloren.

Zur Übertragung von Nutzdaten oder Signalisierungsdaten werden die Zeitschlitzte in logische Kanäle eingeteilt. Ein Nutzdatenkanal für die Übertragung von Sprachdaten ist z. B. ein logischer Kanal. Auf der ersten Trägerfrequenz einer Zelle werden die ersten beiden Timeslots üblicherweise für allgemeine logische Signalisierungskanäle reserviert, die restlichen können für 6 unabhängige Nutzkanäle oder GPRS verwendet werden. Da es wesentlich mehr logische Signalisierungskanäle als physische Kanäle (Timeslots) für die Signalisierung gibt, wurden im 3GPP Standard TS 45.002 für die Signalisierung 51 Frames zu einem Multiframe zusammengefasst. In einem solchen Multiframe, der sich ständig wiederholt, ist genau festgelegt, in welchen Bursts von Timeslot 0 und 1 welche logischen Kanäle übertragen werden. Über diese Vorschrift werden also viele logische Kanäle auf wenige physische Kanäle übertragen. Für Timeslots, die für Nutzdatenübertragung (also z. B. Sprache) verwendet werden, wird ein 26 Multiframe-Muster verwendet.

Um dies grafisch darzustellen, werden alle Bursts eines Timeslots untereinander angeordnet, die 8 Timeslots eines Frames nebeneinander. Abb. 6.23 zeigt dieses Prinzip, mit dem dann in Abb. 6.24 die Zuordnung der logischen Kanäle zu physischen Kanälen dargestellt ist.

Logische Kanäle werden in zwei Gruppen eingeteilt. Sind Daten auf einem logischen Kanal nur für einen einzelnen Nutzer bestimmt, handelt es sich um einen Dedicated

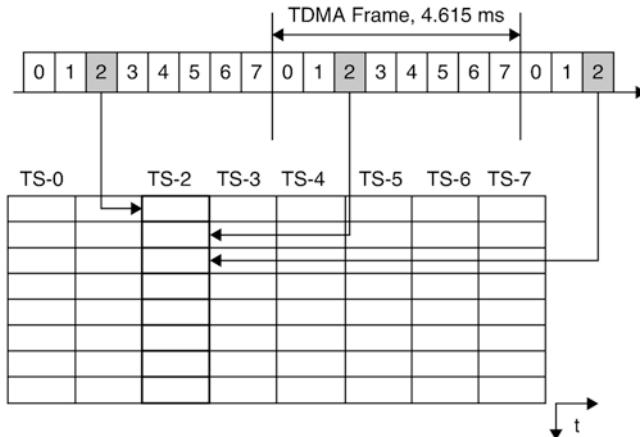


Abb. 6.23 Zusammenhängende Anordnung von Bursts eines Timeslots für die Darstellung der logischen Kanäle in Abb. 6.24

Channel. Werden auf einem Kanal Daten für mehrere Benutzer übertragen, wird dieser Common Channel genannt.

Im Anschluss werden nun zuerst die Dedicated Channels betrachtet:

Der Traffic Channel (TCH) ist ein Nutzdatenkanal in GSM. Über diesen können entweder digitalisierte Sprachdaten oder leitungsvermittelnde Datendienste mit bis zu 14,4 kbit/s oder 9,6 kbit/s FAX übertragen werden.

Der Fast Associated Control Channel (FACCH) wird auf dem gleichen Timeslot wie der TCH übertragen. Er dient zur Übermittlung dringender Signalisierungsnachrichten wie z. B. einem Handover-Kommando. Da dringende Signalisierungsnachrichten nur selten zu übertragen sind, wurden dem FACCH keine eigenen Bursts zugeteilt. Bei Bedarf werden Nutzdaten aus einzelnen Bursts des Timeslots entfernt und FACCH Daten übertragen. Um dies dem Endgerät bzw. dem Netzwerk zu signalisieren, werden die in Abb. 6.22 gezeigten Stealing Flags eines Bursts entsprechend gesetzt. Aus diesem Grund ist der FACCH in Abb. 6.24 auch nicht dargestellt.

Der Slow Associated Control Channel (SACCH) ist ebenfalls einem aktiven Benutzer zugeordnet. Dieser wird im Uplink verwendet, um während der aktiven Verbindung ständig Messergebnisse der Signalpegelmessungen der aktiven Zelle, sowie der Nachbarzellen an das Netzwerk zu senden. Die Messergebnisse werden vom Netzwerk dann für die Handover-Entscheidung sowie für die Leistungsregelung verwendet. Im Downlink werden auf dem SACCH im Gegenzug Befehle für die Leistungsregelung der Mobilstation übermittelt. Außerdem erhält das Endgerät über den SACCH Informationen für die Timing Advance-Regelung, die in Abschn. 6.7.4 und Abb. 6.29 näher beschrieben werden. Da diese Daten keine hohe Priorität haben und die Datenrate sehr gering ist,

FN	TS-0	TS-1		FN	TS-2	...	TS-7
0	FCCH	SDCCH/0		0	TCH		TCH
1	SCH	SDCCH/0		1	TCH		TCH
2	BCCH	SDCCH/0		2	TCH		TCH
3	BCCH	SDCCH/0		3	TCH		TCH
4	BCCH	SDCCH/1		4	TCH		TCH
5	BCCH	SDCCH/1		5	TCH		TCH
6	AGCH/PCH	SDCCH/1		6	TCH		TCH
7	AGCH/PCH	SDCCH/1		7	TCH		TCH
8	AGCH/PCH	SDCCH/2		8	TCH		TCH
9	AGCH/PCH	SDCCH/2		9	TCH		TCH
10	FCCH	SDCCH/2		10	TCH		TCH
11	SCH	SDCCH/2		11	TCH		TCH
12	AGCH/PCH	SDCCH/3		12	SACCH		SACCH
13	AGCH/PCH	SDCCH/3		13	TCH		TCH
14	AGCH/PCH	SDCCH/3		14	TCH		TCH
15	AGCH/PCH	SDCCH/3		15	TCH		TCH
16	AGCH/PCH	SDCCH/4		16	TCH		TCH
17	AGCH/PCH	SDCCH/4		17	TCH		TCH
18	AGCH/PCH	SDCCH/4		18	TCH		TCH
19	AGCH/PCH	SDCCH/4		19	TCH		TCH
20	FCCH	SDCCH/5		20	TCH		TCH
21	SCH	SDCCH/5		21	TCH		TCH
22	SDCCH/0	SDCCH/5		22	TCH		TCH
23	SDCCH/0	SDCCH/5		23	TCH		TCH
24	SDCCH/0	SDCCH/6		24	TCH		TCH
25	SDCCH/0	SDCCH/6		25	free		free
26	SDCCH/1	SDCCH/6		0	TCH		TCH
27	SDCCH/1	SDCCH/6		1	TCH		TCH
28	SDCCH/1	SDCCH/7		2	TCH		TCH
29	SDCCH/1	SDCCH/7		3	TCH		TCH
30	FCCH	SDCCH/7		4	TCH		TCH
31	SCH	SDCCH/7		5	TCH		TCH
32	SDCCH/2	SACCH/0		6	TCH		TCH
33	SDCCH/2	SACCH/0		7	TCH		TCH
34	SDCCH/2	SACCH/0		8	TCH		TCH
35	SDCCH/2	SACCH/0		9	TCH		TCH
36	SDCCH/3	SACCH/1		10	TCH		TCH
37	SDCCH/3	SACCH/1		11	TCH		TCH
38	SDCCH/3	SACCH/1		12	SACCH		SACCH
39	SDCCH/3	SACCH/1		13	TCH		TCH
40	FCCH	SACCH/2		14	TCH		TCH
41	SCH	SACCH/2		15	TCH		TCH
42	SACCH/0	SACCH/2		16	TCH		TCH
43	SACCH/0	SACCH/2		17	TCH		TCH
44	SACCH/0	SACCH/3		18	TCH		TCH
45	SACCH/0	SACCH/3		19	TCH		TCH
46	SACCH/1	SACCH/3		20	TCH		TCH
47	SACCH/1	SACCH/3		21	TCH		TCH
48	SACCH/1	free		22	TCH		TCH
49	SACCH/1	free		23	TCH		TCH
50	free	free		24	TCH		TCH

Abb. 6.24 Nutzung der Timeslots im Downlink, in Anlehnung an 3GPP TS 45.002

werden nur wenige Bursts für diesen logischen Kanal in einem 26 Multiframe verwendet.

Der Standalone Dedicated Control Channel (SDCCH) ist ein reiner Signalisierungs-kanal, der während des Gesprächsaufbaus verwendet wird, solange einem Teilnehmer noch kein eigener TCH zugeordnet ist. Außerdem wird dieser Kanal für Signalisierungs-daten verwendet, die nicht zum Aufbau eines Gesprächs und somit auch zu keiner Zuteilung eines TCH führen. Dies sind z. B. ein Location Update oder das Senden oder Empfangen einer SMS.

Neben diesen teilnehmerbezogenen Kanälen gibt es eine Reihe von Common Channels, die von allen Teilnehmern abgehört werden:

Der Synchronization Channel (SCH) wird von Endgeräten bei der Netzwerk- und Zellsuche verwendet.

Der Frequency Correction Channel (FCCH) wird von Endgeräten für die Kalibrierung ihrer Sende- und Empfangseinheiten verwendet und dient außerdem dazu, den Anfang eines 51-Multiframes zu finden.

Der Broadcast Common Control Channel (BCCH) überträgt in verschiedenen SYS_INFO-Nachrichten eine Vielzahl von Systeminformationen, über die alle Teilnehmer, die am Netzwerk angemeldet, aber nicht aktiv sind (Idle Mode), stets informiert sein müssen. Dazu gehören unter anderem:

- Mobile Country Code (MCC) und Mobile Network Code (MNC) der Zelle.
- Identifikation der Zelle bestehend aus dem Location Area Code (LAC) und der Cell ID.
- Um Endgeräten die Suche nach Nachbarzellen zu vereinfachen, werden auf dem BCCH jeder Zelle die verwendeten Frequenzen der Nachbarzellen ausgestrahlt. Somit muss das Mobiltelefon nicht ständig das komplette Frequenzband nach Nachbarzellen durchsuchen.

Der Paging Channel (PCH) wird verwendet, um nicht aktive Teilnehmer bei eingehenden Anrufen oder SMS-Nachrichten zu rufen (pagen). Da das Netzwerk nur weiß, in welcher Location Area sich ein Teilnehmer befindet, wird dieser auf dem Paging Channel jeder Zelle in dieser Location Area gerufen. Wichtigster Teil der Nachricht ist seine IMSI oder eine temporäre ID, die Temporary Mobile Subscriber Identity (TMSI) genannt wird. Diese wird z. B. nach dem Einschalten einem Teilnehmer zugewiesen und kann vom Netzwerk dann bei beliebigen Netzwerkuzugriffen nach Aktivieren der Datenver-schlüsselung wieder geändert werden, um die Anonymität der Teilnehmer im Netzwerk zu gewährleisten. Dies vereitelt externen Beobachtern, Bewegungsprofile von Teilnehmern zu erstellen.

Der Random Access Channel (RACH) ist der einzige Common Channel vom End-gerät in Richtung Netzwerk. Erhält das Endgerät über den PCH eine Nachricht, dass das Netz mit ihm Kontakt aufnehmen will, oder möchte der Benutzer ein Gespräch beginnen, eine SMS senden, usw., nimmt das Endgerät über den RACH mit dem Netz-

werk Kontakt auf. Dies geschieht mit einer Channel Request-Nachricht. Diese muss über den „Zufallskanal“ gesendet werden, da die Teilnehmer einer Zelle nicht untereinander synchronisiert sind. Somit ist nicht gewährleistet, dass nicht zwei Endgeräte versuchen, zur selben Zeit auf das Netzwerk zuzugreifen. Erst wenn auf die Channel Request-Anfrage ein dedizierter Kanal (SDCCH) vom Netzwerk zugeteilt worden ist, können keine Kollisionen mehr auftreten. Tritt eine Kollision beim Zugriff auf den RACH auf, gehen die kollidierenden Nachrichten verloren, und die Teilnehmer erhalten vom Netzwerk keine Antwort. Nach unterschiedlich langen Wartezeiten müssen sie danach ihre Kanalanforderung wiederholen.

Sendet ein Teilnehmer auf dem RACH eine Channel Request-Nachricht, reserviert das Netzwerk daraufhin einen SDCCH oder in Ausnahmefällen direkt einen TCH und benachrichtigt den Teilnehmer daraufhin auf dem Access Grant Channel (AGCH) mit einer Immediate Assignment-Nachricht. Diese Nachricht enthält dann die Information, welchen SDCCH oder TCH der Teilnehmer verwenden darf.

Abb. 6.25 zeigt das Zusammenspiel von PCH, AGCH und SDCCH beim Aufbau einer Signalisierungsverbindung. Der in der Abbildung gezeigte Base Station Controller (BSC) ist für die Vergabe aller SDCCH und TCH-Kanäle einer BTS zuständig und wird im Abschn. 6.7.4 näher beschrieben.

Wie in Abb. 6.24 auch zu sehen ist, werden nicht alle Bursts von Timeslot 2 bis 7 für Traffic Channels (TCH) verwendet. In jedem Timeslot wird jeweils der 12. Burst für den zum TCH zugehörigen Slow Associated Control Channel (SACCH) verwendet. Außerdem werden im 25. Burst keine Daten übertragen. Diese Lücke wurde geschaffen, um dem Endgerät die Möglichkeit zu geben, auch während einer aktiven Verbindung Messungen der Signalstärken der Nachbarzellen auf anderen Frequenzen durchzuführen. Dies ist nötig, damit das Netzwerk die Verbindung eines aktiven Teilnehmers ggf. in

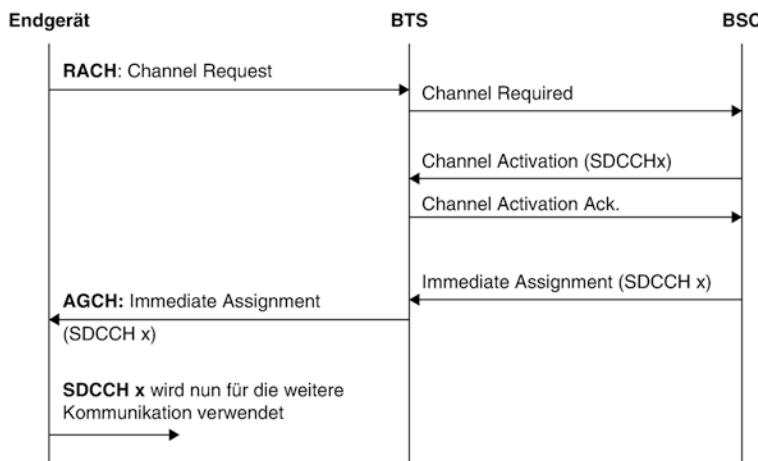


Abb. 6.25 Aufbau einer Signalisierungsverbindung

eine andere Zelle umschalten kann (Handover), falls dort die Übertragungsbedingungen besser als die der aktuellen Zelle werden.

Der GSM-Standard bietet zwei Möglichkeiten der Frequenznutzung. Der einfachste Fall, von dem hier bisher ausgegangen wurde, ist die Verwendung einer konstanten Trägerfrequenz (ARFCN). Um die Übertragungsqualität zu steigern, wurde auch ein Verfahren zum Wechsel der Frequenzen während einer Verbindung, im englischen Frequency Hopping genannt, standardisiert. Wird Frequency Hopping in einer Zelle angewandt, wird nach der Übertragung jedes Bursts die Trägerfrequenz (carrier frequency) gewechselt. Auf diese Weise kann die Wahrscheinlichkeit erhöht werden, nur wenige Daten zu verlieren, wenn in einem Frequenzbereich eine Störung das Nutzdatensignal überlagert. Im schlimmsten Fall ist davon nur ein Burst betroffen, da der nächste Burst eines Teilnehmers schon wieder auf einer anderen Frequenz übertragen wird. Maximal können pro BTS 64 Frequenzen für das Frequency Hopping verwendet werden. Eine Mobilstation bekommt dazu beim Aufbau einer Verbindung in der Immediate Assignment-Nachricht mitgeteilt, welche Frequenzen für seinen Kanal verwendet werden und mit welchem Muster diese gewechselt werden.

Für Carrier, auf denen Broadcast-Kanäle wie SCH, FCCH und BCCH ausgestrahlt werden, darf kein Frequency Hopping verwendet werden. Dies ist zwingend erforderlich, da sonst Endgeräte die Nachbarzellen aufgrund des ständigen Frequenzwechsels nicht finden könnten. In der Praxis zeigt sich, dass Netzbetreiber ihre Zellen sowohl mit, als auch ohne Frequency Hopping betreiben.

Von der BTS werden die Daten aller logischen Kanäle über das Abis Interface und eine E-1-Verbindung an den Base Station Controller weitergeleitet. Die Übertragung erfolgt jedoch in einer gänzlich anderen Rahmenstruktur. Für sämtliche Common Channels sowie die SDCCH und SACCH-Kanäle wird mindestens ein gemeinsamer 64 kbit/s E-1-Timeslot verwendet. Dies ist möglich, da hier nur Signalisierungsdaten übertragen werden, die nicht zeitkritisch sind. Dieser Signalisierungskanal verwendet auf dem BTS – BSC Interface das LAPD-Protokoll. LAPD steht dabei für Link Access Protocol D-Channel und wurde mit wenigen Modifikationen aus der ISDN-Welt übernommen.

Für Traffic Channels, die, wie wir später noch sehen werden, 13 kbit/s an Sprachdaten übertragen, wird jeweils $\frac{1}{4}$ E-1-Timeslot verwendet. Für alle 8 Timeslots eines Air Interface Frames werden somit nur 2 Timeslots auf dem E-1-Interface benötigt (Abb. 6.26). Eine 3-Sektor-Zelle mit jeweils 2 Carrier pro Sektor benötigt somit auf dem Abis Interface 12 Timeslots + 1 Timeslot für die LAPD-Signalisierung. Die restlichen Timeslots können für die Kommunikation zwischen der BSC und einer oder mehreren anderen Basisstationen verwendet werden. Für diesen Anwendungsfall werden diese dann über eine E-1-Leitung in Reihe geschaltet. Dies geschieht heute jedoch nur noch virtuell, da in der Praxis physische E-1 Verbindungen durch virtuelle Verbindungen über IP ersetzt wurden.

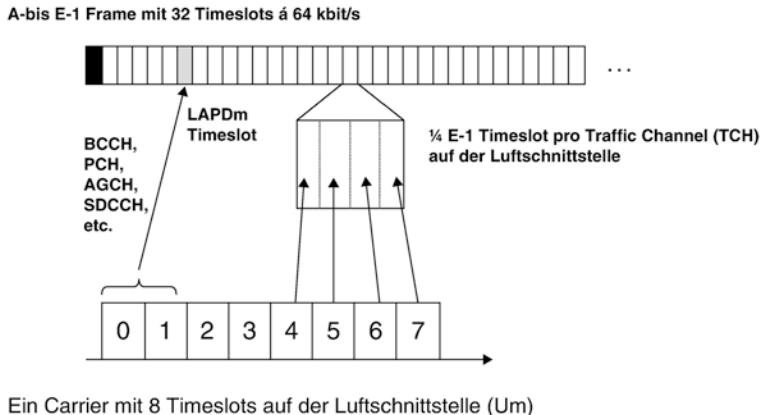


Abb. 6.26 Übertragung der logischen Luftschnittstellenkanäle auf dem A-bis Interface zum BSC

6.7.4 Der Base Station Controller (BSC)

Während die Basisstationen die Schnittstellelemente zu den Endgeräten darstellen, ist der Base Station Controller (BSC) für Aufbau, Abbau und Aufrechterhaltung sämtlicher Verbindungen zu den Endgeräten über alle Basisstationen in seinem Bereich zuständig.

Möchte ein Teilnehmer ein Gespräch beginnen, eine SMS abschicken etc., schickt sein Endgerät dazu wie in Abb. 6.25 dargestellt eine Channel Request-Nachricht an die BSC. Die BSC überprüft daraufhin, ob ein freier Signalisierungskanal (SDCCH) vorhanden ist und aktiviert diesen in der BTS. Danach schickt die BSC auf dem Access Grant Channel (AGCH) eine Immediate Assignment-Nachricht mit der Nummer des zugeteilten SDCCH zum Endgerät zurück. Über die so aufgebaute Signalisierungsverbindung können nun DTAP-Nachrichten transparent zur MSC weitergeleitet werden.

Der zweite Fall für den Aufbau eines Signalisierungskanals ist eine ankommende Verbindung, wie z. B. ein Telefongespräch oder eine SMS. In diesem Fall empfängt der BSC eine Paging-Nachricht von der MSC. Die Paging-Nachricht enthält die IMSI, die TMSI sowie die Location Area, in der sich der gewünschte Teilnehmer momentan aufhält. Die Zellen, die sich in dieser Location Area befinden, sind der Location Area-Datenbank im BSC bekannt. Der BSC leitet daraufhin die Paging-Nachricht an alle Zellen weiter, die sich in dieser Location Area befinden. Nach Empfang der Paging-Nachricht meldet sich das Endgerät beim Netzwerk wiederum wie im ersten Fall gezeigt mit einer Channel Request-Nachricht.

Der Aufbau eines Sprachkanals wird sowohl für ein abgehendes, wie auch für ein ankommendes Gespräch immer von der MSC bei der BSC beantragt. Nachdem sich MSC und Endgerät über die Signalisierungsverbindung (SDCCH) über den Aufbau einer Sprachverbindung verständigt haben, schickt die MSC wie in Abb. 6.27 gezeigt, eine Assignment Request-Nachricht an die BSC.

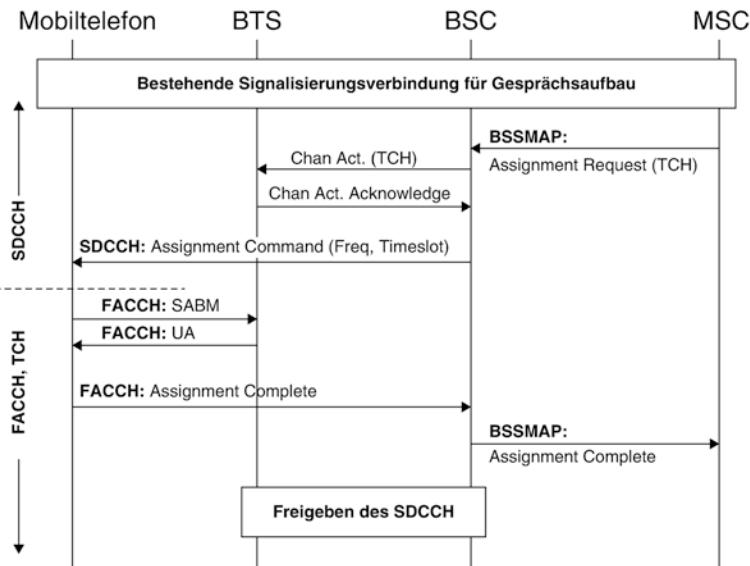


Abb. 6.27 Aufbau eines Sprachkanals (TCH)

Die BSC überprüft daraufhin, ob in der gewünschten Zelle ein freier Traffic Channel (TCH) vorhanden ist und aktiviert diesen in der BTS. Danach wird das Endgerät über den SDCCH benachrichtigt, dass ein TCH für die weitere Kommunikation zur Verfügung steht. Das Endgerät wechselt dann auf den TCH und FACCH und sendet ein SABM Frame zur BTS. Diese sendet daraufhin ein UA Frame als Bestätigung über die korrekte Verbindungsaufnahme an das Endgerät zurück. Danach sendet das Mobiltelefon ein Assignment Complete an die BSC zurück, die diese Nachricht auch an die MSC weitergibt.

Neben dem Auf- und Abbau ist auch die Aufrechterhaltung einer Verbindung eine wichtige Aufgabe des Base Station Controllers. Da Teilnehmer auch während einer Verbindung ihren Standort ändern können, kommt es während einer Verbindung durchaus vor, dass sich Teilnehmer aus dem Versorgungsbereich ihrer aktuellen Zelle hinausbewegen. In diesem Fall muss die BSC einen Wechsel der Verbindung in eine Zelle mit besserer Funkversorgung veranlassen. Dieser Vorgang wird Handover genannt. Um einen Handover durchzuführen, benötigt die BSC Messergebnisse über die Signalqualität auf der Luftschnittstelle. Die Messergebnisse für die Signalqualität im Downlink erhält die BSC vom Endgerät, das die Signalqualität laufend misst und über den SACCH dem Netzwerk mitteilt. Die Uplink-Signalqualität wird ständig von der BTS gemessen und ebenfalls dem BSC mitgeteilt. Neben der Signalqualität der aktuellen Zelle ist es für das Netzwerk weiterhin wichtig zu wissen, wie gut die Nachbarzellen von einem Teilnehmer empfangen werden können. Dazu teilt das Netzwerk dem Endgerät über den SACCH die

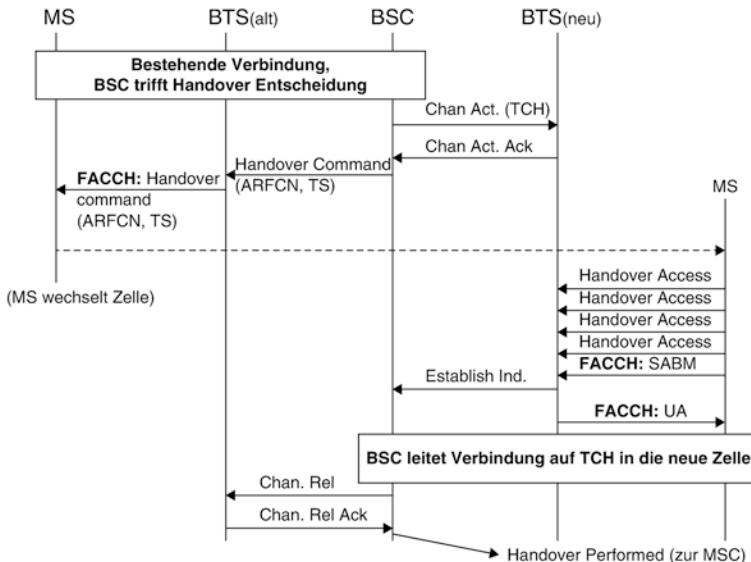


Abb. 6.28 Nachrichtenfluss während eines Handovers

Frequenzen der Nachbarzellen mit, die vom Endgerät dann in den Sendepausen überprüft werden. Auch diese Messergebnisse werden dem Netzwerk über den SACCH mitgeteilt.

Aufgrund dieser Messergebnisse trifft die BSC dann bei Bedarf die Entscheidung, in welche Zelle ein Handover erfolgen soll. Dazu wird als erstes wie in Abb. 6.28 dargestellt in der neuen Zelle ein TCH aktiviert. Danach schickt die BSC dem Endgerät über die alte Zelle ein Handover Command über den Fast Associated Control Channel (FACCH). Wichtige Informationen in dieser Nachricht sind die neue Frequenz und die Nummer des Timeslots des neuen TCH. Das Endgerät ändert dann seine Sende/Empfangsfrequenz, synchronisiert sich ggf. mit der neuen Zelle und sendet in vier aufeinander folgenden Bursts des Timeslots eine Handover Access-Nachricht. Im fünften Burst des Timeslots wird eine SABM-Nachricht gesendet. Hat die BTS den Handover korrekt erkannt, schickt diese eine Establish Indication-Nachricht zum BSC und eine UA-Nachricht zum Endgerät. Die BSC kann daraufhin die Sprachverbindung in die neue Zelle schalten.

Aus Sicht des Endgeräts ist der Handover damit beendet. Die BSC muss jedoch noch den TCH in der alten Zelle abbauen und dem MSC eine Nachricht über den erfolgten Handover schicken. Diese Nachricht ist jedoch nur informativ und hat auf der MSC keinen Einfluss auf den weiteren Verbindungsablauf.

Um Interferenzen möglichst gering zu halten, kontrolliert die BSC während einer Verbindung für jeden Teilnehmer die Sendeleistung auf der Luftschnittstelle. Für Endgeräte hat dies auch den positiven Effekt, dass bei guter Verbindung die Sendeleistung

reduziert werden kann und sich somit die Akkulaufzeit erhöht. Die Regelung erfolgt dabei mithilfe der Signalqualitätsmessungen der BTS. Muss die Sendeleistung erhöht oder abgesenkt werden, sendet die BSC eine entsprechende Änderungsinformation einmalig zur BTS. Die BTS sendet diese dann periodisch am Anfang jedes SACCH Frames zur Mobilstation. Wie sich in der Praxis zeigt, wird eine Leistungsanpassung etwa alle 1–2 s durchgeführt, sofern sich die Signalqualität ändert. Bei Verbindungsaufbau wird dazu immer erst mit einer hohen Sendeleistung begonnen, die dann Schritt für Schritt abgesenkt, bzw. wieder erhöht werden kann. Die nachfolgende Tabelle gibt eine Übersicht über die bei GSM möglichen Leistungsklassen für Endgeräte. Dabei wird zwischen Leistungsklassen für das 900 MHz Band und das 1800 MHz Band unterschieden.

GSM 900 Power Level	GSM 900 Leistung	GSM 1800 Power Level	GSM 1800 Leistung
(0–2)	(8 W)		
5	2 W	0	1 W
6	1,26 W	1	631 mW
7	794 mW	2	398 mW
8	501 mW	3	251 mW
9	316 mW	4	158 mW
10	200 mW	5	100 mW
11	126 mW	6	63 mW
12	79 mW	7	40 mW
13	50 mW	8	25 mW
14	32 mW	9	16 mW
15	20 mW	10	10 mW
16	13 mW	11	6,3 mW
17	8 mW	12	4 mW
18	5 mW	13	2,5 mW
19	3,2 mW	14	1,6 mW
		15	1,0 mW

Während die maximale Sendeleistung für Mobiltelefone im 900 MHz Band 2 W beträgt, ist diese im 1800 MHz Band auf 1 W begrenzt. Für stationäre Geräte oder Autotelefone mit Außenantenne ist im 900 MHz-Bereich eine Sendeleistung bis zu 8 W definiert. Die Leistungsangaben in der Tabelle beziehen sich auf die Leistung, die während der Übertragung in einem einzelnen Timeslot von einem Endgerät erreicht wird. Da das Endgerät aber nur in einem von 8 Timeslots sendet, ist für die gemittelte Leistung der angegebene Wert durch 8 zu teilen. Die maximale durchschnittliche Sendeleistung bei einer Sendeleistung von zwei Watt ist somit nur 250 mW.

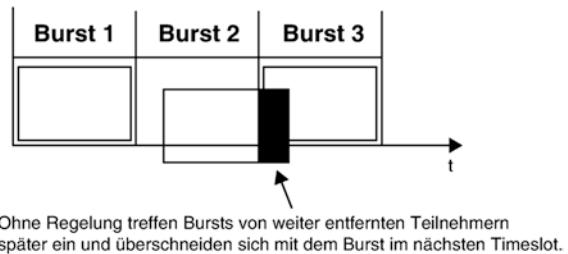
Auch die Sendeleistung der BTS kann von der BSC geregelt werden. Hierfür werden Signalstärke-Messergebnisse des Endgeräts verwendet. Dies ist jedoch in den Standards nur als optional definiert. Die Leistungsregelung im Downlink ist außerdem nur für Timeslots auf Frequenzen möglich, die keine Broadcastkanäle (FCH, SCH, BCCH...) einer Zelle aussenden. Auf solchen Frequenzen muss die Sendeleistung konstant bleiben, damit Teilnehmer in anderen Zellen eine korrekte Nachbarschaftszellenmessung durchführen können. Dies wäre bei einer schwankenden Signalamplitude über die unterschiedlichen Timeslots hinweg nicht möglich.

Entfernt sich ein Teilnehmer während einer aktiven Verbindung von einer Basisstation, benötigen die Funkwellen eines Bursts aufgrund der begrenzten Ausbreitungsgeschwindigkeit der Funkwellen für den längeren Weg mehr Zeit. Würde hier nicht gegengesteuert werden, würde sich der Burst eines Teilnehmers bei zu großer Entfernung trotz der in Abb. 6.22 beschriebenen Guard Time mit dem Burst des Teilnehmers im nächsten Zeitschlitz überschneiden. Aus diesem Grund muss der Sendezeitpunkt für alle Teilnehmer ständig überwacht und angepasst werden. Dabei gilt, dass je weiter ein Teilnehmer entfernt ist er umso früher seinen Burst senden muss, damit dieser zur richtigen Zeit bei der Basisstation eintrifft. Dieses Verfahren wird Timing Advance-Regelung genannt.

Die Regelung erfolgt dabei in 64 Schritten von 0 bis 63. Pro Schritt kann die Entfernung zur Basisstation um 550 m angepasst werden. Die maximale Distanz zwischen einer Basisstation und einem mobilen Teilnehmer kann somit theoretisch $64 \cdot 550 \text{ m} = 35,2 \text{ km}$ betragen. In der Praxis wird eine solche Distanz jedoch nur sehr selten erreicht, da Basisstationen in besiedelten Gebieten wesentlich näher zusammenliegen. Auch reicht die Sendeleistung des Endgeräts nicht aus, diese Entfernung zu überbrücken, da zumeist auch keine direkte Sichtverbindung zwischen Mobiltelefon und Basisstation besteht. Dieser Wert kann allenfalls in Küstennähe von einem Schiff erreicht werden (Abb. 6.29).

Die Regelung des Timing Advance beginnt schon beim ersten Zugriff des Mobiltelefons auf das Netzwerk mit der Channel Request-Nachricht. Diese Nachricht verwendet einen sehr kurzen Burst, der nur sehr wenig Nutzdaten enthalten kann, dafür aber sehr große Guard Periods an Anfang und Ende. Dies ist notwendig, da am Anfang das Mobiltelefon nicht wissen kann, wie weit es von der Basisstation entfernt ist und somit auch noch keinen Timing Advance einstellen kann. Beim Eintreffen der Channel Request-Nachricht bei der BTS misst diese die zeitliche Verzögerung des Bursts. Anschließend leitet die BTS die Channel Request-Nachricht inklusive der gemessenen Verzögerungszeit in Form eines Timing Advance-Wertes an die BSC weiter. Wie in Abb. 6.25 gezeigt wurde, schickt die BSC als Antwort auf die Channel Request-Nachricht eine Immediate Assignment-Nachricht an die Mobilstation zurück. Neben der Nummer des zugeteilten Signalisierungskanals (SDCCH) enthält die Nachricht auch den ersten Timing Advance-Wert, den die Mobilstation für die weitere Kommunikation verwenden soll. Nach erfolgreicher Verbindungsaufnahme über den SDCCH und später evtl. über den TCH misst die BTS ständig die Zeitverzögerung der eintreffenden Bursts und meldet diese in

Abb. 6.29 Zeitverschiebung eines Bursts ohne Timing Advance-Regelung



Form eines Timing Advance-Wertes der BSC weiter. Ändert sich der Timing Advance-Wert, informiert die BSC über den SACCH das Endgerät, das daraufhin seinen Timing Advance-Wert entsprechend korrigiert.

Für Anwendungsfälle wie Küstenkommunikation enthält der GSM-Standard noch eine weitere Timeslotkonfiguration, um die maximale Entfernung zur Basisstation auf bis zu 120 km auszudehnen. Um dies zu ermöglichen, wird in einer Zelle nur jeder zweite Timeslot verwendet und bewusst akzeptiert, dass der Burst sich in den nächsten Timeslot verschiebt. Dies erweitert zwar den Abdeckungsbereich einer Zelle erheblich, dies geht aber sehr zulasten der Anzahl der verfügbaren Kommunikationskanäle. Mobiltelefone, die wie heute üblich auf ein Watt (1800 MHz Band) oder zwei Watt (900 MHz Band) begrenzt sind, mögen zwar den BCCH empfangen können, aufgrund ihrer Sendeleistung wird das Uplink-Signal die Basisstation aber nicht erreichen. Aus diesem Grund können Zellen in solcher Entfernung nur von fest eingebauten Mobiltelefonen verwendet werden, die mit einer Leistung von bis zu 8 W senden können.

6.7.5 Die TRAU für Sprachdatenübertragung

Für die Übertragung eines Sprachdatenkanals über die Luftschnittstelle dient in GSM der in Abschn. 6.7.3 beschriebene Traffic Channel (TCH). Dieser verwendet wie in Abb. 6.24 gezeigt alle Bursts eines 26 Multiframe mit Ausnahme eines Burst für den Slow Associated Control Channel und einen Burst, der für die Nachbarzellen Pegelmessung leer bleibt. Wie im letzten Kapitel außerdem gezeigt wurde, kann ein Burst, der alle 4,615 ms übertragen wird, genau 114 Bit Nutzdaten aufnehmen. Dies entspricht unter Berücksichtigung der zwei nicht für den TCH verwendeten Bursts pro 26-Multiframe einer Bruttodatenrate von 22,8 kbit/s. Wie wir im Laufe dieses Kapitels noch genauer betrachten werden, wird von dieser Bruttodatenrate ein großer Teil für die Fehlererkennung und Fehlerkorrektur verwendet, sodass für die reinen Sprachdaten nur eine Bandbreite von etwa 13 kbit/s zur Verfügung steht.

Dies ist ein Problem, da im Kernnetzwerk immer ein 64 kbit/s E-1-Timeslot für einen Sprachkanal verwendet wird und auch der in Abschn. 6.6.1 vorgestellte PCM-Sprachkodierer diese Bandbreite voll ausnutzt (Abb. 6.30).

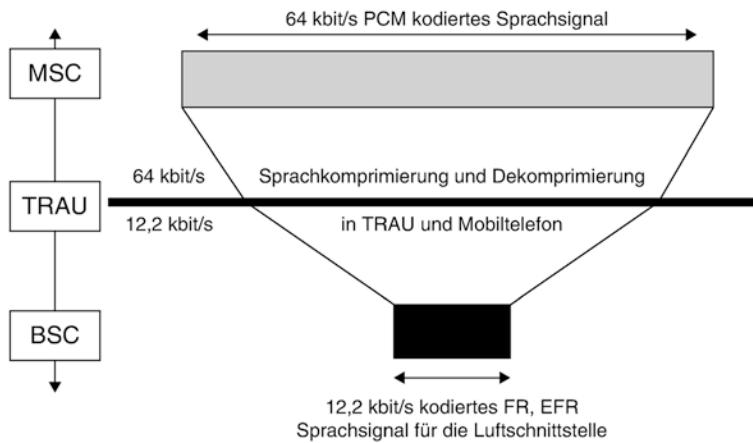


Abb. 6.30 GSM-Sprachdatenkomprimierung

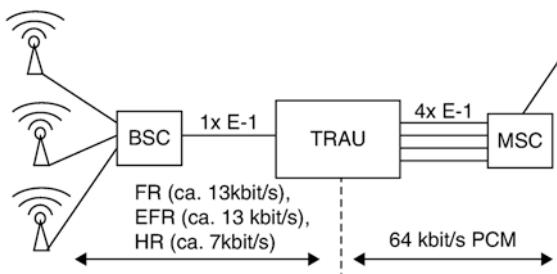


Abb. 6.31 Sprachkompression in Verhältnis 4:1 in der TRAU

Um dieses Problem erst gar nicht entstehen zu lassen, hätte der GSM-Standard auch 64 kbit/s Sprachkanäle auf der Luftschnittstelle definieren können. Die Wahl eines Kanals mit weit geringerer Bandbreite wurde aber ganz bewusst getroffen, um möglichst viele Sprachkanäle über die knappen Ressourcen auf der Luftschnittstelle übertragen zu können. Dies wurde auch deshalb möglich, da zu Beginn der Standardisierung in den 80er-Jahren absehbar war, dass die technischen Möglichkeiten zur Komprimierung der Sprachdaten von 64 kBit/s auf 13 kBit/s in Echtzeit durch neue Hardwareentwicklungen möglich wurde.

Im Mobilfunknetzwerk wird die Komprimierung und Dekomprimierung der Sprachdaten durch die Transcoding and Rate Adaptation Unit (TRAU) durchgeführt. Diese wird zwischen einer MSC und einem BSC geschaltet und von der BSC kontrolliert (Abb. 6.31).

Die MSC schickt dabei die Sprachdaten im 64 kbit/s PCM-Format in Richtung Radio-Netzwerk. In der TRAU wird das Sprachsignal dann auf etwa 13 kbit/s komprimiert und zur BSC weitergeschickt. In der Gegenrichtung dekomprimiert die TRAU das von der

BSC erhaltene 13 kbit/s-Sprachsignal wieder in das 64 kbit/s PCM-Format und gibt es an die MSC weiter. Im Endgerät auf der anderen Seite der Luftschnittstelle sind die Algorithmen für die Komprimierung und Dekomprimierung des Sprachsignals ebenfalls implementiert.

Obwohl die TRAU eine logische Komponente des BSS ist, wird diese in der Praxis normalerweise direkt neben einer MSC aufgestellt. Dies hat den Vorteil, dass nach der Komprimierung des Sprachdatensignals vier Sprachkanäle mit je 13 kbit/s auf einem einzigen E-1-Timeslot übertragen werden können. Jeder Sprachkanal belegt damit einen 16 kbit/s Subtimeslot. Somit wird nur $\frac{1}{4}$ der Übertragungskapazität zwischen MSC und BSC benötigt. Da die BSCs normalerweise in größerer Entfernung zur MSC aufgestellt werden, ergibt sich dadurch eine deutliche Kosteneinsparung für den Netzbetreiber.

Die TRAU bietet eine Anzahl unterschiedlicher Algorithmen für die Sprachkomprimierung. Diese werden auch Sprachcodecs oder Codecs genannt. Der als erstes implementierte Codec wurde Full Rate Codec (FR) genannt und komprimiert das Sprachsignal in Echtzeit auf etwa 13 kbit/s.

Ende der 90er-Jahre wurde ein weiterer Codec eingeführt, der sich Enhanced Full Rate Codec (EFR) nennt. Auch der EFR Codec komprimiert das Sprachsignal auf etwa 13 kbit/s, bietet aber eine bessere Sprachqualität. Nachteil ist der wesentlich komplexere Komprimierungsalgorithmus, der deutlich mehr Rechenkapazität benötigt. Dies spielt aber bei heutigen Mobiltelefonen auch im Niedrigpreissegment aufgrund der gestiegenen Prozessorleistung keine Rolle mehr.

Neben diesen zwei Codecs gibt es den Half Rate Codec (HR), der nur 7 kbit/s Bandbreite benötigt. Während beim Enhanced Full Rate Codec fast kein Unterschied zum original 64 kbit/s PCM-Signal zu hören ist, ist die Sprachqualität beim Half Rate Codec deutlich schlechter. Vorteil für den Netzbetreiber ist jedoch, dass sich die Anzahl der möglichen Sprachverbindungen über eine BTS verdoppelt. Auf einem Timeslot, der normalerweise für einen TCH (EFR) benötigt wird, können auf diese Weise zwei TCH (HR) übertragen werden. In der Praxis scheinen die Netzbetreiber den Half Rate Codec jedoch nicht oft einzusetzen.

Eine weitere Sprachcodec-Entwicklung ist der Adaptive Multi Rate -Algorithmus, auch AMR genannt. Dieser ist in 3GPP TS 26.071 spezifiziert und wird heute von den meisten Endgeräten und Netzwerken unterstützt. Statt sich wie bei FR, EFR und HR bei Beginn der Sprachverbindung auf einen Codec festzulegen, erlaubt der Adaptive Multi Rate-Algorithmus den Wechsel des verwendeten Codecs auch während der Verbindung. Ein wesentlicher Vorteil dieses Verfahrens ist, bei einer schlechten Verbindung auf einen Sprachcodec mit höherer Kompression umzuschalten und dafür die Anzahl der Bits für Fehlererkennung und Fehlerkorrektur zu erhöhen. Andererseits kann bei einer guten Verbindung die Kapazität der Zelle gesteigert werden, in dem ein Codec mit niederer Bitrate gewählt wird und nur ein Timeslot in jedem zweiten Frame für ein Gespräch verwendet wird (Half-Rate AMR). Weitere Informationen über AMR sind im Kapitel über UMTS zu finden.

Die neueste Sprachcodec-Entwicklung für GSM Netzwerke stellt der AMR-Wideband Codec dar, der in ITU G.722.2 und 3GPP TS 26.190 spezifiziert ist. Wie sein Name schon andeutet, wird mit diesem Algorithmus ein breiteres Frequenzspektrum digitalisiert, als dies mit dem vorgestellten PCM Algorithmus möglich ist. Statt bis 3400 Hz wie mit dem PCM Codec wird das Sprachsignal mit AMR-WB bis 7000 Hz digitalisiert, was eine wesentlich natürlichere Sprachwiedergabe beim Empfänger ermöglicht. In der Praxis wird zumeist eine hohe Datenkompressionsrate gewählt, um eine Datenrate von 12,65 kbit/s zu erreichen. Somit passt ein AMR-WB Datenstrom problemlos in einen GSM Timeslot und erfordert auch bei UMTS keine zusätzlichen Übertragungskapazitäten. Da sich AMR-WB aufgrund des erweiterten Frequenzspektrums nicht in den zwischen der TRAU und der MSC verwendeten PCM Codec umsetzen lässt, überträgt die TRAU den AMR-WB Datenstrom nicht als PCM Codec in einem 64 kbit/s Timeslot, sondern fügt diesen transparent ein. Dies bedeutet, dass die meisten Bits im Timeslot ungenutzt bleiben, da die Datenrate ja nur 12,65 kbit/s beträgt.

Während der bereits vorgestellte PCM-Algorithmus im wesentlichen analoge Pegel über eine vorgegebene Kurve in digitale Werte umwandelt, ist die GSM-Sprachdigitalisierung wesentlich komplexer aufgebaut, um die gewünschte Kompression zu erreichen. Im Falle des Full Rate Codecs, der im GSM-Standard 46.010 spezifiziert ist, erfolgt die Komprimierung durch Nachbildung der menschlichen Spracherzeugung. Als mathematische Grundlage dient ein Quelle–Filter-Modell. Die menschliche Spracherzeugung im Kehlkopf und mit den Stimmbändern wird in diesem Modell durch die Quelle repräsentiert. Die Filter repräsentieren die Signalformung, die beim Mensch im Rachen und Mundraum stattfindet.

Mathematisch wird die Sprachformung durch zwei zeitvariante Filter nachgebildet. Der Periodenfilter bildet dabei die periodischen Vibratoren der menschlichen Sprache nach, der VokaltraktfILTER simuliert die Hüllkurve der menschlichen Sprache. Die für die Filter notwendigen Parameter werden aus dem Eingangssignal gebildet. Um menschliche Sprache zu digitalisieren und zu komprimieren, wird dieses Modell wie in Abb. 6.32 gezeigt in umgekehrter Reihenfolge angewandt. Da zeitvariante Filter schwer nachzubilden sind, wird das Modell noch deutlich vereinfacht, in dem die Filterparameter für die Zeit von 20 ms als konstant betrachtet werden.

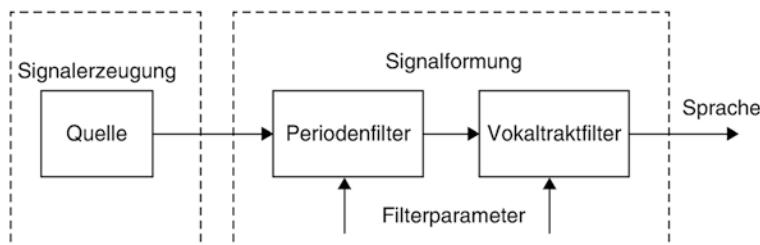


Abb. 6.32 Quelle-Filter-Modell des Full Rate Codecs

Als Eingangssignal dient dem Kompressionsalgorithmus ein nach dem PCM-Verfahren digitalisiertes Sprachsignal, das wie bereits gezeigt pro Wert 8 (oder 13) Bit verwendet. Da der PCM-Algorithmus pro Sekunde 8000 Werte liefert, benötigt der Full Rate Codec für die Berechnung der Filterparameter alle 20 ms genau 160 Werte. Bei 8 Bit pro Wert ergibt dies $8 \text{ Bit} * 160 \text{ Werte} = 1280$ Eingangsbits, bei 13 Bits pro Wert entsprechend mehr. Für den Periodenfilter wird aus diesen Eingangsbits dann ein 36 Bit langer Filterparameter berechnet. Danach wird dieser Filter auf das Eingangssignal angewandt. Mit dem daraus entstandenen Ergebnis wird ein weiterer 36 Bit langer Filterparameter für den Vokaltraktfilter berechnet und der Filter daraufhin wieder entsprechend auf das Signal angewandt. Das so entstandene Restsignal wird in insgesamt 188 Bit kodiert.

Übertragen werden anschließend die Filterparameter mit jeweils 36 Bit Länge, sowie das in 188 Bit kodierte Restsignal. Somit werden statt den ursprünglichen 1280 Eingangsbits nur $36 + 36 + 188 = 260$ Bits übertragen. Auf der Gegenseite wird der Filtervorgang in umgekehrter Reihenfolge auf das Restsignal durchgeführt und das ursprüngliche Sprachsignal somit wiederhergestellt. Da das Verfahren verlustbehaftet arbeitet, ist das wiederhergestellte Signal nicht mehr mit dem Original identisch. Dies ist der Grund, warum sich ein mit dem Full Rate Decoder komprimiertes und wieder dekomprimiertes

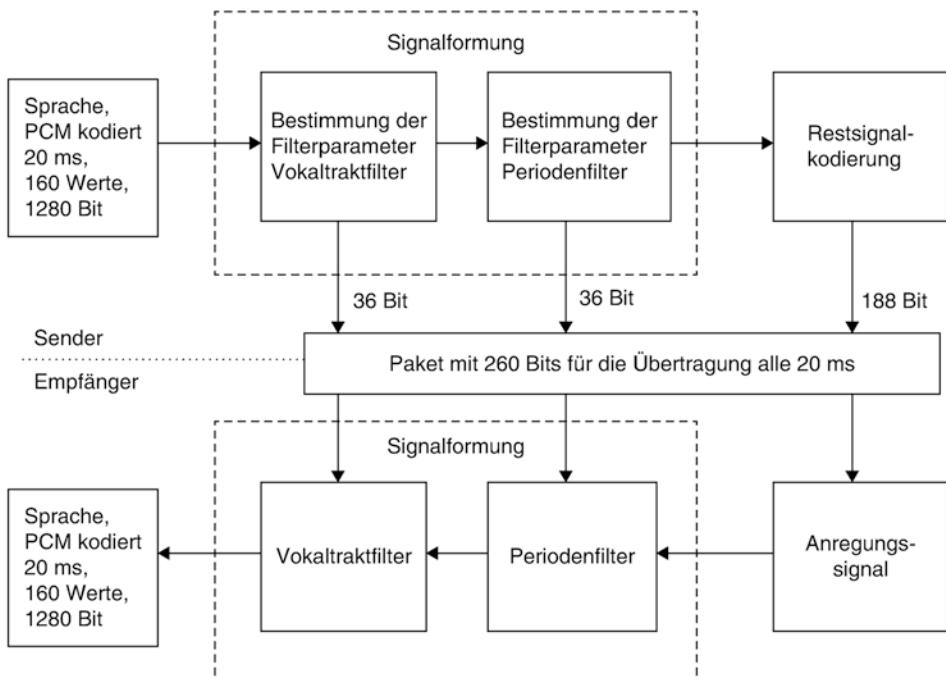


Abb. 6.33 Komplette Übertragungskette mit Sender und Empfänger des GSM Full Rate Codec

Sprachsignal hörbar vom ursprünglichen PCM-Signal unterscheidet. Mit dem Enhanced Full Rate Coder, der nach einem komplexeren Algorithmus arbeitet, ist dieser Unterschied jedoch fast unhörbar geworden (Abb. 6.33).

6.7.6 Channel Coder und Interleaver in der BTS

Bevor dieses 260 Bit-Datenpaket alle 20 ms über die Luftschnittstelle übertragen wird, durchläuft es noch eine Reihe von weiteren Verarbeitungsschritten, die nicht in der TRAU, sondern in der Basisstation durchgeführt werden. Diese sind im Überblick in Abb. 6.34 dargestellt.

Im Kanalkodierer werden dem eigentlichen Nutzdatenstrom Fehlererkennungs- und Fehlerkorrekturinformationen hinzugefügt. Dies ist sehr wichtig, da die Übertragung über die Luftschnittstelle aufgrund der sich ständig ändernden Bedingungen sehr störanfällig ist. Außerdem machen sich aufgrund der stark komprimierten Sprachdatenübertragung schon wenige Fehler später deutlich bemerkbar. Um dies zu vermeiden, werden die 260 Bits des Sprachdatenblocks wie in Abb. 6.35 gezeigt in drei unterschiedliche Klassen eingeteilt:

50 Bits des 260 Bit-Sprachpaketes werden zur ersten Klasse (Class Ia) gezählt. Sie sind extrem wichtig und dürfen unter keinen Umständen bei der Übertragung verfälscht werden. Solche Bits sind z. B. die höherwertigen Bits der FR Coder-Filterparameter. Um dies zu gewährleisten, wird eine 3 Bit CRC-Checksumme gebildet und in den Datenstrom eingefügt. Wird auf der Empfängerseite festgestellt, dass hier ein Fehler aufgetreten ist, wird das komplette Datenpaket verworfen.

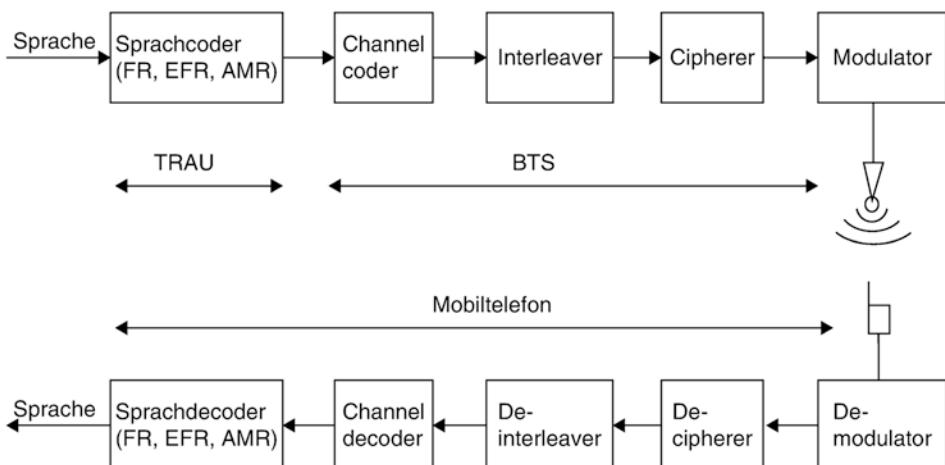
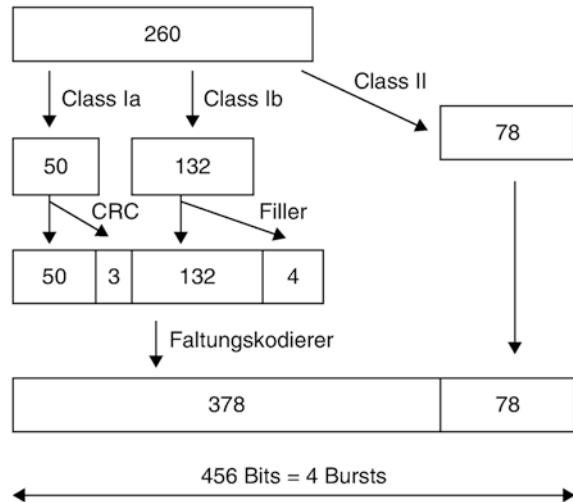


Abb. 6.34 Übertragungsschritte im Downlink zwischen Netzwerk und Mobiltelefon

Abb. 6.35 GSM-Kanalkodierer für FR Sprachdaten



Die 132 Bits der zweiten Klasse (Class Ib) sind auch wichtig, werden aber nicht durch eine Checksumme geschützt. Um später eine vorgegebene Anzahl an Bits am Ausgang des Kanalkodieres zu erhalten, werden am Ende der Klasse Ib vier Füllbits eingefügt. Die Bits der Klasse Ia, die CRC Checksumme, die Bits der Klasse Ib und die vier Füllbits werden dann einem Faltungskodierer übergeben, der den Daten Redundanz hinzufügt. Für jedes Eingangsbit berechnet der Faltungskodierer, im englischen Convolutional Coder genannt, zwei Ausgangsbits. Für die Berechnung der zwei Ausgangsbits wird nicht nur der Wert des aktuellen Bits herangezogen, sondern auch die der vorangegangenen Bits. Da für jedes Eingangsbit genau zwei Ausgangsbits berechnet werden, spricht man auch von einem $\frac{1}{2}$ -Rate Convolutional Coder.

Zur dritten Klasse (Class II) gehören 78 Bits des ursprünglichen 260 Bit-Datenpakets. Diese werden ohne Checksumme und ohne Redundanz übertragen. Fehler, die hier auftreten, können weder erkannt noch korrigiert werden.

Aus den ursprünglichen 260 Bits erstellt der Kanalkodierer somit 456 Bits. Da pro Burst auf der Luftschnittstelle 114 Bits an Daten übertragen werden, entspricht dies somit genau 4 Bursts. Da ein Burst eines TCHs alle 4,6152 ms übertragen wird, ergibt dies somit in etwa wieder 20 ms. Um exakt auf eine Übertragungszeit von 20 ms für diese Daten zu kommen, muss noch der Burst für den SACCH und der leere Burst für die Nachbarzellenmessung eines 26 Multiframe in die Rechnung einbezogen werden.

Durch die im Kanalkodierer hinzugefügte Redundanz ist es möglich, auch eine größere Anzahl an Fehlern pro Datenblock zu korrigieren. Der Faltungskodierer hat jedoch eine Schwachstelle: Werden direkt aufeinander folgende Bits während der Übertragung auf der Luftschnittstelle verfälscht, kann der Faltungsdecodierer auf der anderen Seite die ursprünglichen Daten nicht korrekt wiederherstellen. Dieser Effekt tritt aber

sehr häufig bei ungünstigen Übertragungsbedingungen auf, da Übertragungsstörungen dann meist länger als eine Bitperiode dauern.

Um diesen Effekt zu vermeiden, verteilt der Interleaver die Bits eines 456 Bit-Datenblocks nach einem vorgegebenen Muster über insgesamt 8 Bursts. Aufeinanderfolgende Datenblöcke greifen somit ineinander. Auf der Empfängerseite werden die Datenbits dann wieder durch den Deinterleaver in die richtige Reihenfolge gebracht. Werden nun an einer Stelle viele Bits hintereinander verfälscht, verteilt der Deinterleaver diese somit über das ganze Datenpaket, und der Faltungskodierer kann dies entsprechend korrigieren (Abb. 6.36).

Ein Nachteil dieses Verfahrens ist jedoch eine längere Verzögerung (Delay) des Sprachsignals. Zusätzlich zu den 20 ms des Full Rate Coders, kommen im Interleaver noch weitere 40 ms hinzu, da ein Sprachblock nun über 8 Bursts verteilt wird und nicht direkt in 4 Blocks übertragen wird. Bei einem Gespräch von Mobiltelefon zu Festnetzanschluss ergibt sich dadurch somit mindestens eine Verzögerung von 60 ms. Von Mobiltelefon zu Mobiltelefon sind es dagegen schon mindestens 120 ms, da hier die Kette zweimal durchlaufen wird.

6.7.7 Verschlüsselung

Als nächster Schritt in der Übertragungskette folgt der Cipherer, der vom Interleaver erhaltene Datenpakete verschlüsselt. GSM verwendet dazu einen Stream Cipher-Algorithmus. Dazu wird im Authentication Center und auf der SIM-Karte aus einer

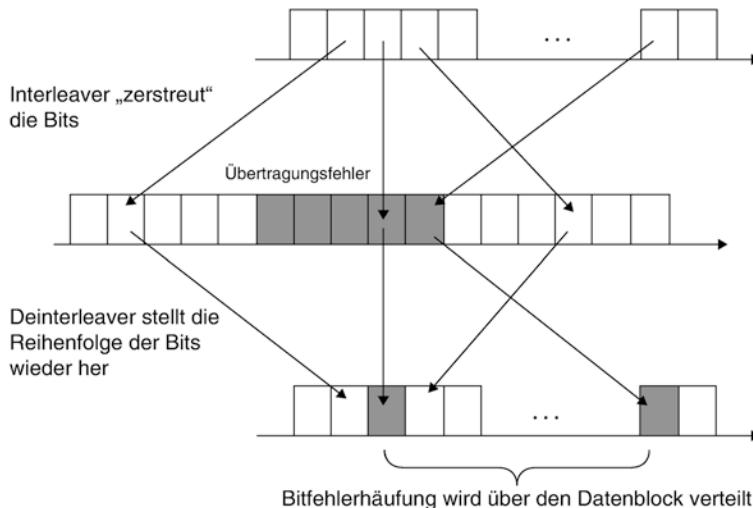


Abb. 6.36 Funktionsweise des Interleavers

Zufallszahl (RAND), dem geheimen Schlüssel K_i und dem Algorithmus A8 der Ciphering Key K_c errechnet. Zusammen mit der GSM-Framenummer, die nach der Übertragung jedes Frames erhöht wird, bildet K_c die Eingangsparameter für den Verschlüsselungsalgorithmus A5. Dieser berechnet nun eine 114 Bit lange Sequenz, mit der die Originaldaten für einen Burst dann Bit für Bit Exklusiv Oder (XOR) verknüpft werden. Da sich die Frame Nummer bei jedem Burst ändert, ist gewährleistet, dass sich auch die 114 Bit Schlüsselsequenz für jeden Burst ändert und somit die Sicherheit des Verfahrens weiter erhöht wird (Abb. 6.37).

Um möglichst flexibel zu sein, können bei GSM mehrere Ciphering-Algorithmen verwendet werden, die A5/1, A5/2, A5/3, A5/4... genannt wurden. Eine der Absichten dahinter war es, GSM-Netze auch in Länder exportieren zu können, in die manche Verschlüsselungsalgorithmen nicht exportiert werden durften. Außerdem sorgt diese Flexibilität dafür, dass in einem bestehenden Netzwerk jederzeit ein neuer Verschlüsselungsalgorithmus eingeführt werden kann, um Sicherheitsprobleme durch die Verwendung eines neuen Algorithmus zu lösen. Dies wurde z. B. durch die Einführung von A5/4 Anfang der 2020er Jahre von vielen Netzbetreibern genutzt. Die Wahl des verwendeten Algorithmus hängt jedoch auch vom Endgerät ab. Damit das Netzwerk einen geeigneten Verschlüsselungsalgorithmus für eine Verbindung wählen kann, informiert das Endgerät dafür bei Verbindungsaufnahme das Netzwerk über die unterstützten Algorithmen.

Da bei Beginn der Kommunikation die Identität des Teilnehmers dem Netzwerk nicht bekannt ist, muss sich das Endgerät vor dem Aktivieren der Verschlüsselung zuerst authentifizieren. Dieser Vorgang wurde in Abschn. 6.6.4 beschrieben. Die Aktivierung

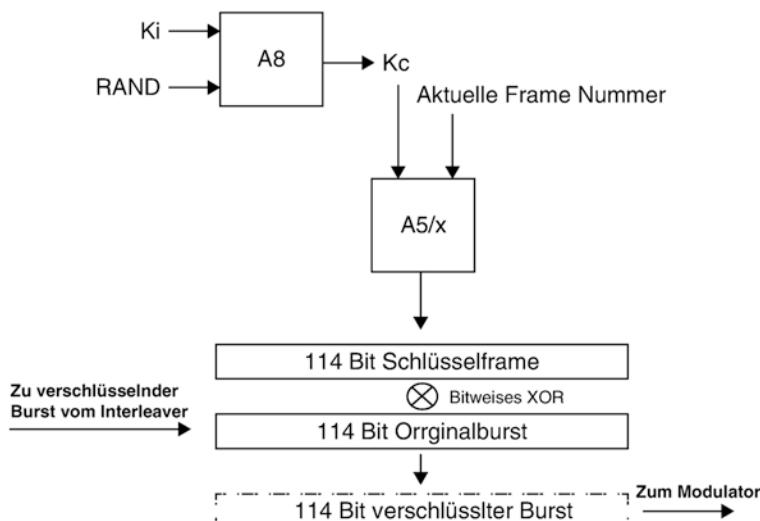


Abb. 6.37 Verschlüsselung eines Datenbursts

der Verschlüsselung erfolgt danach mit einer Ciphering Command-Nachricht durch die MSC. Diese Nachricht enthält unter anderem Kc, der von der BTS für die Verschlüsselung verwendet wird. Bevor die Nachricht zum Mobiltelefon weitergeleitet wird, entfernt das BSS jedoch Kc aus der Nachricht, da dieser nicht über die Luftschnittstelle übertragen werden darf. Die Übermittlung von Kc an das Mobiltelefon ist auch nicht notwendig, da die SIM-Karte diesen selber errechnen kann. In Abb. 6.40 wird gezeigt, wie bei der Kommunikation für ein Location Update die Verschlüsselung aktiviert wird.

6.7.8 Modulation

Als letzter Schritt in der Übertragungskette steht der Modulator. Dieser überträgt die digitalen Daten auf einen Träger (Carrier) mit einer Bandbreite von 200 kHz durch Änderung der Trägerfrequenz. Da die Trägerfrequenz nicht beliebig schnell geändert werden kann, kommt hierfür ein Verfahren namens Gaussian Minimum Shift Keying (GMSK) zum Einsatz, das die Flanken der Frequenzänderung abrundet. Dieses Verfahren wurde zum einen aufgrund seiner Modulations- und Demodulationseigenschaften gewählt, die einfach in Hardwarekomponenten umzusetzen ist, und zum anderen, weil es nur geringe Interferenzen auf Nachbarkanälen erzeugt.

6.7.9 Voice Activity Detection

Um die Interferenz auf der Luftschnittstelle zu reduzieren und die Akkulaufzeiten in den Endgeräten zu erhöhen, werden nur Datenbursts gesendet, wenn auch tatsächlich gesprochen wird. Dieses Verfahren wird Discontinuous Transmission (DTX) genannt und kann unabhängig im Uplink und Downlink aktiviert werden. Da üblicherweise nur ein Gesprächspartner zu einer Zeit spricht, kann somit fast immer die Übertragung zumindest in einer der beiden Richtungen abgeschaltet werden. Dies wird von der TRAU im Downlink und vom Endgerät im Uplink durch die Voice Activity Detection (VAD) gesteuert (Abb. 6.38).

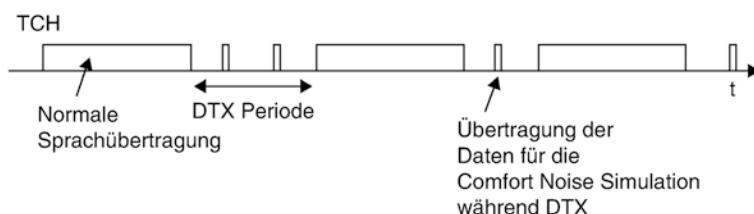


Abb. 6.38 Discontinuous Transmission (DTX)

Würde jedoch der Übertragungskanal einfach abgeschaltet, hätte das eine sehr unangenehme Nebenwirkung. Da nichts mehr übertragen wird, hört der Teilnehmer auch das Hintergrundrauschen des Gesprächspartners nicht mehr. Dies kann sehr irritierend sein, vor allem wenn das Hintergrundrauschen des Gesprächsteilnehmers aufgrund einer Zug- oder Autofahrt sehr laut ist. Deshalb ist es notwendig, während solcher Übertragungspausen ein künstliches Rauschen einzuspielen, das Comfort Noise genannt wird. Da Hintergrundgeräusche jedoch sehr verschieden sind und sich auch mit der Zeit ändern können, analysiert dazu das Mobiltelefon bzw. das Netzwerk das Hintergrundrauschen auf dem Kanal und berechnet eine Approximation. Diese Approximation wird dann nur alle 480 ms zwischen den Teilnehmern ausgetauscht. Außerdem sind diese Frames für Signalstärke und Timing Advance-Messungen notwendig. Wie gut dieses Verfahren arbeitet, ist schon daran zu erkennen, dass die Simulation so gut wie nicht vom Original zu unterscheiden ist.

Trotz ausgefeilter Mechanismen zur Fehlerkorrektur kann nicht ausgeschlossen werden, dass Daten bei der Übertragung unwiederbringlich zerstört werden. In solchen Fällen wird der komplette 20 ms Sprachdatenblock vom Empfänger verworfen und stattdessen der vorige Datenblock nochmals verwendet. Meist bleiben Fehler, die mit diesem Trick ausgebessert werden, unhörbar. Dieser Trick funktioniert aber nicht auf Dauer. Wird auch nach 320 ms kein korrekter Datenblock empfangen, wird der Sprachkanal stumm geschaltet und weiter versucht, einen Datenblock korrekt zu dekodieren. Wird innerhalb der nächsten Sekunden dann weiterhin kein korrekter Datenblock empfangen, wird die Verbindung abgebrochen.

Viele der vorgestellten Verfahren wurden speziell für Sprachdaten entwickelt. Für leitungsvermittelnde Datenverbindungen wie z. B. für die Faxübertragung müssen diese modifiziert bzw. können gar nicht angewandt werden. Die im letzten Absatz besprochenen Verfahren bei nicht korrigierbaren Übertragungsfehlern können beispielsweise nicht für die Datenübertragung angewandt werden. Werden Bits nicht korrekt übertragen, müssen diese von neuem übertragen werden, da ein Datenverlust von den meisten Anwendungen im Unterschied zur Sprachübertragung nicht akzeptiert werden kann. Um die Wahrscheinlichkeit für die korrekte Wiederherstellung der Daten zu erhöhen, wird ein Datenblock über wesentlich mehr als 8 Bursts vom Interleaver gestreut. Auch der Kanalkodierer, der die Bits in Klassen nach deren Wichtigkeit sortiert, muss für die Datenübertragung modifiziert werden, da hier alle Bits gleich wichtig sind und somit der Faltungskodierer auf alle Bits angewandt werden muss. Schließlich kann auch keine Datenreduktion wie bei der Sprache stattfinden, die TRAU verhält sich somit bei Datenübertragungen transparent. Sollten die Daten komprimierbar sein, ist dies von der jeweiligen Anwendung vor der Übertragung selber durchzuführen.

Mit einem Radioempfänger bzw. Stereoanlagenverstärker können die in den vorangegangenen Absätzen beschriebenen Sendezustände während eines Gesprächs auch gehört werden. Dies ist möglich, da das An- und Abschalten des Senders im Endgerät Störungen in der Verstärkerstufe verursachen. Hält man ein GSM-Telefon nahe an ein eingeschaltetes Radio oder einen Verstärker, ist beim Gesprächsaufbau zuerst das

typische Geräuschemuster zu hören, das ein GSM-Telefon auf einem Signalisierungs-kanal (SDCCH) verursacht. Bei Aufbau eines Sprachkanals ist dann der Wechsel auf einen Traffic Channel (TCH) deutlich zu hören. Da für einen TCH alle 4,615 ms ein Burst gesendet wird, wird der Sender mit einer Frequenz von etwa 217 Hz kontinuierlich an- und abgeschaltet. Sind die Hintergrundgeräusche gering oder wird das Mikrofon abgeschaltet, wechselt das Endgerät nach kurzer Zeit in den DTX-Zustand. Auch dies kann gehört werden, da dann statt dem kontinuierlichen 217 Hz Rauschen nur noch etwa alle 0,5 s Bursts gesendet werden.

Bei ankommenden Gesprächen kann man mit dieser Methode auch feststellen, dass das Mobiltelefon schon 1–2 s vor dem eigentlichen ‚Klingeln‘ auf dem SDCCH aktiv wird. Diese Verzögerung kommt dadurch zustande, da das Endgerät den Benutzer erst nach erfolgreicher Authentifizierung, Aktivierung der Verschlüsselung und Aufbau eines Traffic Channels über den Anruf informieren kann. Dies ist auch der Grund, warum der Gesprächsaufbau zu einem mobilen Endgerät länger dauert als zu einem Endgerät im Festnetz.

Manche Endgeräte verfügen über versteckte Netzmonitorfunktionen, die über das normale Menü nicht zugänglich sind. Mit einem solchen Netzmonitor können viele der in diesem Kapitel vorgestellten Abläufe und Parameter wie Timing Advance, Kanal-zuteilung, Leistungsregelung, Cell-ID, Nachbarzelleninformation, Handover, Cell Reselections und vieles mehr beobachtet werden. Im Internet gibt es diverse Websites, die beschreiben, wie dieser Monitormode aktiviert werden kann. Da dies nicht bei allen Endgerätetypen möglich ist und sich die Aktivierungsprozedur von Typ zu Typ unterscheidet, kann hier keine allgemeingültige Anleitung gegeben werden. Im Internet sind jedoch Anleitungen für diverse Endgeräte mit Suchbegriffen wie „Netzmonitor“, „Netmonitor“, „monitoring mode“, etc. zu finden.

6.8 Mobility Management und Call Control

Nachdem in den vorangegangenen Abschnitten alle Komponenten eines Mobilfunknetzwerkes vorgestellt wurden, zeigt dieser Abschnitt einige Vorgänge, um die Mobilität der Teilnehmer zu gewährleisten. In einem GSM-Mobilfunknetzwerk gibt es dazu drei wesentliche Abläufe:

6.8.1 Cell Reselection und Location Area Update

Damit das Netzwerk eingehende Verbindungen an einen Teilnehmer weitervermitteln kann, muss dessen Aufenthaltsort bekannt sein. Direkt nach dem Einschalten meldet sich das Endgerät beim Netz an. Damit kennt das Netzwerk den genauen Aufenthaltsort des Teilnehmers, der sich aber danach jederzeit ändern kann. Besteht zu dieser Zeit keine aktive Sprach- oder Datenverbindung, muss sich das Endgerät beim Netzwerk melden.

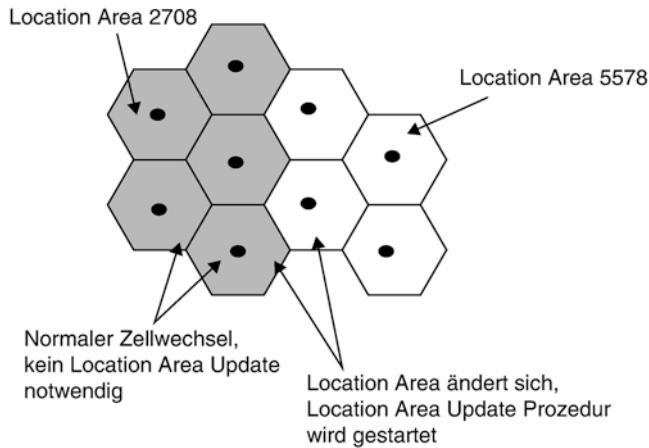


Abb. 6.39 Zellen in verschiedenen Location Areas

Um zu vermeiden, dass dies bei jedem Zellwechsel geschehen muss, werden mehrere Zellen in einer Location Area zusammengefasst. Über den Broadcast Channel (BCCH) informiert das Netzwerk alle Teilnehmer, zu welcher Location Area die aktuelle Zelle gehört. Dazu wird neben der Cell-ID der Zelle auch ständig die Location Area ID ausgestrahlt (Abb. 6.39).

Wenn das Endgerät in eine Zelle einer anderen Location Area wechselt, muss dem Netzwerk dies mit einer Location Area Update-Nachricht mitgeteilt werden. Dieses Verfahren reduziert zum einen die Signalisierungslast des Netzwerkes deutlich und spart zum anderen auch Energie im Endgerät. Nachteil ist jedoch, dass das Netzwerk nur noch die aktuelle Location Area des Teilnehmers kennt, nicht aber die aktuelle Zelle. Bei einem ankommenden Gespräch oder einer SMS muss das Netzwerk dann den Teilnehmer in allen Zellen einer Location Area suchen (Paging). Die Größe der Location Areas kann vom Netzwerkbetreiber festgelegt werden und umfasst meist mehrere dutzend Zellen.

Abb. 6.40 zeigt einen solchen Location Area Update. Nach erfolgreicher Verbindungsaufnahme sendet das Endgerät eine Location Update Request-Nachricht an das Netzwerk. Bevor das Netzwerk diese bearbeitet, wird der Teilnehmer zuerst authentifiziert und danach die Verschlüsselung (Ciphering) aktiviert.

Nachdem die Verbindung so gegen Abhörversuche gesichert ist, wird dem Endgerät eine neue Temporäre ID (TMSI) zugeteilt, die auf der Luftschnittstelle beim Verbindungsaufbau und Paging statt der IMSI verwendet wird. Da eine ständig wechselnde TMSI den Teilnehmer beim nächsten Verbindungsaufbau identifiziert, ist sichergestellt, dass die Identität des Teilnehmers auch während des nicht verschlüsselten Teils der Kommunikation geschützt ist. Nachdem auch diese Prozedur erfolgreich ausgeführt wurde, wird dem Endgerät der erfolgreiche Location Area Update bestätigt und die Verbindung beendet.

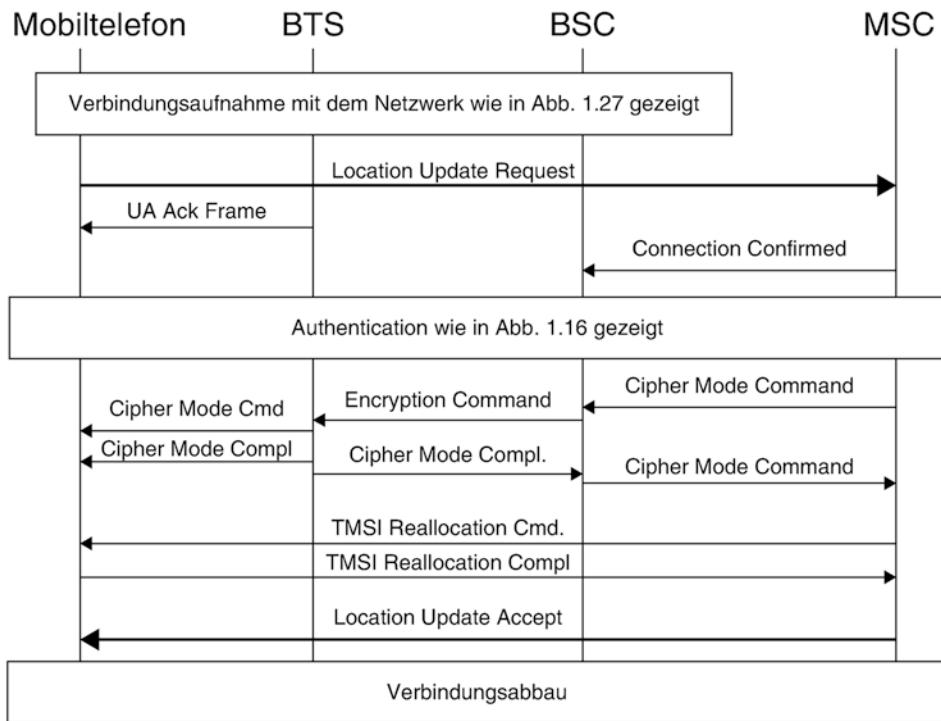


Abb. 6.40 Location Update

Für den Fall, dass die alte und neue Location Area von zwei unterschiedlichen MSC/VLR verwaltet werden, sind noch weitere Schritte notwendig. In diesem Fall muss das neue MSC/VLR das HLR über den Wechsel des Teilnehmers in die neue Area informieren. Das HLR löscht die Daten des Teilnehmers daraufhin im alten MSC/VLR. Dieser Vorgang wird Inter MSC Location Update genannt.

6.8.2 Mobile Terminated Call

Ein Anruf, der bei einem mobilen Teilnehmer eingeht, wird bei GSM als Mobile Terminated Call bezeichnet. Ein wesentlicher Unterschied zwischen Mobilfunknetz und Festnetz ist dabei, dass die Telefonnummer des Teilnehmers keinen Aufschluss mehr über den Aufenthaltsort des Gesprächspartners enthält. Im Mobilfunknetz muss deshalb über das Home Location Register der aktuelle Aufenthaltsort des Teilnehmers ermittelt werden, bevor das Gespräch weitervermittelt werden kann.

Abb. 6.41 zeigt den ersten Teil eines Mobile Terminated Calls, der in diesem Beispiel von einem Festnetzteilnehmer ausgelöst wird. Aus dem Festnetz bekommt dabei die Gateway-MSC (G-MSC) über die schon in Abb. 6.6 gezeigte ISUP-Signalisierung und

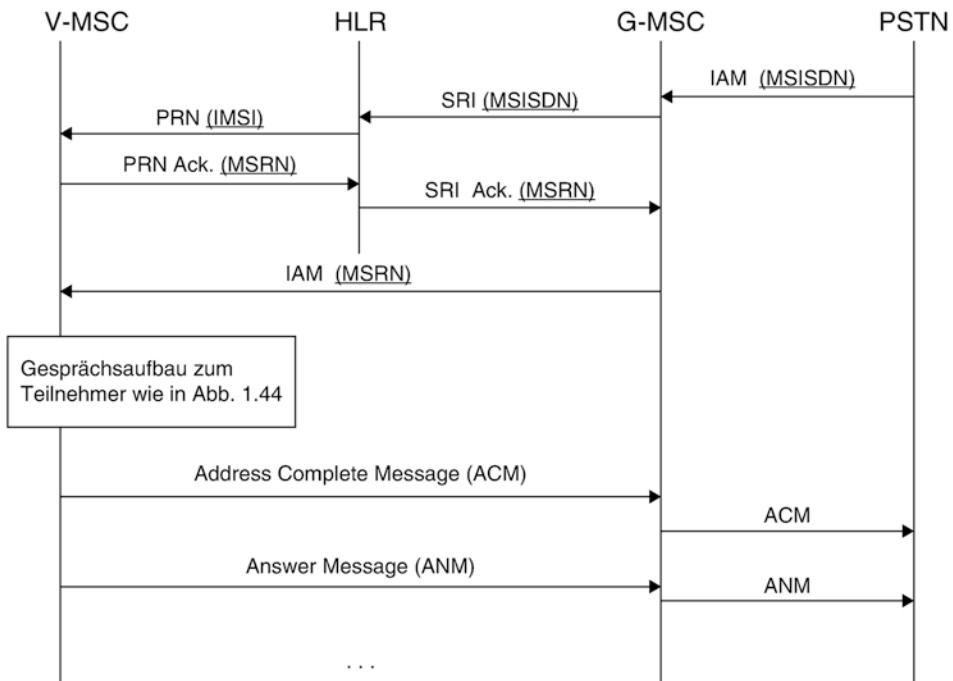


Abb. 6.41 Gesprächsaufbau zu einem mobilen Teilnehmer, Teil 1

die IAM-Nachricht die Telefonnummer (MSISDN) des Gesprächspartners übermittelt. Eine G-MSC, wie in diesem Beispiel gezeigt, ist eine normale MSC mit zusätzlichen Verbindungen in andere Netze. Die G-MSC sendet nach Erhalt der IAM-Nachricht eine Send Routing Information (SRI)-Nachricht an das Home Location Register (HLR), um die aktuelle MSC des Teilnehmers zu ermitteln. Die aktuelle MSC des Teilnehmers wird auch Visited MSC (V-MSC) genannt.

Das HLR ermittelt anhand der übergebenen MSISDN die IMSI des Teilnehmers und findet somit auch seine aktuelle V-MSC und deren VLR. Daraufhin sendet das HLR eine Provide Roaming Number-Nachricht an das V-MSC/VLR, um diese über den ankommenden Anruf zu informieren. Im V-MSC/VLR wird die übergebene IMSI einer temporären Mobile Station Roaming Number (MSRN) zugeordnet, die dann an das HLR zurückgegeben wird. Das HLR gibt die MSRN schließlich transparent an die G-MSC zurück.

Die G-MSC verwendet die so erhaltene MSRN für die Weitervermittlung des Gesprächs an die V-MSC. Dies ist möglich, da die MSRN nicht nur temporär den Teilnehmer in der V-MSC/VLR identifiziert, sondern auch so aufgebaut ist, dass die V-MSC eindeutig identifiziert werden kann. Zwischen G-MSC und V-MSC wird dazu wiederum die ISUP-Signalisierung verwendet. Statt der ursprünglichen MSISDN des Teilnehmers enthält diese IAM-Nachricht jedoch die MSRN. Die MSISDN kann hier nicht mehr verwendet werden, da zwischen G-MSC und V-MSC durchaus noch mehrere Vermittlungsstellen geschaltet sein können.

Da die MSRN nicht nur im nationalen Netz, sondern auch international eindeutig ist, kann über dieses Verfahren auch ein Teilnehmer erreicht werden, der sich gerade im Ausland aufhält. Für das Netzwerk macht es also keinen Unterschied, ob sich ein Teilnehmer im eigenen Netzwerk oder im Ausland befindet. Da die MSRN für die spätere Abrechnung im Billing Record gespeichert wird, ist es auch möglich, dem Teilnehmer eine Gebühr für die Weitervermittlung ins Ausland in Rechnung zu stellen und einen Teil dieser Gebühr an den ausländischen Netzbetreiber zu überweisen.

In der V-MSC/VLR wird die MSRN dann verwendet, um die IMSI des Teilnehmers und seine Daten im VLR zu finden. Dies ist möglich, da bei der Zuteilung der MSRN bei der Anfrage des HLR diese Beziehung gespeichert wurde. Nachdem die Teilnehmerdaten im VLR gefunden wurden, wird nun der Teilnehmer von der MSC in der Location Area im Radionetzwerk gesucht, die in seinem VLR-Eintrag gespeichert ist. Dieser Vorgang wird Paging genannt und ist in Abb. 6.42 dargestellt. Dazu schickt die V-MSC eine Paging-Nachricht an die entsprechende BSC. Die BSC wiederum schickt daraufhin in jede Zelle der betreffenden Location Area eine Paging-Nachricht, die dann auf dem Paging Channel (PCH) ausgestrahlt wird. Meldet sich der Teilnehmer nicht innerhalb weniger Sekunden, wird die Paging-Nachricht wiederholt.

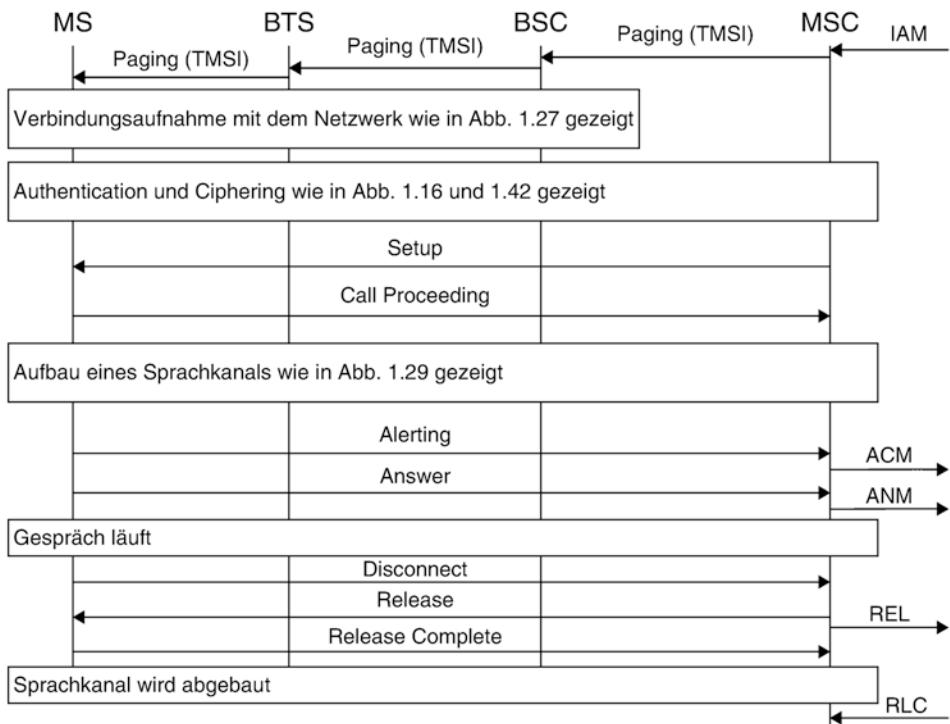


Abb. 6.42 Gesprächsaufbau zu einem mobilen Teilnehmer, Teil 2

Nachdem sich das Endgerät beim Netzwerk gemeldet hat, finden wie beim Location Update wieder eine Authentifizierung und Aktivierung der Verschlüsselung statt. Erst danach wird das Endgerät über den eingehenden Anruf über eine Setup-Nachricht informiert. Teil dieser Nachricht ist z. B. die Telefonnummer des Anrufers, falls dieses Dienstmerkmal aktiviert ist (CLIP) und nicht von der Anruferseite unterdrückt wird (CLIR).

Bestätigt das Endgerät den eingehenden Anruf mit einer Call Proceeding-Nachricht, beantragt die MSC bei der BSC den Aufbau eines Sprachkanals (TCH).

Nach erfolgreichem Aufbau des Sprachkanals schickt das Endgerät eine Alerting-Nachricht zur MSC und teilt ihr dadurch mit, dass der Teilnehmer über den eingehenden Anruf informiert wird (das Telefon „klingelt“). Die V-MSC ihrerseits gibt diese Information über die Address Complete-Nachricht (ACM) an die G-MSC weiter. Auch diese gibt die Information über eine ACM-Nachricht an das Festnetz weiter.

Nimmt der mobile Teilnehmer das Gespräch an, schickt das Endgerät eine Answer-Nachricht zur V-MSC. Diese leitet die Information dann über eine Answer-Nachricht (ANM) zur G-MSC weiter. Von dort aus wird dann das Festnetz wiederum durch eine ISUP ANM darüber informiert, dass das Gespräch durchgeschaltet wurde.

Auch während der eigentlichen Sprachverbindung werden ständig Signalisierungsnachrichten ausgetauscht. Am häufigsten werden zweifellos Nachrichten mit Messergebnissen zwischen Endgerät, BTS und BSC ausgetauscht. Wenn nötig, kann die BSC während der bestehenden Verbindung ein Handover zu einer anderen Zelle veranlassen. Mehr dazu in Abschn. 6.8.3.

Beendet einer der beiden Teilnehmer das Gespräch, schickt die jeweilige Seite eine Disconnect-Nachricht. Nach Abbau des Sprachkanals zum Endgerät und dem Senden einer ISUP-Release Complete-Nachricht ist die Verbindung dann komplett beendet.

In diesem Beispiel wurde davon ausgegangen, dass sich der mobile Teilnehmer nicht im Bereich der G-MSC aufhält. Dies kann aber durchaus vorkommen, wenn z. B. ein Gespräch von einem Festnetzteilnehmer zu einem Mobilfunkteilnehmer aufgebaut wird, der sich in der gleichen Region befindet. Da Festnetzvermittlungsstellen das Gespräch aus Kostengründen meist an die nächstgelegene Mobilfunkvermittlungsstelle weitergeben, kann somit die G-MSC auch gleichzeitig die V-MSC sein. Dies erkennt die G-MSC nach Erhalt der MSRN in der SRI Acknowledge-Nachricht. In diesem Fall wird das Gespräch dann gleich intern behandelt, und die ISUP-Signalisierung (IAM, ACM, ANM...) entfällt.

6.8.3 Handoverszenarien

Verschlechtert sich der Empfang während einer Verbindung z. B. aufgrund einer Positionsänderung des Teilnehmers zusehends, leitet die BSC einen Handover ein. Das grundsätzliche Verfahren und die dazu notwendigen Nachrichten wurden bereits in

Abb. 6.28 dargestellt. Im Netzwerk werden laut 3GPP TS 23.009 folgende Handoverfälle unterschieden:

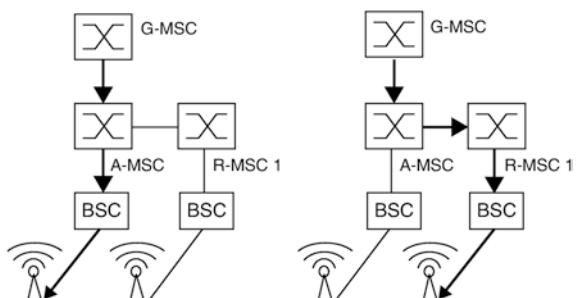
Beim Intra BSC Handover sind die aktuelle Zelle und die neue Zelle an der gleichen BSC angeschlossen. Diese Situation ist in Abb. 6.28 dargestellt.

Bei einem Wechsel in eine Zelle einer anderen BSC kann der Handover nicht durch die aktuelle BSC gesteuert werden, da keine direkte Signalisierungsverbindung zwischen den BSCs existiert. Deshalb beantragt die aktuelle BSC den Handover in die neue Zelle bei ihrer MSC über eine Handover Request-Nachricht. Teil dieser Nachricht ist die Cell-ID und der Location Area Code (LAC) der neuen Zelle. Da die MSC eine Liste aller LACs und Cell-IDs seiner Zellen hat, kann sie die dazugehörige BSC ermitteln, dort einen Sprachkanal in der gewünschten Zelle aufbauen und die BSC sowie die neue Zelle auf den Handover vorbereiten. Nachdem der Sprachkanal vorbereitet wurde, schickt die MSC ein Handover-Kommando zum Endgerät über die noch existierende alte Verbindung. Das Endgerät wechselt daraufhin in die neue Zelle. Erkennt die neue BTS und BSC den erfolgreichen Handover, wird dies der MSC mitgeteilt und die MSC kann den Sprachkanal auf die neue Verbindung umschalten. Danach wird der Sprachkanal in der alten BTS und BSC abgebaut, der Handover ist beendet (Abb. 6.43).

Noch aufwendiger wird es, wenn sich die neue Zelle nicht im Bereich der aktuellen MSC befindet. Aufgrund der Handover Request-Nachricht des aktuellen Base Station Controllers erkennt die MSC, dass sich die Location Area der neuen Zelle nicht in ihrem Versorgungsgebiet befindet. Über eine weitere Datenbank der MSC, die alle Location Areas der benachbarten MSCs enthält, kann die für diese Zelle verantwortliche MSC gefunden werden. Die aktuelle MSC wird in diesem Szenario auch als Anchor MSC (A-MSC) bezeichnet, die für die neue Zelle zuständige MSC wird Relay MSC (R-MSC) genannt.

Um den Handover durchzuführen, schickt die Anchor MSC der ermittelten Relay MSC eine Handover Nachricht über das MAP-Protokoll. Die Relay MSC baut daraufhin in der gewünschten Zelle einen Sprachkanal für den Handover auf und meldet dies der Anchor MSC. Die Anchor MSC leitet daraufhin den Handover durch Senden einer Handover Command-Nachricht an das Endgerät ein.

Abb. 6.43 Inter-MSC Handover



Nach erfolgreichem Handover in die neue Zelle meldet die Relay MSC der Anchor MSC den erfolgreichen Handover. Diese kann daraufhin den Sprachkanal zur Relay MSC durchstellen. Danach wird der Sprachkanal zur alten Zelle abgebaut.

Wechselt ein Teilnehmer nach einem Inter-MSC Handover in eine Zelle, die von einem dritten MSC verwaltet wird, spricht man von einem Subsequent Inter-MSC Handover (Abb. 6.44).

Für diesen Fall meldet die aktuelle Relay MSC (R-MSC 1) der Anchor MSC, dass ein Subsequent Inter MSC Handover zu einer anderen Relay MSC (R-MSC 2) notwendig ist. Die Anchor MSC beauftragt dann R-MSC 2 mit dem Aufbau der nötigen Ressourcen. Nachdem die neue Zelle vorbereitet wurde, schickt die Anchor MSC über R-MSC 1 den Handover-Befehl an das Endgerät. Dieses wechselt in die Zelle von R-MSC 2 und meldet den erfolgreichen Handover der Anchor MSC über die neue Verbindung. Diese kann dann R-MSC 1 anweisen, den nicht mehr benötigten Sprachkanal abzubauen. Auf diese Weise wird erreicht, dass es keine weitere Verkettung von MSCs gibt. An einem Gespräch sind somit immer nur die ursprüngliche Gateway MSC, die Anchor MSC und maximal eine Relay MSC beteiligt. Die Gateway und Anchor MSCs bleiben damit während des ganzen Gespräches unter Umständen die einzigen festen Komponenten.

Und schließlich gibt es auch noch den Fall, dass der Teilnehmer aus dem Gebiet der Relay MSC wieder in das Gebiet der Anchor MSC zurückkehrt. Nach einem solchen Handover ist die Anchor MSC neben der Gateway MSC wieder die einzige an der Verbindung beteiligte MSC. Da die Relay MSC das Gespräch wieder an die Anchor MSC zurückgibt, wird in diesem Fall von einem Subsequent Handback gesprochen. Weitere Details zu den unterschiedlichen Szenarien finden sich in 3GPP TS 23.009.

Aus Sicht des Endgeräts unterscheiden sich die vorgestellten Handovervarianten nicht, da die Handover-Nachricht für alle Fälle identisch ist.

Um einen Handover jedoch so schnell wie möglich durchzuführen, gibt es in GSM die Möglichkeit, Synchronisationsinformationen zwischen aktueller und neuer Zelle in der Handover-Nachricht zu übermitteln. Dies ermöglicht der Mobilstation, sofort auf den ihr zugeteilten Timeslot in der neuen Zelle zuzugreifen, statt sich zuerst auf die neue Zelle zu synchronisieren. Dazu müssen jedoch die aktuelle und neue Zelle synchronisiert

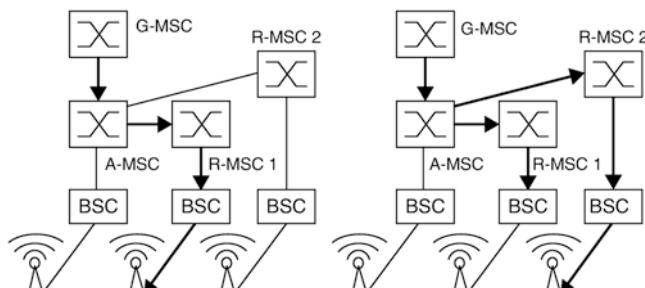


Abb. 6.44 Subsequent Inter-MSC Handover

sein, was z. B. bei einem Inter-MSC-Handover nicht möglich ist, da die zwei Zellen von unterschiedlichen MSCs und BSCs verwaltet werden. Da aber auch zwei Zellen, die mit der gleichen BSC verbunden sind, nicht unbedingt synchronisiert sein müssen, kann das Endgerät auch daran nicht erkennen, um welche Art Handover es sich im Netzwerk handelt.

6.9 Mobile Endgeräte

Durch die fortschreitende Miniaturisierung war es Mitte der 80er- Jahre erstmals möglich, alle für ein Mobiltelefon nötigen Bauteile in einem tragbaren Gerät unterzubringen. Wenige Jahre später konnte man Mobiltelefone dann soweit verkleinern, dass der limitierende Faktor für die Größe eines Mobiltelefons nicht mehr unbedingt die Größe der elektronischen Bauteile war. Vielmehr wurde die Größe eines Endgeräts hauptsächlich durch die notwendige Größe der Bedienteile wie Tastatur und Display bestimmt. Durch die ständige Weiterentwicklung und Miniaturisierung der elektronischen Bauteile ist es möglich, immer mehr Funktionalitäten und Bedienkomfort in ein Mobiltelefon zu integrieren. Wurden Mobiltelefone anfangs hauptsächlich zum Telefonieren verwendet, stehen heute mobile Multimediasgeräte „mit eingebauter Telefoniefunktion“ im Vordergrund. Nachfolgend wird deshalb zunächst die Architektur eines aus heutiger Sicht einfachen Sprach- und SMS Telefons beschrieben. Danach folgt dann ein Überblick über die Hardware und Funktionsweise eines komplexeren Smartphones.

6.9.1 Aufbau eines einfachen GSM Telefons

Einfache GSM Mobiltelefone für Sprach- und SMS Kommunikation haben eine Architektur, die heute mit nur sehr wenigen Bauteilen und somit entsprechend günstig produziert werden kann. Einfachste GSM Telefone sind deshalb schon bereits unter 20 € erhältlich.

Abb. 6.45 zeigt den grundsätzlichen Aufbau eines solchen Geräts mit dem Baseband-Prozessor als Kernkomponente, der eine RISC CPU und einen Digitalen Signalprozessor (DSP) enthält.

Der RISC-Prozessor kümmert sich dabei um:

- Die Verarbeitung der Informationen, die auf den Signialisierungskanälen (BCCH, PCH, AGCH, PCH, etc.) empfangen werden.
- Die Gesprächssignalisierung (DTAP)
- GPRS Management und GPRS-Daten
- Teile der Datenübertragungskette: Kanalkodierer, Interleaver, Cipherer (evtl. eigene Hardwareeinheit)

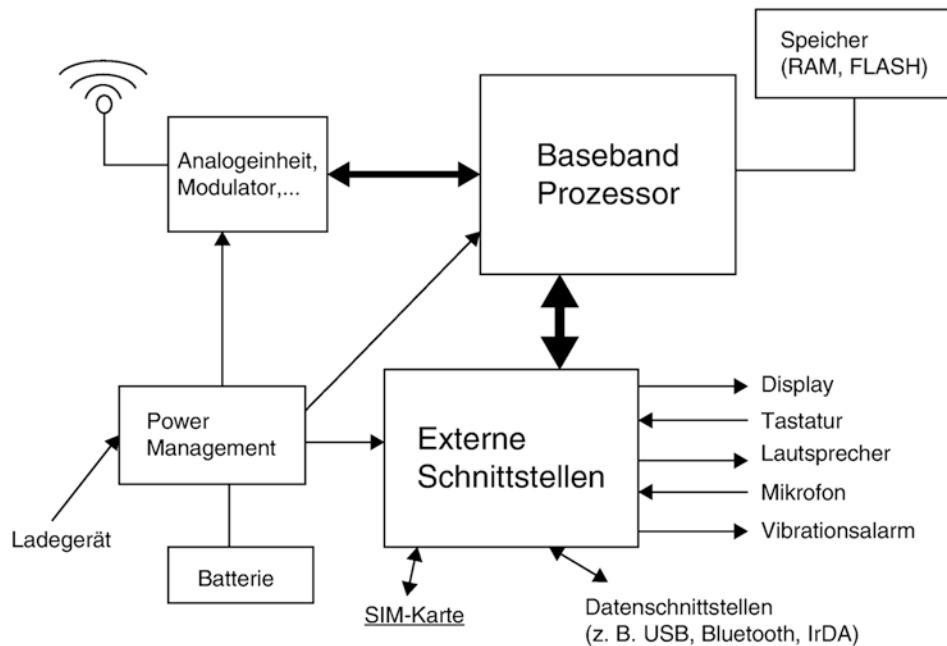


Abb. 6.45 Grundsätzlicher Aufbau eines Mobiltelefons

- Mobility Management (Netzwerksuche, Cell Reselection, Location Update, Handover, Timing Advance, etc.)
- Kommunikation mit externen Schnittstellen wie USB, Infrarot und Bluetooth.
- Userinterface (Tastatur, Display, Bedienungssoftware)

Da viele dieser Aufgaben gleichzeitig zu bearbeiten sind, kommt auf dem RISC-Prozessor ein echtzeitfähiges Embedded Multitasking Betriebssystem zum Einsatz. Die Echtzeitfähigkeit ist notwendig, da der Prozessor zur richtigen Zeit Daten für die Übertragung über die GSM-Rahmenstruktur zur Verfügung stellen und auch empfangen muss. Die restliche Peripherie wie Tastatur oder Display sowie das User Interface hat dagegen eine niedrigere Priorität.

Für den Baseband Prozessor wird heute hauptsächlich die ARM Reduced Instruction Set (RISC) Architektur verwendet, die Prozessorgeschwindigkeiten von wenigen Megaherz bis über 2 GHz bietet. Produziert werden solche Prozessoren von unterschiedlichen Herstellern. Für einfache Endgeräte kommen Ultra Low Power Versionen mit aus heutiger Sicht nur wenigen MHz Taktfrequenz zum Einsatz. Diese sind sehr leistungseffizient und benötigen im Ruhezustand und für das gelegentliche Abhören des Paging Kanals nur sehr wenig Energie. Somit können Standby-Zeiten von über einer Woche

erreicht werden. Auch die benötigte RAM und ROM Ausstattung eines solchen Gerätes mit nur wenigen hundert Kilobyte ist aus heutiger Sicht sehr gering.

Der Digitale Signalprozessor (DSP) ist ein weiterer wichtiger Bestandteil eines GSM Chipsatzes. Seine Hauptaufgabe ist die Sprachdatenkomprimierung mit den unterschiedlichen Sprachcodecs wie FR, EFR, HR oder AMR. Daneben wird er in Empfangsrichtung eingesetzt, um das empfangene Signal, das bereits digitalisiert wurde, vor der Dekodierung zu bearbeiten. Dazu verwendet der DSP die Trainingssequenz eines Bursts, die in Abschn. 6.7.3 vorgestellt wurde. Da dem DSP die Bits der Trainingssequenz bekannt sind, kann dieser einen Filter berechnen, der auf den restlichen Burst angewandt wird, um die darin enthaltenen Daten zu rekonstruieren.

Abb. 6.46 zeigt, welche Aufgaben der RISC-Prozessor und der DSP in einem Endgerät übernehmen. Vergleicht man die Bearbeitungskette des Sprachsignals im Mobiltelefon mit der im Netzwerk, stellt man fest, dass die Aufgabe der TRAU zum größten Teil vom DSP und den analogen Bauelementen im Endgerät übernommen werden. Alle anderen Bearbeitungsschritte wie die Kanalkodierung etc., die im Netzwerk von der BTS durchgeführt werden, finden ihr Gegenstück in der RISC CPU des Endgeräts.

6.9.2 Aufbau eines Smartphones

Während bei einem auf Sprache und SMS ausgerichteten GSM Telefon nur ein Prozessor für die Modemfunktionalität, das User Interface und die Grafikausgabe

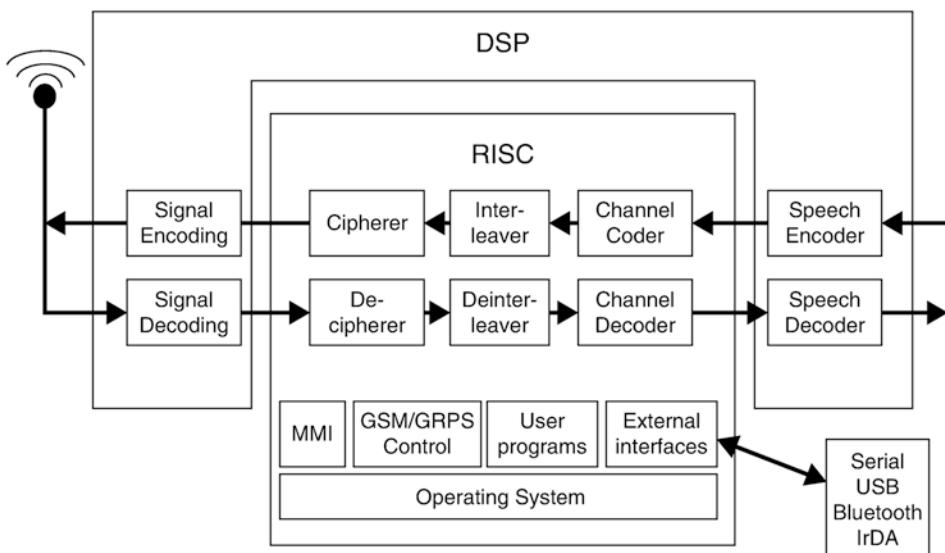


Abb. 6.46 RISC und DSP Funktionen im Überblick

zuständig ist, wurden diese Funktionalitäten bei Smartphones auf unterschiedliche Funktionsblöcke verteilt, die voneinander unabhängig arbeiten. Dies wurde nötig, da jeder Funktionsblock für sich stetig komplexer wurde und die Anforderungen an den Prozessor sich sehr unterschiedlich entwickelten. Außerdem kamen bei Smartphones wesentliche Funktionalitäten hinzu, die es bei einem einfachen Sprachtelefon nicht gibt. Abb. 6.47 gibt einen Überblick über die typischen Funktionsblöcke in einem Smartphone. Aufgrund zunehmender Miniaturisierung sind heute viele oder sogar alle Funktionen, die sich in der Abbildung im System On a Chip (SoC) Rahmen befinden, in einem einzigen Chip untergebracht.

Der Baseband Prozessor ist für die Kommunikation mit dem Mobilfunknetzwerk zuständig und beherrscht heute nicht nur GSM, sondern auch UMTS, LTE und oftmals auch 5G. Einige analoge Bauteile für den Mobilfunkteil können aufgrund ihrer Größe und Beschaffenheit nicht im SoC enthalten sein und sind deshalb in der Abbildung auch separat eingezeichnet. Das User Betriebssystem für die Benutzeroberfläche, wie z. B. Android oder iOS, wird im Applikationsprozessor ausgeführt, der üblicherweise aus ein oder mehreren ARM Prozessorkernen besteht. Diese ARM CPUs haben zwar einen ähnlichen Befehlssatz wie die Prozessoren in den zuvor beschriebenen GSM Telefonen, sind aber durch ihren komplexeren Aufbau und wesentlich höhere Taktung schneller und benötigen auch mehr Strom. Baseband Prozessor und Applikationsprozessor operieren

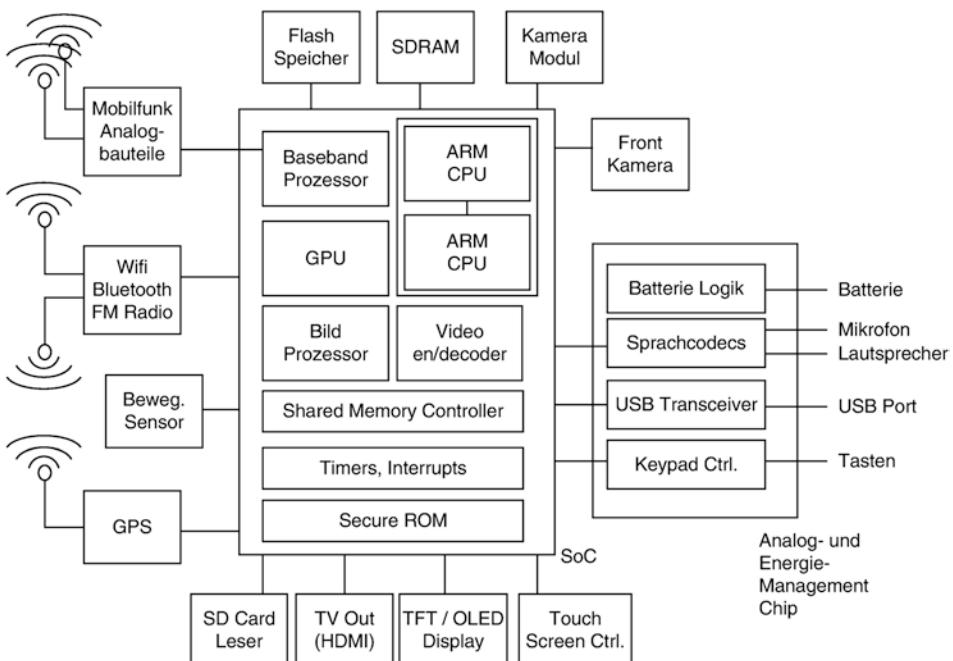


Abb. 6.47 Hardwarebaugruppen eines Smartphones

völlig getrennt voneinander und sind auch in einem SoC getrennte Funktionseinheiten, die über eine schnelle serielle Schnittstelle miteinander kommunizieren. Als dritte wichtige Funktionseinheit in einem Smartphone ist üblicherweise eine dedizierte Grafikeinheit vorhanden, die Graphics Processing Unit (GPU) genannt wird. Des Weiteren sind im SoC Chip noch weitere Hilfseinheiten für die Speicherverwaltung, für Timer und für Interrupts vorhanden, sowie dedizierte Hardware für das externe Kameramodul, um aufgenommene Bilder schnell für das User Betriebssystem zu verarbeiten. Außerdem werden an den SoC noch externe Speicherchips, sowie zahlreiche externe Input/Ouput Einheiten wie das Kamera Modul für Bilder, ein Kameramodul für Videotelefonie auf der Vorderseite, Wi-Fi, Bluetooth, GPS, SD Kartenleser, TV Out, das Display, der Touchscreen, etc., angeschlossen.

6.10 Die SIM-Karte

Trotz ihrer geringen Größe ist auch die SIM-Karte, offiziell Universal Integrated Circuit Card (UICC) genannt, ein wichtiger Bestandteil des GSM-Netzwerkes. Da sie alle Daten eines Teilnehmers enthält, kann der Teilnehmer mit seiner SIM-Karte jedes beliebige GSM-Endgerät verwenden. Ausnahmen sind Endgeräte mit SIM-Sperre, die nur mit SIM-Karten eines bestimmten Netzbetreibers funktionieren. Dies ist aber keine GSM-Einschränkung, sondern wurde von den Mobilfunkbetreibern eingeführt, um den Betrieb eines subventionierten Endgerätes nur mit eigenen SIM-Karten zu gestatten.

Die wichtigsten Informationen auf der SIM-Karte sind unter anderem die International Mobile Subscriber Identity (IMSI) des Teilnehmers, sowie dessen geheimer Schlüssel (Ki), der für die Authentifizierung und Generierung des Verschlüsselungskeys Kc benötigt wird.

Mit diversen im Internet kostenlos erhältlichen Tools können alle nicht lesegeschützten Informationen ausgelesen werden. Abb. 6.48 zeigt ein solches Tool. Sensitive Informationen, wie z. B. der geheime Schlüssel Ki können jedoch auch mit diesen Tools nicht ausgelesen werden.

Erstaunlicherweise ist eine SIM-Karte jedoch weit mehr als nur eine einfache Speicherkarte, denn sie enthält ein komplettes Mikrocontrollersystem, dessen Basisdaten in der folgenden Tabelle gezeigt werden:

CPU	8 oder 16 Bit CPU
Größe des ROM	40–100 kB
Größe des RAM	1–3 kB
EEPROM Größe	16–64 kB
Taktfrequenz	10 MHz, wird aus Mobiltelefontakt generiert
Betriebsspannung	1,8 V, 3 V oder 5 V. Moderne Endgeräte nutzen heute 1,8 V, unterstützen aber auch noch SIM Karten, die höhere Spannungen benötigen

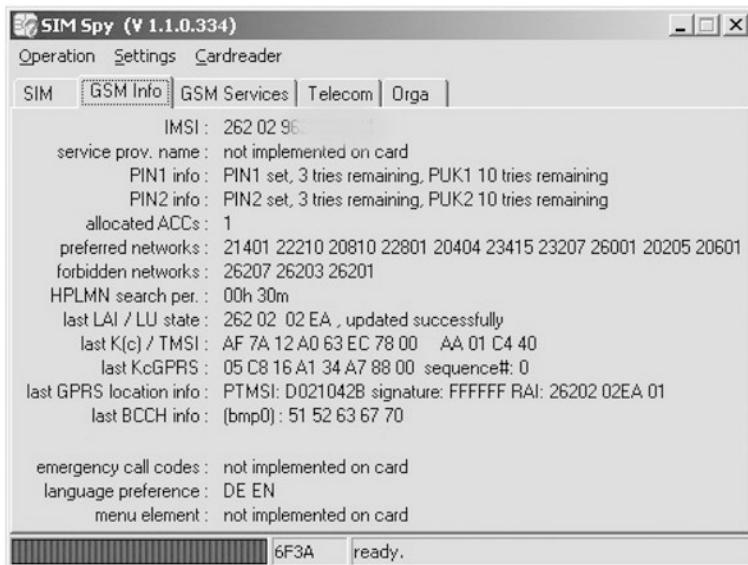


Abb. 6.48 Beispiel eines Tools zum Auslesen der Daten auf der SIM-Karte

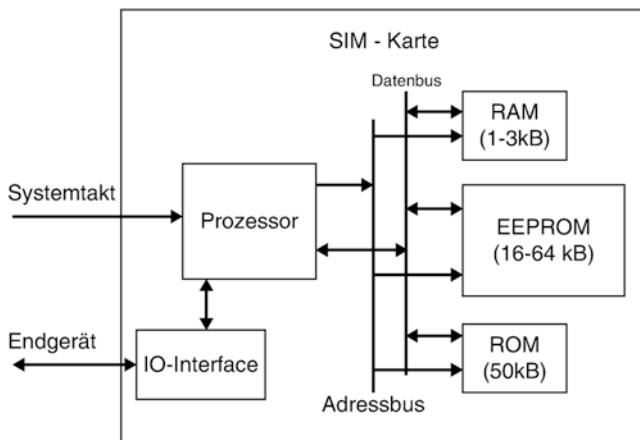


Abb. 6.49 Blockschaltbild der Komponenten einer SIM-Karte

Wie in Abb. 6.49 gezeigt wird, kann von extern nur über die CPU auf die nicht-flüchtigen Daten im EEPROM zugegriffen werden. Somit ist sichergestellt, dass von außerhalb kein direkter Zugriff auf die Daten erfolgen kann und somit sensitive Daten geschützt sind. Weiterhin wird der SIM-Prozessor verwendet, um die Signed Response (SRES) aus der Zufallszahl (RAND) zu generieren, die bei der Authentifizierung (vgl.

Abschn. 6.6.4) an das Mobiltelefon übermittelt wird. Die Berechnung von SRES muss zwingend in der SIM-Karte und nicht im Mobiltelefon durchgeführt werden, da sonst der geheime Schlüssel Ki an das Endgerät übergeben werden müsste. Könnte das Endgerät jedoch Ki auslesen, wäre dies auch mit anderen Geräten oder der in Abb. 6.48 gezeigten Software möglich und wäre somit ein großes Sicherheitsrisiko.

Außerdem kann der Microcontroller auf der SIM-Karte auch Programme ausführen, die vom Netzwerkbetreiber in die SIM-Karte übertragen wurden. Über die SIM Application Toolkit-Schnittstelle, die in der ETSI-Spezifikation 11.14 standardisiert ist, können diese Programme auch auf diverse Funktionen des Mobiltelefons zugreifen und z. B. auf Benutzereingaben reagieren oder Texte oder Menüs auf dem Display darstellen.

Während diese Funktionalität in der Vergangenheit durch Netzbetreiber für viele Dienste genutzt wurde, wird das SAT Interface heute hauptsächlich für Hintergrundaufgaben verwendet, wie z. B. dem Erkennen und Melden des Einlegens der SIM Karte in ein neues Endgerät und für das Übertragen von ‚stillen‘ SMS Nachrichten an die SIM Karte für den Update der Liste von bevorzugten Roamingnetzen.

Die Daten auf einer GSM-SIM-Karte werden aus logischer Sicht ähnlich wie bei einer Festplatte in Verzeichnissen und Dateien verwaltet. Die Datei- und Verzeichnisstruktur ist dabei fest vorgegeben und im ETSI Standard 11.11 spezifiziert. Das Hauptverzeichnis (Root Directory) wird darin unglücklicherweise Main File (MF) genannt, ein Unterverzeichnis wird als Dedicated File (DF) bezeichnet und eine normale Datei wird Elementary File (EF) genannt. Da auf der SIM-Karte nur wenig Speicherplatz zur Verfügung steht, haben die einzelnen Dateien und Verzeichnisse keine Datei- und Verzeichnisnamen, sondern nur 4-stellige Hex Nummern mit einer Länge von 2 Bytes. Diese wurden in der Spezifikation dann Namen gegeben, die jedoch nicht auf der SIM-Karte gespeichert sind. So wurde zum Beispiel dem Root Directory die ID 0x3F00 gegeben, dem GSM-Unterverzeichnis die ID 0x7F20 und der Datei, die die IMSI enthält, die ID 0x6F07. Um die IMSI auszulesen, muss das Endgerät somit auf folgenden Pfad zugreifen: \\0x3F00\\0x7F20\\0x6F07.

Um den Umgang mit Daten auf der SIM-Karte für das Endgerät so einfach wie möglich zu halten, kann jede Datei auf der SIM-Karte eine der folgenden drei Dateiformate haben:

- Transparent: Die Datei enthält nur eine Sequenz aus Bytes. Die Datei für die IMSI verwendet zum Beispiel dieses Format. Wie das Endgerät den Inhalt dieser Datei zu interpretieren hat, um die IMSI zu erhalten, ist wiederum im ETSI Standard 11.11 festgelegt.
- Linear Fixed: Diese Datei enthält Einträge (Records), die eine feste Länge besitzen. Dieses Format wird zum Beispiel für das Telefonbuch der SIM-Karte verwendet. Jeder Telefonbucheintrag ist dabei in einem Record der Telefonbuchdatei abgelegt.
- Cyclic: Ähnlich wie Linear Fixed, das Format enthält jedoch einen Zeiger auf den zuletzt geschriebenen Record. Ist das Ende der Datei erreicht wird der Zeiger auto-

matisch wieder auf den ersten Record gesetzt. Dieses Format wird z. B. für die Datei verwendet, die die zuletzt angerufenen Telefonnummern enthält.

Um Dateien zu schützen, ist jede Datei mit Zugriffsrechten ausgestattet. Dabei kann individuell kontrolliert werden, ob eine Datei gelesen oder geschrieben werden darf. Grundsätzlich ist der Zugriff auf die Dateien der SIM-Karte nur möglich, wenn sich der Teilnehmer zuvor per PIN authentifiziert hat. SIM-Karten mancher Netzbetreiber bieten jedoch die Möglichkeit, diesen Schutz zu deaktivieren, damit die PIN beim Einschalten des Endgeräts nicht eingegeben werden muss.

Nach Übergabe der PIN an die SIM-Karte kann dann das Lesen und Schreiben einzelner Dateien freigegeben oder gesperrt sein. So ist zum Beispiel trotz korrekter PIN das Lesen oder gar Schreiben der Datei für den geheimen Schlüssel Ki nicht möglich.

Neben der Dateistruktur der SIM-Karte legt die ETSI-Spezifikation 11.11 auch fest, wie mit der SIM-Karte kommuniziert wird. Auf Layer 2 wurden dazu Kommando- und Antwortnachrichten spezifiziert, die ganz allgemein als Application Protocol Data Units (APDU) bezeichnet werden. Sollen Daten zwischen einem Endgerät und einer SIM-Karte ausgetauscht werden, sendet das Endgerät eine Command APDU an die SIM-Karte. Diese muss darauf mit einer Response APDU antworten. Die SIM-Karte nimmt bei dieser Kommunikation eine passive Rolle ein, da sie nur Response APDUs schicken kann.

Sollen Daten gelesen werden, enthalten die Command APDUs unter anderem die Datei ID sowie die Anzahl der zu lesenden Bytes oder die Nummer des gewünschten Records. In Response APDUs werden dann die gewünschten Daten zurückgegeben.

Sollen Daten auf die SIM-Karte geschrieben werden, enthalten die Command APDUs neben der Datei ID die zu schreibenden Daten. In den Response APDUs befinden sich dann Statusmeldungen, ob der Schreibvorgang erfolgreich war.

Abb. 6.50 zeigt das Format einer Command APDU. Das erste Feld ist dabei die Class of Instruction und enthält bei GSM immer den Wert 0xA0. Das Instruction-Feld enthält die ID des Befehls, der von der SIM-Karte ausgeführt werden soll.

Die nachfolgende Tabelle zeigt einige Befehle und deren IDs. Die Felder P1 und P2 dienen zur Übergabe von Parametern für den gewählten Befehl. P3 gibt die Länge des nachfolgenden Datenfeldes an, das z. B. bei einem Schreibbefehl die zu schreibenden Daten enthält.

CLA	INS	P1	P2	P3	Data
-----	-----	----	----	----	------

Abb. 6.50 Command APDU



Abb. 6.51 Response APDU

Befehl	ID	P1	P2	Länge
SELECT (Datei öffnen)	A4	00	00	02
READ BINARY (Datei lesen)	B0	Offset High	Offset Low	Länge
UPDATE BINARY (Datei schreiben)	D6	Offset High	Offset Low	Länge
VERIFY CHV (PIN Eingabe)	20	00	ID	08
CHANGE CHV (PIN ändern)	24	00	ID	10
RUN GSM ALGORITHM (RAND, SRES, Kc ...)	88	00	00	10

Das Format einer Response PDU ist in Abb. 6.51 dargestellt. Neben einem Datenfeld enthält die Response APDU auch die Felder SW1 und SW2. Diese werden von der SIM-Karte verwendet, um dem Endgerät mitzuteilen, ob der zuvor gesendete Befehl korrekt ausgeführt werden konnte.

Um eine Datei für das Lesen oder Schreiben von Daten zu öffnen, sendet das Endgerät ein SELECT-Kommando an die SIM-Karte. Die SELECT APDU hat dabei den wie in Abb. 6.52 dargestellten Inhalt.

Als Antwort bekommt das Endgerät von der SIM-Karte eine Response APDU, die unter anderem folgende Datenfelder enthält:

Byte	Description	Länge
3–4	File Size	2
5–6	File ID	2
7	Type of File (Transparent, Linear Fixed, Cyclic)	1
9–11	Zugriffsberechtigungen	3
12	Dateistatus	1

Für eine vollständige Auflistung der zurückgegebenen Informationen siehe ETSI 11.11.

Im nächsten Schritt kann dann z. B. mit einer READ BINARY oder WRITE BINARY APDU die Datei gelesen oder modifiziert werden.

Um mit der SIM-Karte zu kommunizieren, hat diese auf ihrer Oberfläche 8 Kontaktstellen. Eine GSM-SIM-Karte verwendet davon jedoch nur 5 für folgende Zwecke:

- C1: Spannungsversorgung
- C2: Resetleitung
- C3: Takt (1–5 MHz)
- C5: Masse
- C7: Input/Output-Leitung

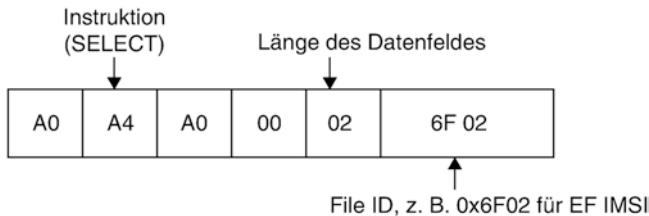


Abb. 6.52 Select Command

Da nur eine Leitung für Ein- und Ausgabe von Command und Status APDUs verwendet wird, erfolgt die Übertragung der Kommandos seriell und nur abwechselnd im Halbduplexverfahren. Die Taktgeschwindigkeit für die Datenübertragung ist dabei mit C3 Takt/372 definiert worden. Bei einem C3 Takt von 5 MHz beträgt somit die Übertragungsgeschwindigkeit 13.440 Bit/s.

6.11 Das Intelligent Network Subsystem und CAMEL

Alle bisher in diesem Kapitel beschriebenen Komponenten sind zwingend für den Betrieb eines Mobilfunknetzwerkes notwendig. Mobilfunkbetreiber bieten jedoch über die grundsätzliche Kommunikation hinaus zusätzliche Dienste an, für die zusätzliche Logik und Datenbanken notwendig sind. Dazu zählen insbesondere die Prepaid-Dienste. Diese erfreuen sich seit deren Einführung Mitte der 90er-Jahre großer Beliebtheit. Statt einmal im Monat eine Rechnung zu erhalten, besitzt ein Prepaid-Kunde ein Konto bei seinem Mobilfunkbetreiber, das er vorab aufladen kann. Für die aufgeladene Summe kann dann telefoniert werden. Während jedes Gespräches wird dabei der Kontostand laufend aktualisiert und nach dem Aufbrauchen des Verbindungsguthabens beendet. Weiterhin ist das Prepaid-System auch mit dem SMSC und dem paketvermittelnden Teil des Netzwerks (GPRS, UMTS, LTE) verbunden und kann somit auch für die sofortige Abrechnung von Kurznachrichten und für die Abrechnung von Datenverbindungen verwendet werden.

Diese und viele andere Dienste können mit Hilfe des Intelligent Network (IN) Subsystem gelöst werden. Die Logik und die entsprechenden Datenbanken befinden sich dabei auf einem Service Control Point (SCP), dessen grundsätzliche Funktionsweise schon am Anfang des Kapitels kurz vorgestellt wurde.

In den Anfangsjahren der GSM-Entwicklung wurde für diese Dienste in Ermangelung eines Standards auf herstellerspezifische Entwicklungen gesetzt. Großer Nachteil dieser Lösungen war jedoch, dass sie nur zwischen Komponenten des gleichen Herstellers verwendet werden konnten. Dies bedeutet, dass die Dienste im Ausland nicht funktionierten, wenn die Komponenten dort von anderen Herstellern stammten. Dies

war z. B. für den Prepaid-Dienst sehr ärgerlich, da Prepaid-Teilnehmer somit vom International Roaming ausgeschlossen waren.

Um die Interoperabilität zwischen Netzwerkkomponenten unterschiedlicher Hersteller und zwischen unterschiedlichen Mobilfunknetzen zu gewährleisten, wurde von 3GPP in TS 23.078 ein Protokoll und Verfahren spezifiziert, die den Namen CAMEL tragen. CAMEL steht dabei für ‚Customized Applications for Mobile network Enhanced Logic‘, ist aber in seinen Grundzügen deutlich einfacher, als sein Name suggeriert.

Während CAMEL auch Funktionalitäten für SMS und GPRS bietet, wird nachfolgend jedoch nur auf die grundsätzliche Funktionsweise für leitungsvermittelnde Verbindungen eingegangen.

CAMEL selbst ist keine Applikation oder Dienst, sondern die Grundlage, Dienste (Customized Applications) auf einem SCP zu entwickeln, die mit Netzwerkelementen anderer Hersteller national und international kompatibel sind. Diese Eigenschaft lässt sich z. B. mit dem http-Protokoll vergleichen. HTTP wird für die Übertragung von Web-Seiten zwischen einem Web Server und einem Web Client verwendet. Dabei stellt HTTP sicher, dass jeder beliebige Web Server mit jedem beliebigen Web Client Daten austauschen kann. Ob es sich bei den Daten nun um Web-Seiten oder Bilder handelt ist HTTP egal, denn die Interpretation ist Sache des Web Clients, bzw. Web Server.

CAMEL spezifiziert dazu im Wesentlichen das Protokoll zwischen den Netzwerkelementen wie MSC und SCP, sowie ein Zustandsmodell für einen Verbindungsablauf.

Dieses Zustandsmodell wird bei CAMEL Basic Call State Model (BCSM) genannt. Ein Verbindungsablauf wird dabei in eine Anzahl Zustände unterteilt. Für den Anrufer (Originator BCSM) gibt es unter anderem folgende Zustände:

- Anrufaufbau
- Analyse der Zielrufnummer
- Routing der Verbindung
- Benachrichtigen des Ziels (Alerting)
- Gespräch läuft (Active)
- Beenden der Verbindung (Disconnect)
- Keine Antwort des Zielteilnehmers
- Zielteilnehmer besetzt

Auch für einen angerufenen Teilnehmer (Terminator) gibt es ein Zustandsmodell, das entsprechend T-BCSM genannt wird. Das T-BCSM wird z. B. für Prepaid-Teilnehmer im Ausland benötigt, um die Weiterleitung des Gesprächs ins Ausland steuern und abrechnen zu können (Abb. 6.53).

Beim Übergang zwischen den Zuständen definiert CAMEL Detection Points (DPs). Ist ein Detection Point für einen Teilnehmer aktiviert, wird der SCP über den Zustandsübergang informiert. Teil dieser Nachricht sind die IMSI des Anrufers, seine aktuelle Position (Cell ID), Zielrufnummer und vieles mehr. Ob ein DP für einen Teilnehmer

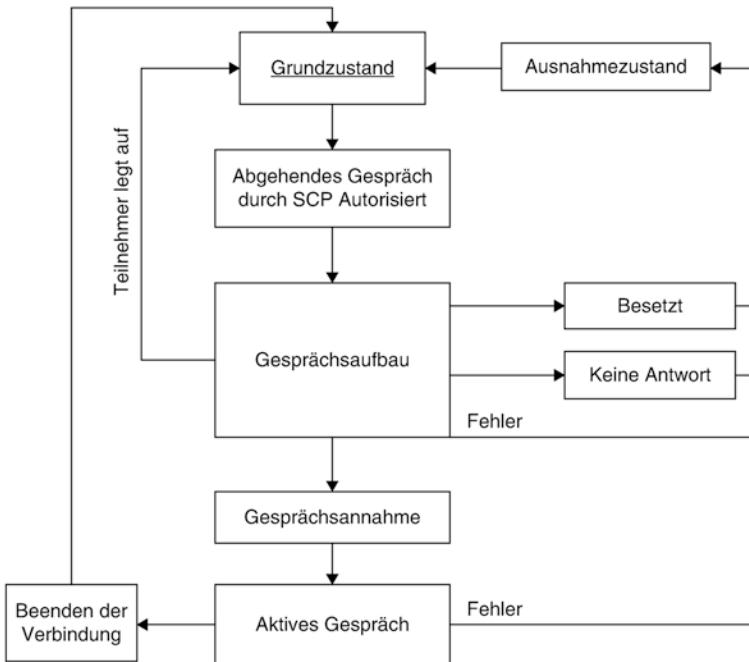


Abb. 6.53 Vereinfachtes Anrufer-Zustandsmodell (O-BCSM) nach 3GPP TS 23.078

aktiviert ist, wird im HLR für jeden Teilnehmer individuell eingetragen. Der SCP hat dann bei Empfang einer solchen Nachricht aufgrund der enthaltenen Daten die Möglichkeit, den weiteren Ablauf des Gesprächs zu beeinflussen. Der SCP hat zum Beispiel die Möglichkeit, das Gespräch zu beenden, die Zielrufnummer zu ändern oder Informationen an die MSC zurückzugeben, die in den Billing Record aufgenommen werden und somit später Einfluss auf die Gesprächsabrechnung haben.

Für einen Prepaid-Dienst kann das Zustandsmodell und das CAMEL-Protokoll zwischen MSC und SCP wie folgt verwendet werden:

Ein Teilnehmer möchte ein Gespräch aufbauen. Die MSC stellt am Anfang des Gesprächsaufbaus fest, dass der Detection Point ‚Authorize Origination‘ in dessen HLR-Eintrag gesetzt ist und sendet daraufhin eine Nachricht zum SCP. Anhand der darin enthaltenen IMSI und der gewünschten CAMEL-Dienstnummer erkennt der SCP, dass es sich um einen Prepaid-Teilnehmer handelt. Mit der übergebenen Zielrufnummer, der aktuellen Uhrzeit, etc., ermittelt der SCP dann den Minutenpreis für das Gespräch. Hat der Teilnehmer noch genug Guthaben auf seinem Konto, gestattet der SCP den Gesprächsaufbau und teilt der MSC mit, für wie viele Minuten diese Freigabe Gültigkeit hat. Die MSC verbindet daraufhin das Gespräch. Nach Ende des Gesprächs schickt

die MSC erneut eine Nachricht zum SCP und teilt ihm die Dauer des Gesprächs mit. Der SCP aktualisiert daraufhin das Guthaben des Teilnehmers entsprechend.

Läuft die vom SCP übergebene Zeit während des Gesprächs ab, benachrichtigt die MSC wiederum den SCP. Dieser hat dann die Möglichkeit, der MSC eine weitere Zeitspanne für die Weiterführung des Gesprächs zu übergeben. Der SCP kann die MSC jedoch auch anweisen, das Gespräch zu beenden oder einen Ton oder eine Ansage einzuspielen. Im Prepaid-Fall kann dieser Ton zum Beispiel ein Hinweis sein, dass das Guthaben fast erschöpft ist.

6.12 Fragen und Aufgaben

1. Mit welchem Verfahren und typischen Übertragungsgeschwindigkeiten wurden Sprachdaten in einem leitungsvermittelten Netzwerk übertragen?
2. Welche wichtigen Komponenten gibt es im GSM-Network Subsystem (NSS) und welche Aufgaben erfüllen sie?
3. Welche wichtigen Komponenten gibt es im GSM-Radionetzwerk (BSS) und welche Aufgaben erfüllen sie?
4. Mit welchen Verfahren kann eine BTS gleichzeitig mit mehreren Teilnehmern kommunizieren?
5. Welche Verarbeitungsschritte durchläuft die menschliche Sprache in einem Mobiltelefon, bevor sie über die GSM-Luftschnittstelle versandt werden kann?
6. Was ist ein Handover und welche Komponenten können daran beteiligt sein?
7. Wie wird bei einem eingehenden Gespräch der aktuelle Aufenthaltsort eines Teilnehmers ermittelt und wie wird das Gespräch im Netzwerk zugestellt?
8. Wie wird eine SMS-Nachricht zwischen zwei Teilnehmern ausgetauscht?
9. Wie wird ein Teilnehmer im GSM-Netzwerk authentifiziert? Warum ist eine Authentifizierung notwendig?
10. Welche Aufgaben haben der RISC-Prozessor und der DSP in einem einfachen GSM Endgerät?
11. Wie werden Daten auf einer SIM-Karte abgelegt?
12. Was ist CAMEL und für welche Dienste wird es verwendet?

Lösungen sind auf der Website zum Buch unter <http://www.cm-networks.de> zu finden.



GPRS und EDGE

7

Mitte der 80er- Jahre war die Sprachübertragung die wichtigste Anwendung für drahtgebundene und mobile Netzwerke. Aus diesem Grund wurde das GSM-Netz auch hauptsächlich für die Sprachübertragung konzipiert und optimiert. Seit Mitte der 90er-Jahre spielt jedoch das Internet und somit die Datenübertragung eine immer größere Rolle. GPRS, der General Packet Radio Service, erweiterte den GSM-Standard für eine effiziente Datenübertragung und ermöglichte somit mobilen Geräten den Zugriff auf das Internet. Mit der Enhanced Data Rates for GSM Evolution (EDGE)-Erweiterung wurden die möglichen Datenraten dann weiter erhöht und die Verzögerungszeit bei der Datenübertragung deutlich verbessert. Während GPRS und auch EDGE heute aufgrund der gestiegenen Geschwindigkeitsanforderungen an Bedeutung verloren hat, werden diese Netze weiterhin von Machine-To-Machine (M2M) Modulen im industriellen Umfeld für den Datenaustausch verwendet. Hier spielt oftmals der Preis des Datenmoduls eine größere Rolle als die Geschwindigkeit der Datenübertragung. Mit GPRS wurden zudem die Grundlagen für die paketvermittelte Datenübertragung in modernen Mobilfunknetzwerken gelegt, die auch bei UMTS, LTE und auch 5G noch Anwendung finden. Somit lohnt sich auch heute noch ein Blick auf dieses System. Im ersten Teil dieses Kapitels werden die Vor- und Nachteile von GPRS gegenüber der in den 1990er Jahren üblichen GSM-Datenübertragung und der Datenübertragung in leitungsvermittelten Netzen erläutert. Teil zwei des Kapitels beschreibt dann, wie GPRS und EDGE standardisiert und in der Praxis implementiert wurden.

7.1 Leitungsvermittelte Datenübertragung

Da das GSM-Netzwerk ursprünglich als leitungsvermittelndes Netzwerk konzipiert wurde, wird für eine herkömmliche Sprach- oder Datenverbindung zwischen zwei Teilnehmern ein exklusiver Kanal geschaltet. Dieser kann während der Verbindung nur von den zwei miteinander verbundenen Teilnehmern verwendet werden (Abb. 7.1).

Dieser exklusive Kanal hat eine konstante Bandbreite und eine konstante Verzögerungszeit. Für den Anwender hat dies eine Reihe von Vorteilen:

Nach Aufbau der Verbindung können Daten in beide Richtungen ohne weitere Signalisierungsinformationen für das Weiterleiten der Daten (Routing) gesendet werden. Da die Verbindung fest geschaltet ist, leitet jede Komponente im Netzwerk die Daten über den für die Verbindung reservierten Kanal transparent an die nächste Komponente weiter.

Der zugeteilte Kanal hat eine konstante Bandbreite, die Geschwindigkeit der Datenübertragung variiert also nicht. Dies ist besonders für die Sprachdatenübertragung wichtig, da hier die Daten nicht in Netzwerkelementen zwischengepuffert werden sollten oder gar in Überlastsituationen verworfen werden dürfen.

Eine weitere wichtige Eigenschaft einer leitungsvermittelten Verbindung für die Sprachübertragung ist die konstante Verzögerungszeit. Die Verzögerungszeit ist dabei die Zeit zwischen dem Senden und Empfangen eines Bits oder eines Datenblocks. Ohne eine konstante Verzögerungszeit müsste beim Empfänger ein Empfangspuffer vorhanden sein, der die schwankenden Verzögerungszeiten ausgleicht. Dies ist vor allem für Sprache sehr unerwünscht, da diese so schnell wie möglich am anderen Ende empfangen und wiedergegeben werden soll.

Während die Leitungsvermittlung ideal für die Sprachübertragung geeignet ist, gibt es jedoch einen großen Nachteil für die Datenübertragung mit variablen Übertragungsraten:

Das Webbrowsern im Internet ist eine typische Datenanwendung mit variablen Übertragungsraten. Während der Anforderung einer Webseite sollte dem Anwender eine möglichst große Bandbreite zur Verfügung stehen, um die Webseite möglichst schnell zu empfangen. Während des anschließenden Lesens der Webseite werden dann für einige Zeit keine Daten übertragen. Da bei einer leitungsvermittelten Verbindung die Bandbreite weder erhöht noch während des Lesens wieder freigegeben werden kann, ist die Leitungsvermittlung für diese Art der Datenübertragung nicht ideal geeignet. Vor allem

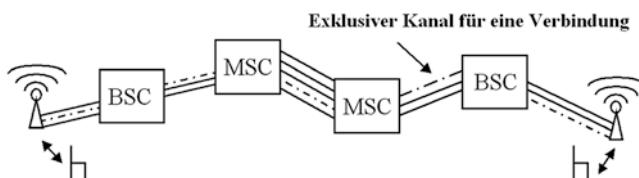


Abb. 7.1 Exklusive Verbindung bei der Leitungsvermittlung

die ungenutzte Bandbreite während des Lesens der Webseite ist für ein Mobilfunknetzwerk problematisch, da auf der Luftschnittstelle die Übertragungskapazität sehr begrenzt ist.

7.2 Paketorientierte Datenübertragung

Für Anwendungen wie dem Webbrowsen ist es viel effizienter, wenn der Übertragungskanal nur während der eigentlichen Übertragung der Daten für einen Teilnehmer verwendet wird und danach wieder für andere freigegeben wird. Um dies zu erreichen, wird bei der paketorientierten Datenübertragung der Übertragungskanal nicht mehr in kleinere Kanäle für einzelne Benutzer aufgeteilt und fest zugeordnet. Stattdessen werden die Daten der unterschiedlichen Benutzer in Datenpaketen nacheinander über den Übertragungskanal gesendet. Zwar kann zu einer Zeit nur ein Teilnehmer senden oder empfangen, die Datenpakete werden dafür aber schneller übertragen, da die Bandbreite des gesamten Übertragungskanals zur Verfügung steht. Da bei dieser Übertragung nicht Leitungen vermittelt werden, sondern einzelne Pakete, wird diese Art der Datenübertragung Paketvermittlung oder Packet Switching genannt (Abb. 7.2).

Da es bei der Paketvermittlung keine festen Kanäle gibt, muss jedes Paket eine Information über Absender (Source) und Empfänger (Destination) enthalten. Die Empfängeradresse, auch Zieladresse genannt, wird dann innerhalb des Netzwerkes für die Weiterleitung der Datenpakete an den richtigen Empfänger verwendet. Auf diese Weise wird z. B. auch eine Webseite im Internet übertragen. Die Webseite wird dazu vom Webserver (Sender) in viele IP-Pakete aufgeteilt und danach zum Webbrowser (Empfänger) übertragen.

Die paketorientierte Übertragung hat außerdem den Vorteil, dass ein Webbrowser auch Webseiten von verschiedenen Servern empfangen kann, ohne dafür wie bei der Leitungsvermittlung mehrere physische Verbindungen (Leitungen) explizit nacheinander aufzubauen.

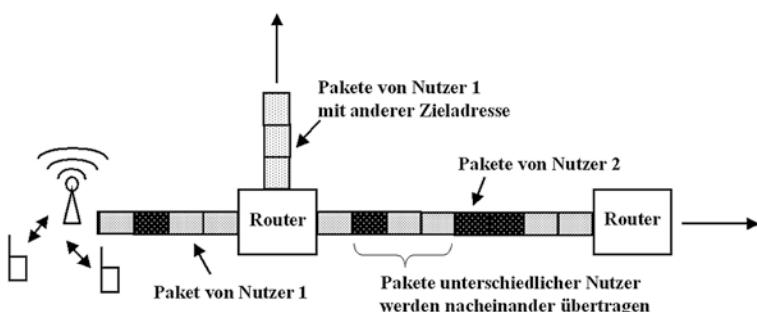


Abb. 7.2 Paketorientierte Datenübertragung

Um die Paketdatenübertragung auch in GSM-Netzwerken zu ermöglichen, wurde der General Packet Radio Service (GPRS) entwickelt. Dabei wurde besonderen Wert darauf gelegt, dass für GPRS keine neuen Basisstationen (BTS'en) notwendig sind. Dies war eine wichtige Voraussetzung, um die paketorientierte Datenübertragung am Anfang der 2000er Jahre kostengünstig in bereits existierenden Netzwerken einzuführen.

GPRS bietet durch seinen paketorientierten Ansatz für mobile Applikationen mit dynamischer Bandbreitennutzung außerdem folgende Vorteile gegenüber der Leitungsvermittlung:

Flexible Zuteilung der Bandbreite auf der Luftschnittstelle: Da mehr als nur ein Zeitschlitz pro Teilnehmer zugeteilt werden kann, übertrifft die GPRS-Übertragungsgeschwindigkeit die eines leitungsvermittelten Kanals von 9.6 oder 14.4 kbit/s bei weitem. GPRS bietet eine Übertragungsgeschwindigkeit von theoretisch 170 kbit/s, in der Praxis wurden jedoch nur Geschwindigkeiten von etwa 80 kbit/s erreicht. Dies entsprach etwa der Geschwindigkeit von Festnetzmodems, die zum Zeitpunkt der GPRS Einführung weit verbreitet waren.

Mit EDGE (Enhanced Data Rates for GSM Evolution), das den GPRS-Standard unter anderem um eine zusätzliche Modulationsart erweitert, kann die Übertragungsgeschwindigkeit auf bis zu 270 kbit/s gesteigert werden. In der Praxis liegt die erreichbare Geschwindigkeit jedoch heute deutlich niedriger, da nur noch sehr wenig Spektrum für GSM und GPRS/EDGE verwendet wird und viele Geräte gleichzeitig auf das Netzwerk zugreifen. Da EDGE auch Neuerungen für den leitungsvermittelten Teil des Netzwerkes brachte, werden die GPRS-Erweiterungen als EGPRS bezeichnet. Im täglichen Umgang dominiert jedoch die Abkürzung EDGE (Abb. 7.3).

Bei GPRS wird ausschließlich nach Datenvolumen statt Onlinezeit abgerechnet. Großer Vorteil hierbei ist, dass z. B. beim Websurfen „nur“ für das übertragene Datenvolumen bezahlt werden muss und nicht für die Zeit, in der die Webseite gelesen wird. Da während des Lesens keine Daten übertragen werden, kann die frei gewordene Bandbreite für andere Nutzer verwendet werden. Dies ist bei der Leitungsvermittlung grundsätzlich nicht möglich (Abb. 7.4).

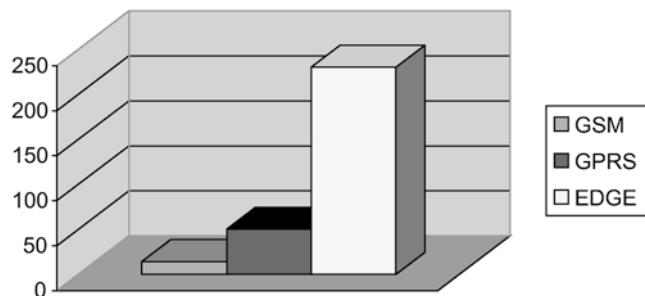


Abb. 7.3 Geschwindigkeit von GSM, GPRS und EDGE in kbit/s

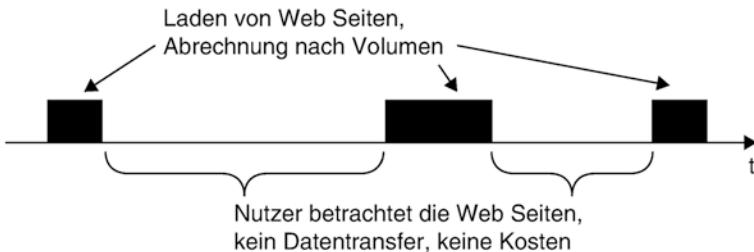


Abb. 7.4 Abrechnung nach Volumen und nicht nach Onlinezeit

GPRS reduziert die Zeit für die Interneteinwahl erheblich. Während eine GSM-leitungsvermittelte Internetverbindung ähnlich einer analogen Modemverbindung im Festnetz bis zu 20 s für den Verbindungsauftakt benötigte, kann eine GPRS-Verbindung in weniger als 5 s aufgebaut werden.

Da der Benutzer nicht für die Zeit bezahlt, in der keine Daten übertragen werden, muss die Internetverbindung auch bei langen Übertragungspausen nicht abgebaut werden. Dieser „Always On“-Modus ermöglichte erstmals die mobile Nutzung vieler Anwendungen wie z. B. von eMail-Programmen, die automatisch neue eMail-Nachrichten empfangen oder Mobile Messaging Clients, die ständig auf neue Nachrichten warten können.

Während einer Zug- oder Autofahrt kommt es auch heute noch mehr oder weniger regelmäßig oft zu schlechtem Empfang oder sogar zu Empfangsverlust. In solchen Fällen brechen leitungsvermittelte Internetverbindungen ab und müssen vom Anwender erneut aufgebaut werden. GPRS-Verbindungen dagegen werden bei Empfangsverlust nicht abgebrochen, da die logische Verbindung unabhängig von der Verfügbarkeit der physischen Verbindung weiterhin besteht. Wurden bei Empfangsverlust gerade Daten übertragen, kann der Transfer sofort wieder aufgenommen werden, sobald das Endgerät eine neue Zelle des Netzwerks entdeckt. Da jedoch beim Webbrowsern die meiste Zeit für das Lesen der Webseiten verwendet wird und somit die meiste Zeit keine Daten übertragen werden, bemerkt ein Anwender einen Empfangsverlust oft gar nicht.

GPRS wurde ursprünglich für die Übertragung von verschiedenen paketorientierten Protokollen entworfen. Mit dem großen Erfolg des Internets, das ausschließlich auf dem paketorientierten Internet Protokoll (IP) basiert, ist IP auch das einzige Protokoll, das GPRS heute unterstützt. Deshalb werden Begriffe wie „Nutzdatenübertragung“, „paketorientierte Datenübertragung“ oder „packet switching“ in diesem Kapitel als Synonyme für „Übertragung von IP-Paketen“ verwendet.

7.3 GPRS auf der Luftschnittstelle

Nach einem Überblick im letzten Abschnitt über die grundsätzliche Funktionsweise von GPRS zeigt dieser Abschnitt nun, welche Verfahren die paketdatenorientierte Datenübertragung über die Luftschnittstelle des GSM-Mobilfunknetzwerkes ermöglichen.

7.3.1 GPRS Timeslot-Nutzung im Vergleich zu GSM

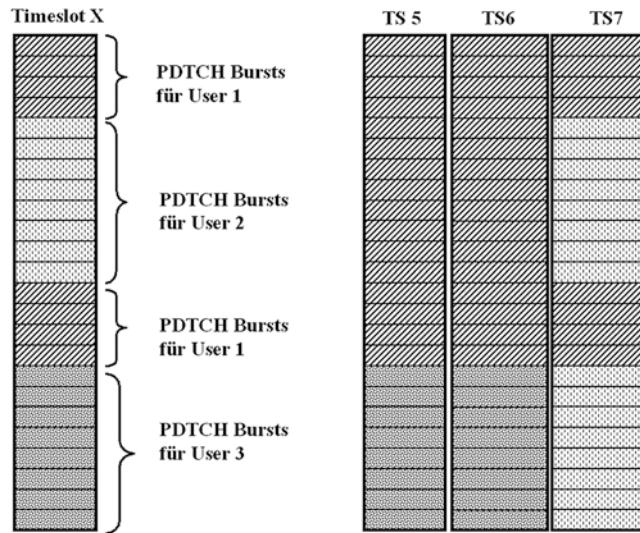
Wie im ersten Kapitel gezeigt wurde, verwendet GSM-Zeitschlüsse (Timeslots) auf der Luftschnittstelle für die Kommunikation mit mehreren Teilnehmern. Während einer leitungsvermittelten Verbindung bekommt ein Teilnehmer einen logischen Traffic Channel (TCH) fest zugeteilt, der auf einem physischen Timeslot übertragen wird. Dieser kann nicht für andere Teilnehmer verwendet werden, auch wenn darauf für eine gewisse Zeit keine Daten übertragen werden.

GPRS ist bei der Zuteilung der physischen Ressourcen sehr viel flexibler. Die kleinste physische Ressource, die bei GPRS einem Teilnehmer zugeteilt werden kann, ist ein Block, der aus 4 Bursts eines Packet Data Traffic Channel (PDTCH) besteht. Ein logischer PDTCH ist dem logischen TCH sehr ähnlich, da auch er auf einem physischen Timeslot übertragen wird. Möchte ein Teilnehmer weitere Daten übertragen, kann das Netzwerk auch die nachfolgenden Blöcke des PDTCH dem Teilnehmer zuweisen. Die nachfolgenden Blöcke können aber auch an andere Teilnehmer vergeben oder für Kontrollinformationen verwendet werden. Abb. 7.5 zeigt auf der linken Seite, wie die Blöcke eines PDTCH unterschiedlichen Teilnehmern dynamisch zugewiesen werden können. Die Darstellungsweise entspricht dabei dem Prinzip aus Abb. 6.25.

GPRS verwendet ähnlich wie GSM eine Rahmenstruktur für die Anordnung der Frames. Statt jedoch 26 bzw. 51 Frames zu einem Multiframe zu gruppieren, werden bei GPRS 52 Frames zu einem Multiframe zusammengefasst. In einem GPRS 52 Multiframe werden fast alle Frames für den logischen PDTCH sowie für Kontrollinformationen verwendet. Ausnahmen bilden lediglich Frame 24 und 51, die von aktiven Endgeräten für die Pegelmessung von Nachbarzellen verwendet werden. Frame 12 und 38 schließlich werden für Timing Advance-Berechnungen verwendet. Details über den PDTCH und andere logische GPRS-Kanäle werden in Abschn. 7.3.7 beschrieben.

Um die Übertragungsgeschwindigkeit eines Teilnehmers zu steigern, können wie im rechten Teil von Abb. 7.5 gezeigt, mehrere Zeitschlüsse gleichzeitig verwendet werden. Sind Daten für einen Teilnehmer zu übertragen, entscheidet das Netzwerk anhand der verfügbaren Timeslots und technischen Möglichkeiten des Endgeräts, wie viele Timeslots verwendet werden können. Dieses Verfahren wird Multislot-Datenübertragung oder auch Timeslot Aggregation genannt.

Die Anzahl der Timeslots, die einem Endgerät gleichzeitig zugeteilt werden konnten, war in den Anfangszeiten von GPRS hauptsächlich von dessen technischen Eigenschaften wie z. B. der Prozessorgeschwindigkeit abhängig. Aus diesem Grund wurden

**Abb. 7.5** PDTCH Vergabe und Timeslot Aggregation

Multislot-Klasse	Max. Timeslots		
	Downlink	Uplink	Summe
8	4	1	5
10	4	2	5
12	4	4	5
32	5	3	6

Abb. 7.6 Beispiele für GPRS Multislot-Klassen aus 3GPP TS 45.002, Annex B.1

Endgeräte deshalb in unterschiedliche Multislot-Klassen eingeteilt, von denen die wichtigsten in Abb. 7.6 dargestellt sind.

In Abhängigkeit der Multislot-Klasse können dann 3, 4 oder mehr Timeslots für die Datenübertragung vom Netzwerk an ein Endgerät zugewiesen werden. Die meisten heute erhältlichen Endgeräte unterstützen die Multislot-Klassen 10, 12 oder 32. Geräte der Klasse 10 können beispielsweise bis zu 4 Timeslots im Downlink und bis zu 2 Timeslots im Uplink bündeln. Das bedeutet, dass die maximale Übertragungsgeschwindigkeit im Uplink wesentlich geringer als im Downlink ist. Für viele Anwendungen ist dies kein Problem, da die Datenmenge die zum Endgerät übertragen wird meistens größer ist als in der umgekehrten Richtung. Für Anwendungen die größere Datenmengen versenden wären jedoch mehr Timeslots in Uplink-Richtung wünschenswert. Bessere Endgeräte unterstützen deshalb heute auch Multislot-Klasse 32, mit der 5 Timeslots

im Downlink und 3 Timeslots im Uplink gebündelt werden können. Gleichzeitig nutzbar sind 6 Timeslots. Das Netzwerk kann somit dynamisch folgende Downlink+Uplink Kombinationen wählen: 5+1, 4+2 oder 3+3.

In Multislot-Klasse 10 ist die Summe der gleichzeitig nutzbaren Timeslots im Uplink und Downlink zusammen maximal 5. Sind 4 Timeslots im Downlink zugeteilt, kann das Endgerät nur einen Timeslot für die Datenübertragung im Uplink verwenden. Bemerkt das Netzwerk, dass auch Daten in Uplink-Richtung zu übertragen sind, wird die Timeslot-Zuteilung allerdings automatisch neu konfiguriert. Das Endgerät bekommt dann 3 Timeslots in Downlink-Richtung und 2 Timeslots in Uplink-Richtung zugewiesen. Beendet das Endgerät die Übertragung im Uplink und werden weiterhin Daten im Downlink übertragen, ändert das Netzwerk erneut die Zuweisung, und es werden wieder 4 Timeslots im Downlink zugewiesen.

Damit das Netzwerk für jeden Teilnehmer die richtige Anzahl Timeslots zuweisen kann, teilt das Endgerät bei Anforderung eines Uplinkkanals dem Netzwerk seine Mobile Station Radio Access Capabilities mit. Teil dieser Information ist auch die unterstützte Multislot-Klasse. Im Netzwerk wird die Multislot-Klasse des Endgeräts gespeichert und kann somit später wieder verwendet werden, wenn Ressourcen im Downlink zugeteilt werden.

7.3.2 Gleichzeitige Nutzung einer Basisstation von GSM und GPRS

Da GPRS als Erweiterung eines GSM-Netzwerks entwickelt wurde, teilen sich GSM und GPRS-Netz auch die 8 Timeslots pro Frequenz einer Basisstation (BTS). Die maximale GPRS-Datenrate hängt also auch davon ab, wie viele Timeslots einer BTS für Sprachkanäle verwendet werden. Es bleibt dabei dem Netzwerkbetreiber überlassen, wie viele Timeslots für welchen Dienst verwendet werden. Auch eine dynamische Zuordnung ist möglich. Bei geringem Sprachaufkommen können viele Timeslots für GPRS verwendet werden, die aber jederzeit für die Sprachübertragung dem GPRS-Netzwerk entzogen werden können. In der Praxis wird üblicherweise eine Mischkonfiguration verwendet. Das bedeutet, dass für GPRS eine gewisse Anzahl an Timeslots zur Verfügung steht, die nicht für Sprachkanäle verwendet werden dürfen. Neben diesen festen Timeslots werden weitere Timeslots dynamisch zugeteilt und können je nach Verkehrsaufkommen entweder für GPRS oder für leitungsvermittelte Verbindungen verwendet werden. So ist sichergestellt, dass trotz hohem Sprachverkehrsaufkommen zu manchen Zeiten weiterhin ein Datentransfer mit GPRS möglich ist.

Abb. 7.7 zeigt ein Beispiel für eine gemischte GSM/GPRS Konfiguration einer Zelle. Wie in Abschn. 6.7.3 gezeigt wurde, besteht in der Praxis eine BTS meist aus mehreren Zellen, in denen zur Kapazitätssteigerung mehrere Frequenzen verwendet werden.

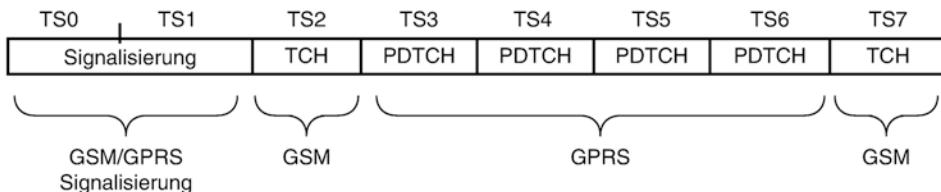


Abb. 7.7 Gemeinsame Nutzung einer Zelle von GSM und GPRS

7.3.3 Coding Schemes

Eine weitere Möglichkeit die Übertragungsgeschwindigkeit der Teilnehmer zu steigern, ist die Anpassung der Bits pro Block für die Fehlerkorrektur an die jeweiligen Übertragungsbedingungen. Zu diesem Zweck wurden in GPRS vier Kodierungsverfahren (Coding Schemes) mit einem unterschiedlichen Verhältnis von Nutzdatenbits zu Fehlerkorrekturbits definiert. Bei schlechten Übertragungsbedingungen kann mit Coding Scheme 1 oder 2 eine Nettodatengeschwindigkeit von 8 bzw. 12 kbit/s pro Timeslot erreicht werden. Bei guten Übertragungsbedingungen können Coding Scheme 3 oder 4 verwendet werden und so Übertragungsraten von bis zu 20 kbit/s pro Timeslot erreicht werden (Abb. 7.8).

In Abb. 7.9 ist dargestellt, wie Nutzdaten mit Coding Scheme 2 und 3 für die Übertragung über die Luftschnittstelle vorbereitet werden. Das Verfahren ist dabei sehr ähnlich wie die Kodierung der Sprachdaten, die bereits in Abschn. 6.7.5 vorgestellt wurde. Während der Half Rate Convolutional Coder bei Sprachdaten nur für eine Auswahl an Bits verwendet wird, wird dieser hier auf alle Datenbits angewandt. Dies ist auch sinnvoll, da bei der Datenübertragung alle Bits gleich wichtig sind und somit auch alle Bits gleichermaßen geschützt werden müssen. Ergebnis dieser Codierung sind 588 (CS-2) bzw. 676 Bits (CS-3), die dann innerhalb eines Blocks übertragen werden sollen. Da aber in einem Block (= 4 Bursts) nur genau $4 * 114$ Bits = 456 Bits übertragen werden können, muss der so erzeugte Datenstrom vor der Übertragung aber noch angepasst werden. Dies geschieht durch weglassen einzelner Bits (Punktierung). Da der Empfänger weiß, welche Bits punktiert, also nicht übertragen wurden, kann dieser an den geeigneten Stellen „Dummy“ Bits einfügen, die der Convolutional Decoder dann als Fehler betrachtet und entsprechend wieder korrigieren kann. Unterschied zwischen CS-2 und CS-3 ist die Anzahl der punktierten Bits. Je mehr punktierte Bits, desto weniger „richtige“ Fehler dürfen während der Übertragung im Datenblock auftreten. Das in Abb. 7.9 gezeigte USF (Uplink State Flag) Precoding in 6 Bits dient den unterschiedlichen Teilnehmern als Sendeerlaubnis und wird in Abschn. 7.5 zusammen mit dem dort eingeführten RLC/MAC Header ausführlicher beschrieben.

	Modulation	Geschwindigkeit pro Timeslot
GPRS CS-1	GMSK	8 kbit/s
GPRS CS-2	GMSK	12 kbit/s
GPRS CS-3	GMSK	14.4 kbit/s
GPRS CS-4	GMSK	20 kbit/s
EDGE MCS-1	GMSK	8.8 kbit/s
EDGE MCS-2	GMSK	11.2 kbit/s
EDGE MCS-3	GMSK	14.8 kbit/s
EDGE MCS-4	GMSK	17.6 kbit/s
EDGE MCS-5	8PSK	22.4 kbit/s
EDGE MCS-6	8PSK	29.6 kbit/s
EDGE MCS-7	8PSK	44.8 kbit/s
EDGE MCS-8	8PSK	54.4 kbit/s
EDGE MCS-9	8PSK	59.2 kbit/s

Abb. 7.8 GPRS Coding Schemes (CS) und EDGE-Modulation and Coding Schemes (MCS) im Überblick

7.3.4 EDGE (EGPRS)

Um die GPRS Übertragungsgeschwindigkeit noch weiter zu steigern, wurde für GPRS als nächste Ausbaustufe ein weiteres Modulationsverfahren nach dem 8PSK-Prinzip unter dem Namen EDGE (Enhanced Data Rates for GSM Evolution) standardisiert.

Statt 1 Bit pro Übertragungsschritt wie bei der bisher für GSM und GPRS verwendeten GMSK-Modulation, werden mit EDGE 3 Bits pro Übertragungsschritt gesendet. Zusammen mit dem höchsten der insgesamt 9 Coding Schemes sind Übertragungsraten von bis zu 60 kbit/s pro Timeslot möglich. Zwar waren für das neue Modulationsverfahren viele der ursprünglich verwendeten Komponenten im Radionetzwerk nicht geeignet, diese wurden jedoch nach und nach durch zyklisch stattfindende Netzwerkerneuerungen ersetzt. Darüber hinaus müssen auch die Endgeräte das EDGE-Modulationsverfahren unterstützen. Auch dies gehört heute zur Standardausstattung. Ein weiterer Vorteil von neun unterschiedlichen Modulation and Coding Schemes (MCS) gegenüber den vier GPRS Coding Schemes ist eine exakte Verwendung der für die aktuelle Übertragungsqualität geeigneten Modulation und Kodierung. Da sich Netz-

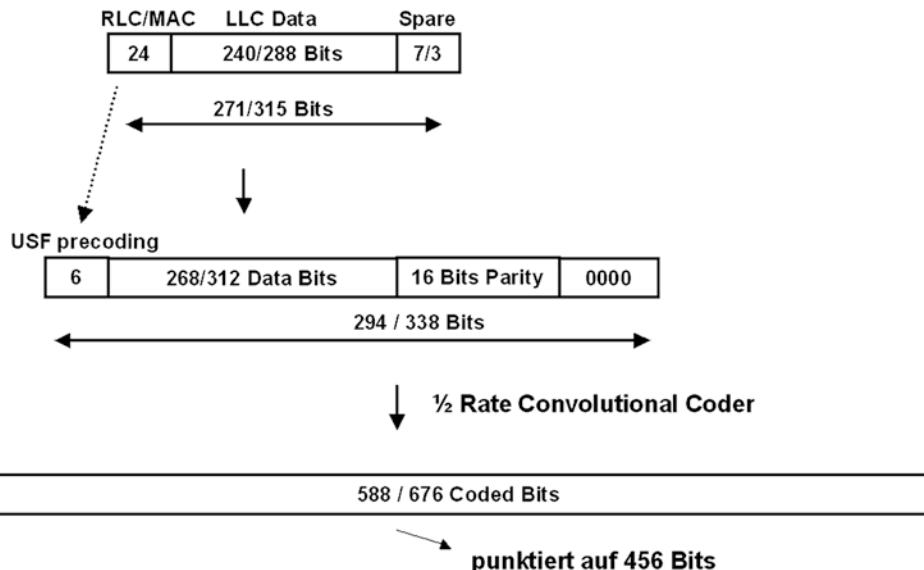


Abb. 7.9 CS-2 und CS-3-Kodierung von GPRS-Daten

werk und Endgerät im Gegensatz zu GPRS auch ständig gegenseitig über die Signalqualität beim Empfang der vorhergehenden Datenpakete informieren, kann somit schnell auf geänderte Übertragungsbedingungen reagiert werden. Dies senkt die Fehlerrate und ermöglicht bei jeder Signalqualität die optimale Geschwindigkeit. Durch diesen Regelmechanismus ist es auch in der Praxis tatsächlich möglich, MCS-8 und 9 bei guten Übertragungsbedingungen zu verwenden.

Trotz schneller Reaktion auf sich ändernde Übertragungsbedingungen ist es natürlich weiterhin möglich, dass Datenblocks nicht korrekt empfangen werden. Auch hier wurde mit EDGE der GPRS-Standard erweitert, um den Datentransfer auf höheren Schichten nicht ins Stocken geraten zu lassen. Um Übertragungsfehler zu beheben, kann z. B. ein Verfahren namens „Incremental Redundancy“ eingesetzt werden. Wie zuvor schon bei den GPRS Coding Schemes gezeigt, werden nicht alle berechneten Fehlerkorrekturbits auch tatsächlich gesendet (puncturing). Tritt ein Übertragungsfehler auf, wird mit Incremental Redundancy das Paket nicht einfach erneut übertragen, sondern es werden nun Fehlerkorrekturbits gesendet, die vorher nicht übertragen wurden. Auf der Empfängerseite können dann die Fehlerkorrekturbits vom ersten und zweiten Übertragungsversuch kombiniert werden. Da nun mehr Fehlerkorrekturbits zur Verfügung stehen, steigen die Chancen, die Übertragungsfehler im Paket zu korrigieren.

Statt Incremental Redundancy ist es mit EDGE auch möglich, den Inhalt eines fehlerhaft empfangenen Paketes, das zuvor mit einem hohen MCS gesendet wurde auf

zwei Pakete mit niedrigerem MCS aufzuteilen. Diese Methode wird Re-Segmentation genannt.

Auch beim Interleaving, also dem Mischen von Bits um punktuelle Übertragungsfehler über den Block zu streuen (vgl. Abschn. 6.7.5), wurden mit EDGE-Erweiterungen am Standard vorgenommen. Bei GPRS wird ein Datenblock unabhängig vom Coding Scheme immer über 4 Bursts gesendet. Bei den EDGE MCS 7–9 jedoch wurde die Länge eines Blocks auf zwei Bursts reduziert und somit das Interleaving verkürzt. Somit müssen bei einem Übertragungsfehler nur zwei Bursts neu übertragen werden und nicht vier. Dies ist vor allem bei der Verwendung von Frequency Hopping ein großer Vorteil. Dieses Verfahren wird verwendet, um nicht konstant auf einer Frequenz zu senden, die gestört oder für die aktuellen Übertragungsbedingungen ungeeignet ist. Während Frequency Hopping für die Sprachübertragung aufgrund der Kaschierung von kleinen Fehlern im Sprachdecoder gut geeignet ist, lassen sich viele falsche Bits in einem Burst bei der Datenübertragung nicht verstecken und der fehlerhafte Block muss komplett neu übertragen werden.

7.3.5 Mobile Device Classes

Alle heute auf dem Markt befindlichen GPRS fähigen Geräte gehören zur sogenannten Mobile Station Class B. Diese Endgeräte können gleichzeitig im GSM leitungsvermittelnden Netz und dem GPRS paketvermittelnden Netz angemeldet sein und beide Dienste nutzen. Einschränkend gilt jedoch, dass während eines Telefongesprächs keine Daten per GPRS gesendet oder empfangen werden können. Umgekehrt bedeutet dies auch, dass während einer laufenden Datenübertragung keine Telefonate geführt werden können. Baut der Anwender während einer GPRS-Datenübertragung ein Telefonat auf, wird die GPRS-Datenübertragung unterbrochen. Nach Ende des Gesprächs wird die Datenübertragung ohne erneute Verbindungsaufnahme weitergeführt. Dies kann vor allem dann eine Einschränkung sein, wenn das Endgerät neben dem Telefonat noch zusätzlich Daten übertragen möchte.

Da ein Endgerät der Klasse B während einer GPRS-Datenübertragung den Paging Channel (PCH) nicht beobachten kann, konnten ursprünglich in dieser Zeit ankommende Telefongespräche oder SMS-Nachrichten nicht empfangen werden. Bei burstartiger Datenübertragung wie dem Internet Browsing war die Chance jedoch recht hoch, das eingehende Gespräch oder die SMS trotzdem zu empfangen, falls nur kurzzeitig Daten übertragen wurden, und die Nachrichten auf dem Paging Channel zumindest einmal wiederholt wurden. Es konnte jedoch nicht ausgeschlossen werden, dass die Paging-Nachricht nicht gesehen wurde. Abhilfe schaffen heute in der Praxis zwei Mechanismen. Die meisten Netzbetreiber koordinieren heute die Datenübertragung mit dem Eingang von Paging Nachrichten für Anrufe und SMS Nachrichten und unterbrechen die Datenübertragung. Eine andere Möglichkeit ist die Nutzung des Network Operation Mode 1 (NOM 1), der im nächsten Abschnitt vorgestellt wird.

Neben der heute in der Praxis eingesetzten Mobile Station Class B beschreibt der GPRS Standard auch noch die Klassen A und C. Während die Klasse C nur die Anmeldung entweder im leitungsvermittelten Teil oder im paketvermittelten Teil des Netzes erlaubt, spezifizierte die Klasse A die simultane Nutzung von paket- und leitungsvermittelnden Diensten. Voraussetzung dafür wären jedoch zwei unabhängige Sende- und Empfangsteile im Endgerät. Dies wurde in der Praxis jedoch nie umgesetzt.

Um ohne komplexere Hardware dennoch gleichzeitig Sprache und Daten zu übertragen, findet sich für Class B-Endgeräte eine Erweiterung namens Dual Transfer Mode (DTM) in den 3GPP Standards. DTM synchronisiert die leitungs- und paketvermittelten Teile des GSM/GPRS-Netzwerkes und ermöglicht somit die gleichzeitige Nutzung mit nur einer Sende- und Empfangseinheit im Endgerät. In der Praxis ist zu beobachten, dass zwar viele Endgeräte heute DTM unterstützen, diese Funktion aber nur in wenigen Netzwerken aktiviert ist.

7.3.6 Network Operation Mode (NOM)

Ähnlich der verschiedenen Endgeräteklassen mit unterschiedlicher Komplexität gibt es auch für das Netzwerk unterschiedliche Betriebsmodi, die Network Operation Mode (NOM) 1, 2 und 3 genannt werden. In der Praxis werden nur die ersten beiden verwendet.

Der Network Operation Mode 2 (NOM 2) ist der einfachste Netzwerk-Modus und wurde deshalb bei der Einführung fast aller GPRS-Netze in der Praxis verwendet. Auch heute ist dieser Modus noch bei den meisten Netzwerken anzutreffen. NOM 2 verwendet für einen Teil der Signalisierung die schon vorhandenen GSM-Signalisierungskanäle wie den RACH, den AGCH und den PCH. Nachrichten auf diesen Kanälen werden vom BSC transparent zwischen Endgerät und GPRS-Netzwerk weitergegeben. Mehr dazu in Abschn. 7.4. Um diesen Mode so einfach wie möglich zu halten, gibt es zwischen den GSM und GPRS-Kernnetzen keine Verbindung. Dies führt wie schon erwähnt dazu, dass ohne zusätzliche Maßnahmen eingehende Telefongespräche und SMS-Nachrichten während einer aktiven GPRS-Datenübertragung mit einem Mobile Station Class B-Endgerät nicht empfangen werden können.

Seltener anzutreffen ist in der Praxis der Network Operation Mode 1 (NOM 1). Auch dieser verwendet die GSM-Signalisierungskanäle wie bei NOM 2.

Zusätzlich wird in NOM-1 das Gs Interface zwischen der MSC im leitungsvermittelnden GSM-Teil des Netzwerkes und dem SGSN (Serving GPRS Support Node) im GPRS-Teil des Netzwerkes eingeführt. Der SGSN ist dabei das paketvermittelnde Gegenstück der MSC und ist neben dem Vermitteln von Datenpaketen auch für das Mobility Management und Session Management der Teilnehmer zuständig. Mehr dazu in Abschn. 7.7.

Über das Gs Interface ist es möglich, GSM und GPRS-Signalisierungsvorgänge zu synchronisieren und zusammenzufassen. Dies hat folgende Vorteile:

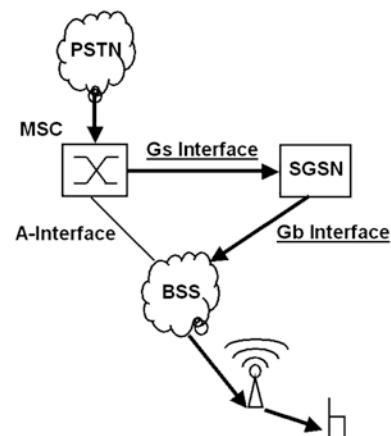
- Bei eingehenden Gesprächen sucht die MSC den Teilnehmer nicht direkt über die BSC mit einer Paging-Nachricht, sondern schickt die Paging-Nachricht stattdessen an den SGSN, der dann seinerseits den Teilnehmer benachrichtigt. Dies hat den großen Vorteil, dass auch ein Endgerät der Mobile Station Class B während eines aktiven GPRS-Datentransfers über das eingehende Gespräch informiert werden kann.
- Location Area Update und Routing Area Update müssen nicht mehr getrennt für GSM und GPRS durchgeführt werden. Bei Bedarf erfolgt ein Combined Location Update mit dem SGSN. Der SGSN gibt die Daten während des Vorgangs dann auch an das MSC/VLR weiter. Während der Vorgang im Netzwerk dadurch etwas komplizierter wird, vereinfacht sich der Vorgang für das Endgerät, da die Prozedur statt zweimal nur noch einmal durchgeführt werden muss. Außerdem werden weniger Signalisierungsressourcen im Radionetzwerk benötigt.

Die Signialisierung auf dem Gs Interface erfolgt über das in Abschn. 6.4.2 vorgestellte MAP-Protokoll, das für diese Zwecke erweitert wurde (Abb. 7.10).

Während bei der Einführung von GPRS im Feld zunächst bei den meisten Netzwerkbetreibern auf NOM 2 gesetzt wurde, verwenden manche Netzwerkbetreiber heute auch NOM 1.

Um eingehende SMS und Gespräche auch während einer aktiven GPRS Datenübertragung zu signalisieren, wird in der Praxis statt NOM 1 von manchen Netzbetreibern noch eine weitere Alternative verwendet, die sich Paging Coordination nennt und auf NOM 2 basiert. In dieser Variante gibt es kein Gs Interface, sondern der Base Station Controller informiert die GPRS Packet Control Unit (PCU), die nachfolgend beschrieben wird, über eine eingehende SMS oder einen Anruf. Bei aktiver Datenübertragung schickt dann die PCU, ähnlich wie bei NOM 1, die Paging Nachricht während der aktuellen Datenübertragung an das Endgerät.

Abb. 7.10 Paging für ein eingehendes Gespräch über das Gs Interface



Welcher Network Operation Mode vom Netzwerk verwendet wird, erfahren die Endgeräte über den Broadcast-Kanal (BCCH bzw. PBCCH) der Zelle in der SYS_INFO 13 Nachricht.

7.3.7 GPRS-Kanalstruktur auf der Luftschnittstelle

Mit GPRS wurden für die Datenübertragung und Signalisierung folgende neue logische Kanäle auf der Luftschnittstelle eingeführt:

Wichtigster Kanal aus Endnutzersicht ist sicherlich der Packet Data Traffic Channel (PDTCH), welcher bei GPRS die eigentlichen Nutzdaten überträgt. Dieser Kanal wird in Uplink- wie auch in Downlinkrichtung verwendet. Up- und Downlinkrichtung werden jedoch unabhängig voneinander vom Netzwerk zugewiesen. Der PDTCH verwendet ähnlich einem leitungsvermittelten GSM Traffic Channel (TCH) bis auf wenige Ausnahmen alle Bursts eines Timeslots (GPRS 52 Multiframe).

Zu jedem PDTCH gehört auch ein Packet Associated Control Channel (PACCH). Der PACCH ist ein bidirekionaler Kanal und wird für die Übertragung von Signalisierungsnachrichten verwendet. Diese sind zum Beispiel notwendig, um den korrekten Empfang von Datenpaketen zu bestätigen. Außerdem wird die Ressourcenzuteilung (also die Zuweisung von Blocks eines PDTCH an einen Teilnehmer) mit Uplink und Downlink Assignment-Nachrichten über diesen Kanal gesteuert.

Ein PACCH wird auf den gleichen Timeslots wie ein PDTCH übertragen. Um den PACCH und den PDTCH zu unterscheiden, gibt es im Header jedes Datenpakets, wie in Abb. 7.11 dargestellt, ein Payload Type-Feld.

Der Packet Timing Advance Control Channel (PTCCH) wird für die Timing Advance-Kontrolle von aktiven Endgeräten verwendet. In vom Netzwerk vorgegebenen Intervallen senden die aktiven Endgeräte in Uplink-Richtung des PTCCH einen kurzen Burst, der vom Netzwerk für die Berechnung des Timing Advance verwendet wird. Das Ergebnis wird dann den Endgeräten in Downlinkrichtung des PTCCH mitgeteilt (Abb. 7.12).

Im Network Operation Mode 2 werden neben diesen Kanälen auch folgende bereits existierende GSM-Kanäle verwendet (vgl. Abschn. 6.7.3):

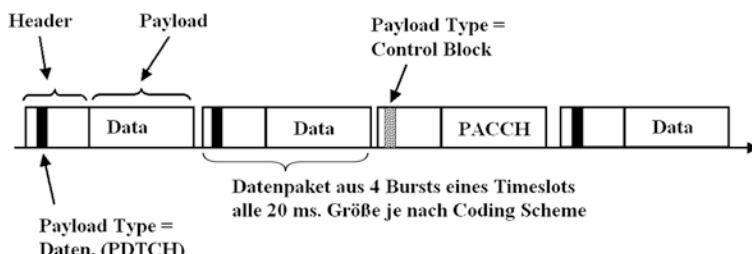


Abb. 7.11 PDTCH und PACCH werden wahlweise auf dem gleichen Timeslot gesendet

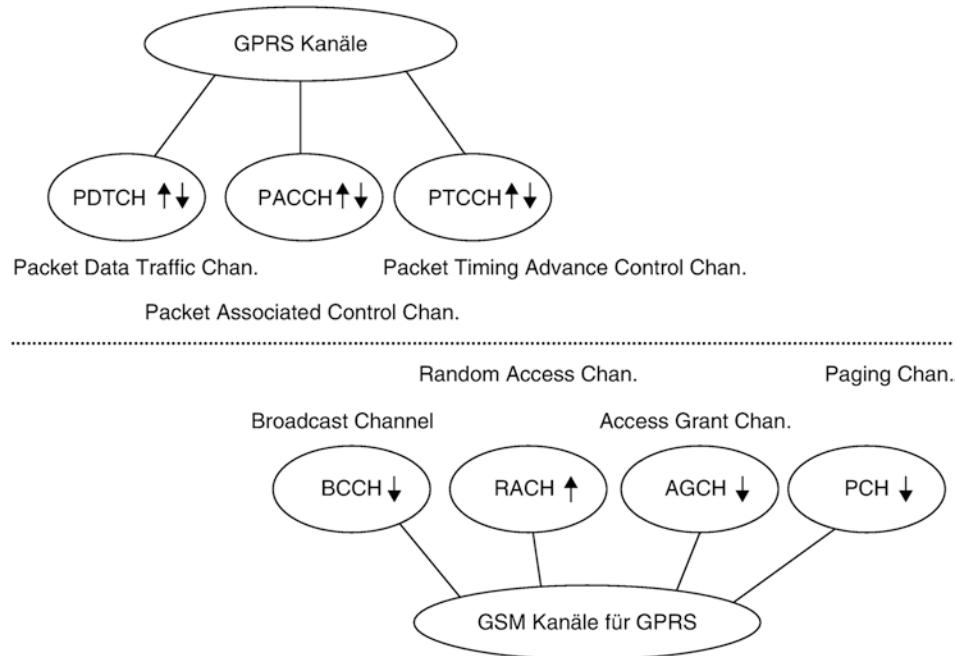


Abb. 7.12 Logische Kanäle in NOM 2 für GPRS

Der Random Access Channel (RACH) wird bei GPRS für die Anforderung von Uplink Ressourcen (Blocks auf dem Uplink PDTCH) verwendet. Statt einer GSM Channel Request-Nachricht wird dafür jedoch eine Packet Channel Request-Nachricht gesendet.

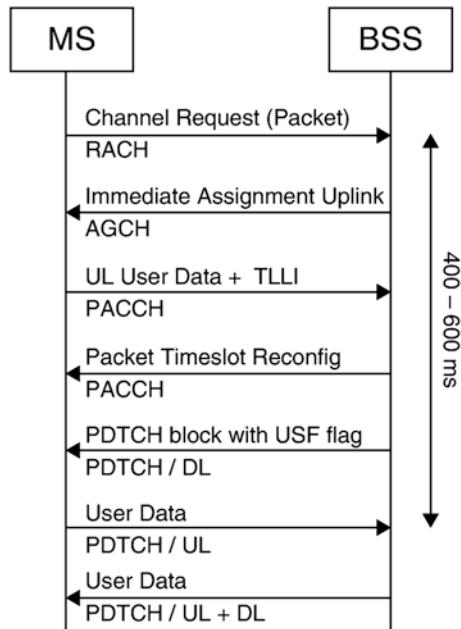
Der PACCH kann nur für die Ressourcenzuteilung an ein Endgerät verwendet werden, wenn diesem bereits Datenblocks in der jeweils anderen Richtung zugeteilt worden sind. Für eine neue Ressourcenzuteilung wird der GSM Access Grant Channel (AGCH) verwendet.

Abb. 7.13 zeigt die Verwendung von RACH, AGCH und PACCH bei der Zuteilung von Uplink Ressourcen und die anschließende Datenübertragung über den PDTCH. Weitere Details hierzu in Abschn. 7.5 über das Radio Resource Management.

Ist ein Endgerät für längere Zeit inaktiv, geht es in den Standby Mode über, der in Abschn. 7.3 näher beschrieben wird. Im Standby Mode ist es nicht mehr möglich, dem Endgerät unmittelbar Ressourcen über den AGCH zuzuteilen. In diesem Zustand muss deshalb das Endgerät zunächst über den Paging Channel gerufen werden.

Auch der Broadcast-Kanal (BCCH) wird für GPRS verwendet. Neben den schon bisher bekannten SYS_INFO-Nachrichten mit GSM-Systeminformationen wird für GPRS nun noch zusätzlich die GPRS spezifische SYS_INFO 13-Nachricht ausgestrahlt. Diese

Abb. 7.13 Anforderung einer Uplink Ressource in NOM 2



enthält alle für das Endgerät wichtigen GPRS-Parameter wie zum Beispiel den Network Operation Mode oder den Routing Area Code.

7.4 GPRS-Zustandsmodell

Bei GSM befindet sich ein am Netzwerk angemeldetes Endgerät entweder im Idle Mode oder im Dedicated Mode. Im Idle Mode gibt es keine Verbindung zwischen Endgerät und Netzwerk und das Endgerät überprüft nur von Zeit zu Zeit den Paging-Kanal. Im Dedicated Mode hingegen existiert zwischen Endgerät und Netzwerk eine aktive Verbindung (z. B. ein Telefongespräch) und es werden in regelmäßigen Abständen Daten gesendet und empfangen. Für GPRS wurde dieses Zustandsmodell für die Anforderungen der Paketdatenübertragung etwas modifiziert:

Im Idle State ist das Endgerät nicht am GPRS-Netzwerk angemeldet, der Aufenthaltsort des Teilnehmers ist nicht bekannt. Somit kann der Teilnehmer keine Daten übertragen, es ist keine GPRS Session (auch PDP-Kontext genannt, siehe Abschn. 7.5.2) aktiv. Leider birgt die Bezeichnung Idle State bei GPRS ein großes Verwechslungsrisiko mit dem GSM Idle Mode. Während im GSM Idle Mode das Endgerät eingebucht aber im Ruhezustand ist und somit jederzeit erreicht werden kann, ist ein Endgerät im GPRS Idle State nicht eingebucht und kann auch vom Netzwerk nicht angesprochen werden (Abb. 7.14).

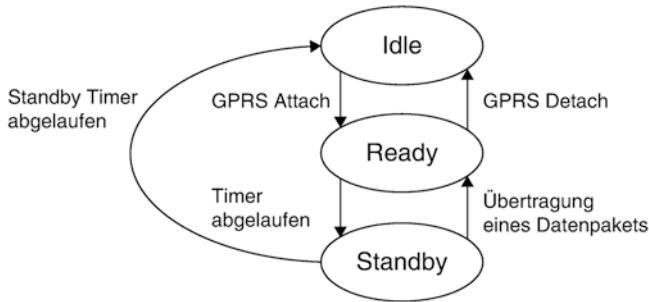


Abb. 7.14 GPRS-Zustandsdiagramm

Möchte sich ein Endgerät im GPRS-Netzwerk einbuchen, wechselt es in den GPRS Ready State mit der Übertragung der Packet Channel Request-Nachricht wie in Abb. 7.13 gezeigt. Im Ready State kann das Netzwerk jederzeit Downlinkressourcen über den AGCH zuweisen und Daten an das Endgerät schicken. Dies ist möglich, da dem GPRS-Netzwerk die Zelle bekannt ist, in dem sich der Teilnehmer gerade aufhält. Dies bedeutet umgekehrt aber auch, dass das Endgerät jeden Zellwechsel dem Netzwerk durch eine Cell Update-Nachricht mitteilen muss. Das Endgerät bleibt im Ready State, solange entweder Signalisierungsnachrichten oder Nutzdaten übertragen werden, sowie noch für einige Zeit nach dem Ende der Übertragung. So wird sichergestellt, dass nachfolgende Datenpakete ohne große Verzögerung zugestellt werden können. Über den Ready Timer wird die Zeit bestimmt, die ein Endgerät nach einer Datenübertragung noch im Ready State verbleibt. Sein maximaler Wert, auf den der Ready Timer nach jeder Datenübertragung zurückgestellt wird, ist Teil der GPRS-Systeminformationen, die auf dem BCCH oder PBCCH ausgestrahlt werden. Ein typischer Wert des Ready Timers, der in vielen Netzwerken verwendet wird, ist 44 s. Nach Ablauf des Ready Timers wechselt das Endgerät in den Standby State.

Der Ready State macht keine Aussage darüber, ob ein Teilnehmer Daten von und zum Internet übertragen kann. Hierfür wird ein sogenannter PDP-Kontext benötigt, der in Abschn. 7.5.2 beschrieben wird. Der Ready State bedeutet für das Endgerät und Netzwerk lediglich, dass Daten und Signalisierungsnachrichten sofort ohne vorheriges Paging an das Endgerät zugestellt werden können.

Der GPRS Ready State ähnelt stark dem GSM Dedicated Mode, da beide Zustände für die Übertragung von Daten gedacht sind. Im Falle des GSM Dedicated Mode sind dies vorwiegend Signalisierungs- und Sprachdaten, im GPRS Ready State hingegen neben der Signialisierung hauptsächlich IP-Pakete. Während jedoch im GSM Dedicated Mode das Mobility Management durch die BSC (vgl. Handover) kontrolliert wird, bleibt diese Aufgabe im GPRS Ready State dem Endgerät überlassen.

Wie bei GSM führt das Endgerät auch bei GPRS-Signalstärkemessungen der aktuellen Zelle sowie aller Nachbarzellen durch. Bei Bedarf startet das Endgerät dann

ohne Hilfe oder Anweisung des Netzwerkes einen Zellwechselvorgang, der Cell Update genannt wird. Nach Wechsel in die neue Zelle werden zunächst die Systeminformationen der Zelle aus deren Broadcast-Kanal (BCCH oder PBCCH) ausgelesen. Danach nimmt das Endgerät über den RACH Kontakt zum Netzwerk auf und sendet ein leeres Datenpaket. Hieran erkennt das Netzwerk, dass der Teilnehmer die Zelle gewechselt hat und ändert die Route für nachfolgende Datenpakete entsprechend. Der komplette Cell Update-Vorgang benötigt etwa 2 s. Wird ein Cell Update-Vorgang während einer laufenden Datenübertragung durchgeführt, entsteht dadurch natürlich eine Unterbrechung der Übertragung von mindestens 2 s. Daten, die in Downlink-Richtung während des Cell Updates in der alten Zelle übertragen wurden, müssen erneut gesendet werden.

Um die Unterbrechungen bei Cell Updates so kurz wie möglich zu halten, wurde in die GPRS-Standards ein Verfahren eingebracht, das sich Network Assisted Cell Change, kurz NACC nennt. Bei diesem Verfahren kündigt das Endgerät einen bevorstehenden Zellwechsel dem Netzwerk an. Dieses kann dann die Systeminformationen der neuen Zelle an das Endgerät schicken und eine laufende Datenübertragung in Downlink-Richtung anhalten. Das Endgerät wechselt danach in die neue Zelle und sendet ein neues Datenpaket. Somit reduziert sich die Unterbrechungszeit wesentlich, da der BCCH der neuen Zelle nicht gelesen werden muss und auch keine Daten im Downlink verloren gehen, die in der neuen Zelle erneut gesendet werden müssten.

Nach Ablauf des Ready Timers wechselt das Endgerät in den Standby State. In diesem Zustand informiert das Endgerät das Netzwerk nur noch über einen Zellwechsel, wenn die neue Zelle zu einer neuen Routing Area gehört. Nachteil ist jedoch, dass bei ankommenden Daten der Teilnehmer erst über den PCH in der gesamten Routing Area gesucht werden muss. Eine Routing Area ist ein Teil einer Location Area, besteht also auch aus einer Anzahl Zellen. Eigentlich hätten auch die Location Areas auch für GPRS weiterverwendet werden können. Durch die feinere Unterteilung einer Location Area in eine oder mehrere Routing Areas gibt man den Netzbetreibern aber die Möglichkeit, die richtige Balance zwischen Nutzung des Paging Channels bei ankommenden Daten und die Häufigkeit der Routing Area Updates unabhängig von GSM zu kontrollieren.

Ändert sich bei einem Zellwechsel nur die Routing Area, führt das Endgerät einen Routing Area Update mit dem GPRS-Netzwerk durch. Ändert sich neben der Routing Area auch gleichzeitig die Location Area, führt das Endgerät zusätzlich noch einen Location Area Update durch. Auch im Ready State ist bei einem Zellwechsel statt eines Cell Updates ein Routing und Location Area Update notwendig, wenn sich diese bei einem Zellwechsel ändern (Abb. 7.15).

Vorteil des Standby State ist der geringere Signalisierungsaufwand und somit für die Endgeräte eine längere Akkulaufzeit. Für das Netzwerk hat dies den Vorteil, dass knappe Signalisierungsressourcen auf dem RACH, AGCH und PDTCH gespart werden. Nachteil ist jedoch, dass bei ankommenden Daten das Endgerät erst über den Paging Channel gerufen werden muss, was zusätzlich Zeit kostet. Da bei üblichen Ready Timer-Werten

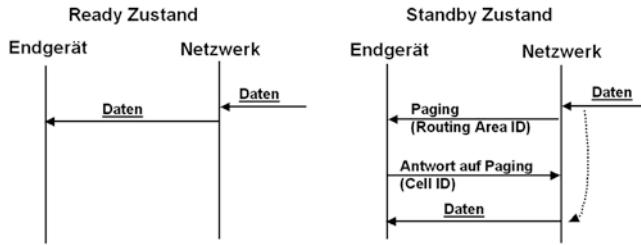


Abb. 7.15 Unterschied zwischen Ready- und Standby-Zustand

von z. B. 44 s in den meisten Fällen keine Daten mehr nachkommen, muss der Paging Channel in der Praxis nicht sehr oft verwendet werden.

In Uplink-Richtung gibt es keinen Unterschied zwischen Ready und Standby State. Möchte ein Endgerät im Standby State Daten schicken, geht das Endgerät mit dem Senden des ersten Pakets automatisch wieder in den Ready State über.

7.5 GPRS-Netzwerkelemente

Durch Einteilung der Timeslots für GSM und GPRS im Radionetzwerk ist es möglich, beide Dienste über die gleichen Basisstationen zu betreiben. Hierzu war lediglich ein Softwareupdate für die Basisstationen und BSCs notwendig. Aufgrund der großen Unterschiede zwischen Leitungs- und Paketvermittlung wurde es jedoch notwendig, drei neue Netzwerkkomponenten für GPRS in das vorhandene Netzwerk zu integrieren. Diese werden in Abb. 7.16 gezeigt und in diesem Abschnitt näher beschrieben.

7.5.1 Die Packet Control Unit (PCU)

Der Base Station Controller (BSC) ist Teil des leitungsvermittelnden GSM-Netzwerkes und verbindet Endgeräte über 16 kbit/s Kanäle mit der MSC im Kernnetzwerk. Außerdem ist der BSC für das Handover der Verbindungen zuständig. Da GPRS-Teilnehmer jedoch keine dedizierte 16 kbit/s-Verbindung mehr mit dem Netzwerk haben, ist die Architektur der BSC nicht für GPRS geeignet. Aus diesem Grund wurde die Packet Control Unit (PCU) im Netzwerk eingeführt, die das paketvermittelnde Gegenstück zur BSC im Radionetzwerk darstellt. Die PCU hat folgende Aufgaben:

Die Vergabe von Timeslots, respektive PDTCHs in Up- und Downlink-Richtung an die einzelnen Teilnehmer: Uplinkressourcen werden mit einer Packet Channel Request-Nachricht über den RACH angefordert. Die BSC empfängt diese Nachrichten und leitet sie an die PCU weiter. Dieser Umweg ist notwendig, da über den RACH auch GSM-Kanalanforderungen gesendet werden, die von der BSC selber bearbeitet werden.

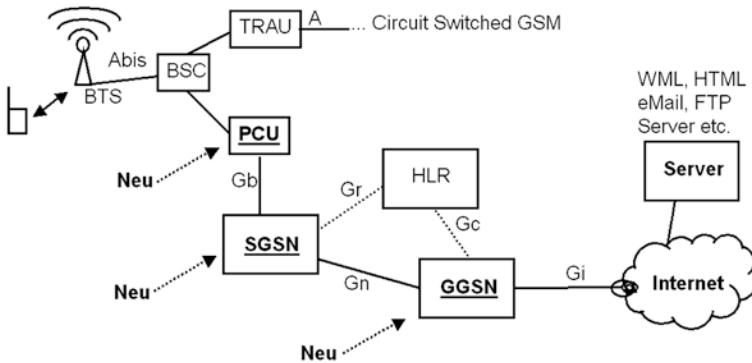


Abb. 7.16 GPRS-Netzkomponenten

Da über eine Zelle viele Teilnehmer gleichzeitig Daten übertragen können, ist die PCU auch für die Flusskontrolle der Daten in Up- und Downlink-Richtung zuständig, sowie für die Priorisierung der unterschiedlichen Datenströme.

In Uplinkrichtung überprüft die PCU die ankommenden Datenblocks und fordert bei Übertragungsfehlern ggf. die Daten erneut beim Endgerät an.

Befindet sich ein Teilnehmer im Standby State, ist die PCU auch für das Paging verantwortlich, wenn vom Netzwerk neue Daten für die Übertragung bereitstehen.

Um sicherzustellen, dass Datenpakete von Endgeräten in ihren vorgesehenen Zeitfenstern an der BTS eintreffen, ist die PCU auch für das grundsätzliche Timing Advance Management zuständig. Bei Vergabe von Up- und Downlinkressourcen teilt die PCU dem Endgerät mit, wann es den Packet Timing Advance Control Channel (PTCCH) für Timing Advance Messungen verwenden darf. Die eigentlichen Timing Advance Messungen auf dem PTCCH sowie die Berechnung und Übertragung der Werte an das Endgerät werden jedoch von der BTS autonom ausgeführt.

Wie in Abb. 7.16 zu sehen ist, ist die PCU direkt mit dem BSC verbunden und ist heute typischerweise nur ein logisches BSC Subsystem. Die für GPRS reservierten Timeslots für PDTCHs und GPRS Signalisierung werden dann von der PCU selbstständig verwaltet.

7.5.2 Der Serving GPRS Support Node (SGSN)

Der Serving GPRS Support Node (SGSN) ist das Gegenstück zur leitungsvermittelnden MSC im paketorientierten GPRS-Netzwerk. Er erfüllt im Wesentlichen die gleichen Aufgaben, die sich in die Teilbereiche User Plane und Signalling Plane unterteilen lassen:

Die User Plane ist für die Übertragung von Nutzdatenpaketen zwischen Teilnehmern und externen Netzwerken wie dem Internet oder einem Firmenintranet zuständig. Alle Pakete, die beim SGSN für einen Teilnehmer eingehen, werden an die für die aktuelle

Zelle des Teilnehmers zuständige PCU weitergeleitet („geroutet“). Liefert die PCU Pakete eines Teilnehmers, reicht der SGSN diese an den nächsten Netzwerkknoten, den Gateway GPRS Support Node (GGSN) weiter, der im nächsten Abschnitt beschrieben wird (Abb. 7.17).

Im GPRS-Kernnetz wird zwischen den unterschiedlichen Netzwerkkomponenten IP als Transportprotokoll verwendet. Dies hat den großen Vorteil, dass eine Vielzahl unterschiedlicher Übertragungstechnologien verwendet werden können.

Von und zur PCU wurde ursprünglich das Frame Relay-Protokoll für den Transport der Userdatenpakete gewählt. Die Entscheidung, hier nicht auch das IP-Protokoll zu verwenden, ist aus heutiger Sicht nur schwer nachvollziehbar. Die Wahl auf Frame Relay fiel unter anderem deswegen, da Datenpakete zwischen SGSN und PCU, wie im BSS damals üblich, über E-1-Leitungen transportiert werden sollten. Frame Relay mit seiner paketorientierten Architektur und vielen Ähnlichkeiten zu ATM eignete sich gut für die Übertragung über 2 Mbit/s E-1-Kanäle und wurde viele Jahren in der Weitverkehrstechnik eingesetzt. Nachteil war allerdings neben einer komplizierteren Netzwerkarchitektur auch, dass der SGSN die Datenpakete von der PCU aus einem Frame Relay-Paket extrahieren und danach per IP weiter an den GGSN schicken muss musste und umgekehrt. Mit der zunehmenden Verbreitung des Internet Protokolls wurde dann in den 3GPP Standards IP als Alternative zu Frame Relay für das Gb Interface spezifiziert. Dieses hat mittlerweile Frame Relay Verbindungen in Mobilfunknetzwerken ersetzt (Abb. 7.17).

Während bei GSM leitungsvermittelte Verbindungen nur die Übertragung auf der Luftschnittstelle zwischen Endgerät und BTS verschlüsselt wird, werden GPRS-Datenpakete zwischen Endgerät und SGSN durchgehend geschützt (Abb. 7.18). Dies hat den Vorteil, dass nun auch das Abis Interface geschützt ist, das oft über eine Mikrowellenverbindung läuft und somit leicht abhörbar ist. Nachteil ist jedoch, dass die Rechenleistung für die Verschlüsselung im Netzwerk nicht mehr über viele Basisstationen verteilt werden kann, sondern im SGSN konzentriert ist.

Neben dem Weiterleiten von Daten zwischen mobilen Teilnehmern und dem GGSN ist eine weitere wichtige Aufgabe des SGSNs die Teilnehmersignalisierung. Diese Aufgabe wird von der Signalling Plane übernommen, die in zwei Bereiche aufgeteilt ist:

Um als mobiler Teilnehmer Daten mit dem Internet austauschen, muss zunächst über das GPRS-Netzwerk eine Datenverbindung aufgebaut werden. Diese Prozedur wird Packet Data Protocol (PDP) Context Activation genannt und ist Teil des Session

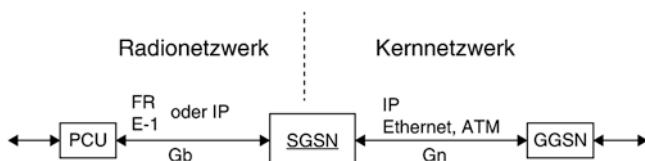


Abb. 7.17 Schnittstellen und Protokolle des SGSN auf Layer 2 + 3

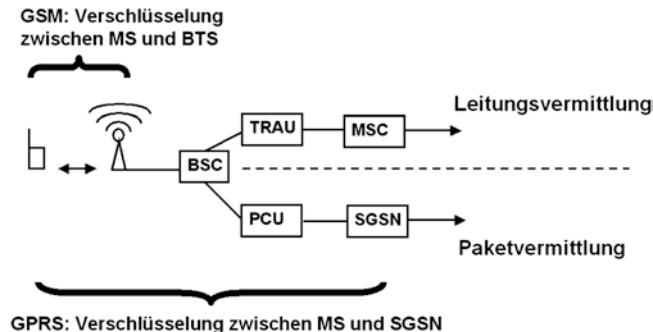


Abb. 7.18 Verschlüsselung in GSM und GPRS im Vergleich

Managements des SGSN. Aus Anwendersicht wird während der PDP Context Activation Prozedur dem Endgerät eine IP-Adresse zugeteilt.

Um am Netzwerk angemeldete Teilnehmer jederzeit erreichen zu können, ist der SGSN auch für die Verwaltung der Position jedes Teilnehmers in seinem Versorgungsbereich zuständig. Diese Aufgabe wird GPRS Mobility Management (GMM) genannt und ist dem Mobility Management der MSC sehr ähnlich. Zusammen mit der Session Management (SM) Komponente wird das zugehörige Protokoll auch GMM/SM genannt.

Um GPRS-Dienste in Rechnung stellen zu können, sammelt der SGSN und der im Anschluss beschriebene GGSN die dazu nötigen Billing Information in Call Detail Records (CDR). Diese werden an einen Billing Server weitergeleitet. Die CDRs des SGSN sind vor allem bei Teilnehmern wichtig, die in einem ausländischen Netz roamen. Wie wir in Abschn. 7.5.2 noch genauer betrachten werden, ist bei einer GPRS-Verbindung im Ausland der SGSN die einzige Komponente, die im ausländischen Netz einen CDR generieren kann. In diesem Fall dienen die CDRs des SGSN dem Netzbetreibern für die Abrechnung des Datenverkehrs des fremden Teilnehmers. Für Teilnehmer, die sich im Heimnetzwerk befinden, erzeugt hingegen auch der GGSN einen CDR, und die Billing Informationen des SGSN sind somit nicht unbedingt erforderlich.

7.5.3 Der Gateway GPRS Support Node (GGSN)

Während der SGSN das Bindeglied zwischen Radionetzwerk und GPRS-Kernnetz darstellt sowie die Mobilität der Teilnehmer verwaltet, verbindet der GGSN das GPRS-Netzwerk mit dem Internet. Für Geschäftskunden kann der GGSN das GPRS-Netzwerk auch direkt mit dem Intranet einer Firma verbinden.

Der GGSN ist am Aufbau einer Internetverbindung, also dem Aufbau eines PDP-Kontextes beteiligt und vergibt die IP-Adressen. Danach fungiert der GGSN als fester Bezugspunkt (Anchor Point) der Verbindung. Bewegt sich ein Teilnehmer mit einem

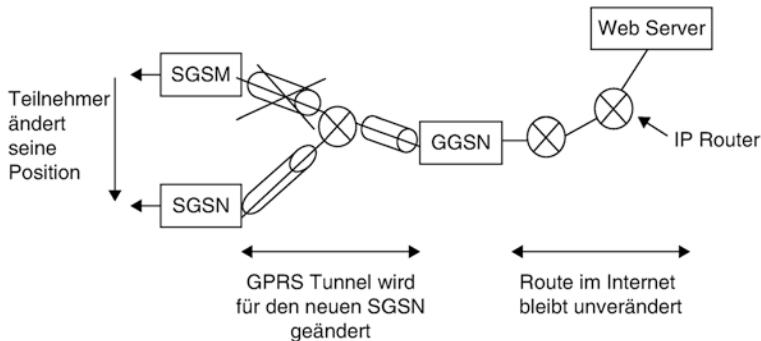


Abb. 7.19 Änderung des Aufenthaltsorts eines GPRS-Teilnehmers

aktiven PDP-Kontext in das Gebiet eines anderen SGSN, ändert das GPRS-Netzwerk entsprechend das Routing der Datenpakete zwischen dem GGSN und dem neuen SGSN. Im Internet ist dies jedoch nicht sichtbar, da der GGSN während der Verbindung niemals gewechselt wird. Dies ist auch notwendig, da Router im Internet Datenpakete für eine IP-Adresse immer an das gleiche Ziel weiterleiten und ihre Routing-Tabellen für mobile Teilnehmer nicht anpassen können. Durch den GGSN wird also die Mobilität des Teilnehmers vor dem Internet versteckt. Abb. 7.19 zeigt, wie sich eine Positionsänderung des Teilnehmers im GPRS-Netzwerk auswirkt.

7.6 GPRS Radio Resource Management

Wie in Abb. 7.5 dargestellt wurde, kann ein Timeslot bei GPRS mehreren Teilnehmern gleichzeitig zugeordnet sein. Daten der unterschiedlichen Teilnehmer werden dann abwechselnd übertragen. Einem Teilnehmer können andererseits aber auch zur Steigerung seiner Übertragungsgeschwindigkeit mehrere Timeslots gleichzeitig zugeordnet sein. Die kleinste GPRS-Übertragungseinheit ist dabei ein Block, der aus 4 Bursts eines Timeslots besteht.

Jeder Datenblock auf dem PDTCH oder PACCH besteht aus einem RLC/MAC (Radio Link Control/Medium Access Control) Header und einem Nutzdatenfeld.

Möchte das Netzwerk Daten an ein Endgerät senden, muss zuvor eine virtuelle Verbindung in Form eines Temporary Block Flow (TBF) zwischen Netzwerk und Endgerät aufgebaut werden. Dies geschieht durch Zuweisung eines Temporary Flow Identifier (TFI) in einer Packet Downlink Assignment-Nachricht. Alle Datenblocks im Downlink, die für diesen Teilnehmer bestimmt sind, enthalten in ihrem RLC/MAC Header dann diesen TFI-Wert. Abb. 7.20 zeigt, wie mehrere Datenblocks nacheinander auf dem gleichen PDTCH übertragen werden. Datenblock 1 und 3 mit TFI=1 sind für das dargestellte Endgerät bestimmt. Datenblock zwei mit TFI=3 im RLC/MAC Header ist für

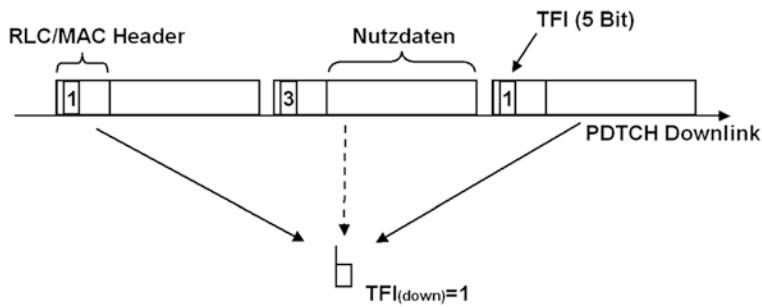


Abb. 7.20 Auswertung des TFI im Endgerät

ein anderes Endgerät bestimmt. Zwar empfängt das dargestellte Endgerät auch diesen Datenblock, ignoriert diesen aber aufgrund des anderen TFI-Wertes.

Die Bestätigung der Downlinkdatenblocks erfolgt über den Packet Associated Control Channel (PACCH) Uplink. Damit das Endgerät weiß, wann eine Bestätigung im Uplink gesendet werden darf, enthält ein Downlinkdatenblock des Teilnehmers im RLC/MAC Header die Information, in welchen Uplinkblocks die Bestätigungsmeldungen gesendet werden dürfen. Somit ist es möglich, Downlinkblocks zu bestätigen, ohne einen Uplink TBF zuzuteilen.

Nachdem die PCU alle Daten eines Teilnehmers aus ihrer Sendequeue übertragen hat, wird der Downlink TBF beendet. Dazu wird im letzten Downlinkblock das Final Block Indicator Bit gesetzt. Nach Empfang des letzten Datenblocks beendet das Endgerät seine Empfangsbereitschaft und überprüft im Ready-Zustand fortan nur noch den Access Grant Channel (AGCH) auf Zuteilung einer neuen Downlink Ressource.

Auch in Uplinkrichtung muss für die Datenübertragung zuerst eine virtuelle Verbindung in Form eines Temporary Block Flows aufgebaut werden. Dem Endgerät wird wiederum wie in Downlinkrichtung ein TFI-Wert zugewiesen. Dies geschieht über eine Immediate Packet Assignment-Nachricht auf dem Access Grant Channel.

Die Zuweisung eines TFI ist jedoch noch keine Sendeerlaubnis. Im Unterschied zu anderen Technologien wie z. B. Ethernet, darf ein GPRS-Endgerät nur Daten senden, wenn es zuvor die Erlaubnis vom Netzwerk erhalten hat. In den Standards gibt es mehrere Möglichkeiten, wie diese Zuteilung (Allocation) erfolgen kann. Die heute gebräuchlichste Art ist die Dynamic Allocation. Dazu wird dem Endgerät in einer Packet Uplink Assignment-Nachricht neben den zugeteilten Timeslotnummern und seinem TFI-Wert auch ein Uplink State Flag (USF) Wert übergeben. Das Endgerät überprüft fortan alle Downlinkdatenblocks in allen zugeteilten (Uplink-) Timeslots, ob im RLC/MAC Header sein USF-Wert enthalten ist. Findet das Endgerät seinen USF-Wert, darf es im nächsten Uplinkblock Daten übertragen (Abb. 7.21).

Endgeräte mit hohen Multislot-Klassen sind nicht in jeder Konfiguration in der Lage, alle Timeslots im Downlink abzuhören, die für Uplink Übertragungen zugewiesen

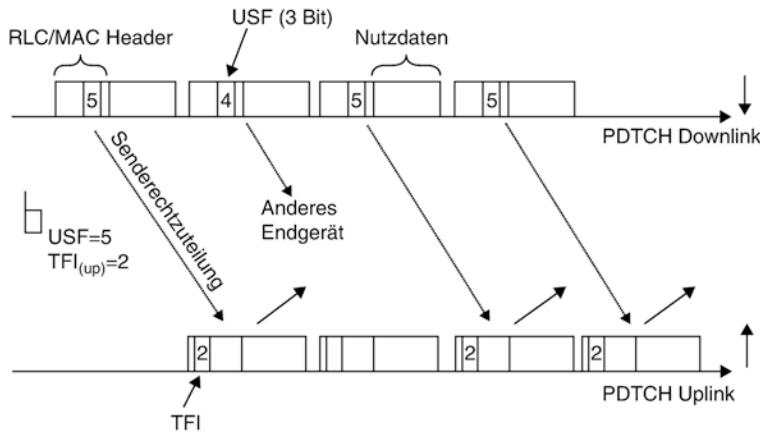


Abb. 7.21 Verwendung des Uplink State Flags

wurden. In diesen Fällen wird das Extended Dynamic Allocation Verfahren verwendet. Hier wird ein einzelnes USF verwendet, um Ressourcen für mehrere Uplink Timeslots gleichzeitig zuzuweisen.

Während einem Endgerät ein Uplink TBF zugewiesen ist, muss das Netzwerk in Downlink-Richtung die empfangenen Datenblocks bestätigen. Dies geschieht über den Packet Associated Control Channel (PACCH) Downlink. Da der PACCH und PDTCH auf den gleichen Timeslots übertragen werden, dient das Payload Type-Feld im RLC/MAC Header des Datenblocks der Unterscheidung dieser zwei logischen Kanäle. Da das Endgerät in den zugeteilten Timeslots wegen des Uplink State Flags sowieso die Datenblocks im Downlink mitlesen muss, können auch diese Kontrollnachrichten ohne zusätzlichen Aufwand empfangen werden.

Nachdem das Endgerät seinen Sendepuffer geleert hat, muss es dem Netzwerk signalisieren, dass keine weiteren Uplinkressourcen notwendig sind. Dies geschieht mit der Countdown-Prozedur. Im RLC/MAC Header jedes Uplink Blocks befindet sich dazu ein 4 Bit Countdown-Zähler, der bei der Übertragung jedes Blocks am Ende des Datentransfers vom Endgerät um 1 reduziert wird. Ist der Zähler bei 0 angekommen, vergibt die PCU keine Uplink Blocks mehr an den Teilnehmer und der Temporary Flow Identifier und das Uplink State Flag werden ungültig.

Zwar ist die beschriebene Handhabung eines Uplink TBF sehr effizient, in der Praxis verursacht dieses Verfahren jedoch eine große Verzögerungszeit, wenn nur sporadisch Datenpakete gesendet werden. Dies wirkt sich z. B. negativ auf das TCP Acknowledgement Verfahren auf Layer 4 des Protokollstacks aus, da Pakete nur mit großer Verzögerung gesendet werden und somit den Aufbau und Start einer TCP Verbindung verlangsamen. Aus diesem Grund wurde der GPRS-Standard nachträglich um das sogenannte Extended Uplink TBF-Verfahren erweitert. Unterstützen sowohl Netzwerk wie auch Endgerät das Verfahren, wird der TBF am Ende der Countdown-

Prozedur nicht automatisch geschlossen, sondern kann vom Netzwerk weiter offen gehalten werden. Dies ermöglicht dem Endgerät, neue Daten im Uplink ohne neue Ressourceanforderung, also ohne Verzögerung zu senden.

Abb. 7.22 zeigt den Inhalt einer Timeslot Reconfiguration-Nachricht, die für die Konfigurationsänderung eines bestehenden TBFs verwendet wird. Die gezeigte Nachricht teilt einem Endgerät 3 Timeslots in Downlink-Richtung zu. Außerdem enthält sie neben den TFIs für Uplink und Downlinkrichtung weitere Informationen wie den zu verwendenden Timing Advance-Wert und den Coding Scheme, den das Endgerät für Datenpakete in Uplinkrichtung verwenden soll.

7.7 GPRS-Schnittstellen und Protokolle

Wie bereits in der Übersicht in Abb. 7.16 dargestellt, werden die GPRS-Netzwerklemente über standardisierte und somit offene Schnittstellen miteinander verbunden. Mit Ausnahme der PCUs, die vom gleichen Hersteller wie die BSCs in einem Netzwerk sein müssen, können alle anderen Netzwerkkomponenten frei gewählt werden. Eine PCU von Nokia kann z. B. an einen SGSN von Ericsson angeschlossen werden, der

```
[...]
      RLC/MAC PACKET TIMESLOT RECONFIGURE
000111-- Message Type : 7 = packet timeslot reconfigure
-----00 Page Mode : 0 = normal paging
Global TFI:
--01111-     Uplink Temporary Flow Identifier : 15
00----- Channel Coding Command : Use CS-1 in Uplink
Global Packet Timing Advance:
----0001     Uplink TA Index : 1
101-----     Uplink TA Timeslot Number : 5
----0001     Downlink TA Index : 1
101-----     Downlink TA Timeslot Number : 5
---0---- Downlink RLC Mode : RLC acknowledged mode
---0---- CTRL ACK : 0 = downlink TBF already established
xxxxxxx  Downlink Temporary Flow ID: 11
xxxxxxx  Uplink Temporary Flow ID: 15
Downlink Timeslot Allocation:
-0-----     Timeslot Number 0 : 0
-0-----     Timeslot Number 1 : 0
-0-----     Timeslot Number 2 : 0
----0---     Timeslot Number 3 : 0
----1--     Timeslot Number 4 : 1 = assigned
-----1-     Timeslot Number 5 : 1 = assigned
-----1-     Timeslot Number 6 : 1 = assigned
0-----     Timeslot Number 7 : 0
Frequency Parameters:
--000---     Training Sequence Code : 0
xxxxxxx  ARFCN : 067
[...]
```

Abb. 7.22 Packet Timeslot Reconfiguration-Nachricht

wiederum mit einem GGSN von Cisco verbunden sein könnte. Natürlich können auch alle Komponenten vom gleichen Netzwerkhersteller sein, da die meisten Hersteller alle Komponenten eines GPRS-Netzwerkes anbieten.

Das Abis Interface verbindet die BTS mit dem BSC. Auf allen Timeslots, in denen im Radionetzwerk GPRS PDTCHs konfiguriert sind, kommt der Protokollstack, wie in Abb. 7.23 gezeigt, zum Einsatz. Üblicherweise werden die Daten transparent auf das nicht standardisierte Interface zwischen BSC und PCU weitergegeben. Auf den unteren Layern des Protokollstacks wird das RLC/MAC-Protokoll für das Radio Ressource Management verwendet. Eine Protokollsicht höher sorgt das Logical Link Control Protocol (LLC) für das Framing der Nutzdatenpakete und Signalisierungsnachrichten (Mobility Management/Session Management). Optional sorgt das LLC-Protokoll auch für eine gesicherte Verbindung zwischen Endgerät und SGSN durch einen Bestätigungsmechanismus für korrekt empfangene Blocks (Acknowledged Mode). Eine Stufe höher verpackt das Subnetwork Dependant Convergence Protocol (SNDCP) die IP Nutzdatenpakete für den korrekten Versand über das Radionetzwerk. Optional führt SNDCP auch eine Kompression der IP Header der Nutzdaten oder eine Kompression der kompletten Nutzdatenpakete durch. Der LLC Layer und alle höheren Schichten sind für die PCU, BSC und BTS transparent, da sie für eine Ende-zu-Ende-Verbindung im Radionetzwerk sorgen.

Das Gb Interface verbindet den SGSN mit der PCU. Auf Layer 1 wurden für dieses Interface ursprünglich 2 Mbit/s E-1 Verbindungen verwendet. Ein SGSN verwaltet in der Praxis mehrere PCUs, die ursprünglich jeweils mit mehreren 2 Mbit/s Leitungen an den SGSN angeschlossen wurden. Auf Layer 2 und 3 des Protokollstacks wurde, wie bereits zuvor beschrieben das Frame Relay-Protokoll verwendet. Heute wird auf dieser Schnittstelle üblicherweise das IP Protokoll über eine optische Verbindung für den Datentransport verwendet.

Das Gn Interface verbindet SGSNs mit GGSNs innerhalb eines GPRS-Netzwerkes und wird in 3GPP TS 29.060 spezifiziert. Je nach Größe des Netzwerkes besteht ein GPRS-Netzwerk aus einem oder mehreren SGSNs. Auch die Anzahl der GGSNs, die

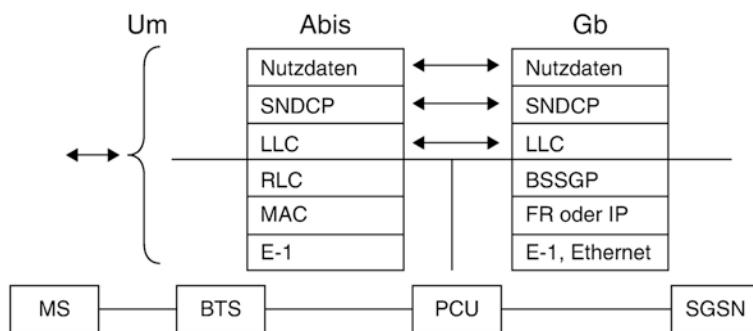


Abb. 7.23 GPRS-Protokollstacks im Radionetzwerk

üblicherweise aber geringer als die Anzahl der SGSNs ist, wird maßgeblich von der Anzahl der Nutzer des Netzwerks bestimmt (Abb. 7.24).

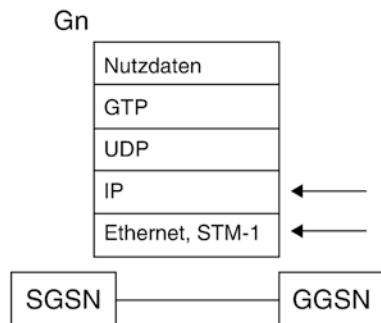
Weiterhin ist es möglich, unterschiedliche GGSNs für unterschiedliche Anwendungen zu verwenden. Während ein oder mehrere GGSNs z. B. für Vertragskunden zuständig sein könnten, sind andere speziell auf die Bereitstellung des GPRS-Dienstes für Prepaid Subscriber spezialisiert. Es spricht aber auch nichts dagegen, für alle Teilnehmer den gleichen GGSN zu verwenden. Aus Last- und auch aus Redundanzgründen werden üblicherweise mehrere GGSN in einem Netzwerk eingesetzt, die dann in unterschiedlichen Städten untergebracht sind. Beim Ausfall eines Standorts können neue Verbindungen automatisch umgelenkt werden.

Auf Layer 3 des OSI-Protokollstacks (Network Layer), wird auf dem Gn Interface das IP-Protokoll für das Routing aller Nutzdatenpakete der Teilnehmer, sowie für Signalisierungsnachrichten zwischen SGSNs und GGSNs verwendet.

Nutzdatenpakete der Teilnehmer werden auf dem Gn Interface nicht direkt, sondern in GPRS Tunneling Protocol (GTP) Paketen verpackt übertragen. Dies erzeugt zwar zusätzlichen Overhead im GPRS-Kernnetz, ist aber aus folgenden Gründen notwendig:

Jeder Router im Internet zwischen GGSN und Ziel entscheidet anhand der IP Zieladresse des Datenpaketes und einer Routing-Tabelle, wohin das Datenpaket weitergeleitet werden soll. Da sich die Position eines Teilnehmers im Internet nicht ändert, ist dieses Verfahren sehr effizient. Im GPRS-Netzwerk kann dieses Verfahren jedoch nicht angewandt werden, da die Teilnehmer jederzeit ihren Standort wechseln können. Somit kann sich, wie schon in Abb. 7.19 gezeigt, die Route für die Datenpakete durch das GPRS-Netzwerk jederzeit ändern. Da zwischen GGSN und SGSN beliebig viele IP Router geschaltet sein können, müsste in jedem Router im GPRS-Netzwerk bei einer Positionsänderung des Teilnehmers das Routing für seine IP-Adresse geändert werden. Um dies zu vermeiden, wird innerhalb des GPRS-Netzwerkes nicht mit der Quell- und Ziel IP-Adresse des Nutzdatenpaketes geroutet, sondern es werden die IP-Adressen von SGSN und GGSN verwendet. Das eigentliche Nutzdatenpaket wird zwischen dem SGSN und GGSN in ein GTP-Paket eingepackt und läuft somit transparent durch das GPRS-Netzwerk. Ändert sich später die Position des Teilnehmers, muss dem GGSN nur

Abb. 7.24 Der Gn Protokoll-Stack



die IP-Adresse des neuen SGSNs mitgeteilt werden. Der große Vorteil dieses Verfahrens ist somit, dass die Router zwischen SGSN und GGSN ihre Routing-Tabellen nicht ändern müssen.

Abb. 7.25 zeigt die wichtigsten Parameter der Protokollsichten eines Pakets auf dem Gn Interface. Die IP-Adressen auf Layer 3 stammen vom SGSN und GGSN, während die IP-Adressen des Nutzdatenpaketes, das in einem GTP-Paket eingepackt ist, die IP-Adressen des Teilnehmers und des angesprochenen Servers im Internet enthält. Dies bedeutet paradoxe Weise, dass in einem GTP Datenpaket zwei IP Header vorhanden sind. Erhält der GGSN ein GTP-Paket von einem SGSN, entfernt dieser alle Header inklusive des GTP Headers. Danach wird das vom Teilnehmer ursprünglich gesendete IP-Paket auf dem Gi Interface zum Internet weitergesendet.

Das Gi Interface verbindet das GPRS-Netzwerk über den GGSN mit einem externen Netzwerk. Aus Sicht des externen Netzwerkes verhält sich der GGSN wie ein ganz normaler IP Router. Während die Nutzdaten der Teilnehmer innerhalb des GPRS-Netzwerkes in GTP-Pakete eingepackt werden, sind die Nutzdatenpakete auf dem Gi Interface wieder in ihrer originalen Form als IP-Pakete auf Layer 3 präsent. Aus Redundanzgründen oder um die verfügbare Bandbreite zu erhöhen, kann auch dieses Interface gleichzeitig über mehrere Verbindungen an den oder die nächsten Router im Internet oder dem Firmennetzwerk angeschlossen sein.

Über das Gr Interface kommuniziert der SGSN mit dem HLR. Diese Verbindung ist nötig, da das HLR für jeden Teilnehmer dessen Berechtigungen für GPRS speichert. Dazu gehört unter anderem:

- Ob ein Teilnehmer den GPRS-Dienst nutzen darf
- Welche Dienste von einem Teilnehmer verwendet werden dürfen (Access Point Names, APN)
- Internationales GPRS Roaming und Beschränkungen

Wie in Kap. 6 gezeigt wurde, ist das HLR ein SS-7 Signalling Control Point (SCP). Aus diesem Grund war das Gr Interface ursprünglich auf E-1 Trunks und dem SS-7-Protokoll

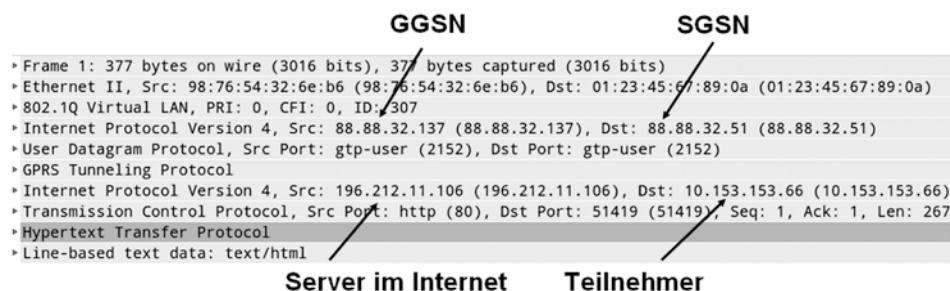


Abb. 7.25 GTP-Paket auf dem Gn Interface

aufgebaut, wird aber heute üblicherweise auch über IP Verbindungen übertragen. Für die Signalisierungsnachrichten kommt das Mobile Application Part (MAP)-Protokoll zur Anwendung, das auch die MSC für die Kommunikation mit dem HLR verwendet. Nachfolgend einige Beispiele für Nachrichten zwischen SGSN und HLR:

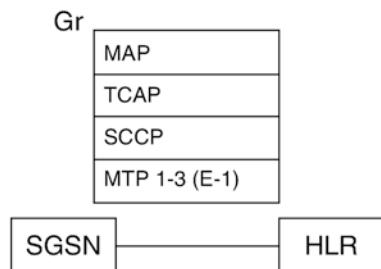
- Send Authentication Information: Diese Nachricht wird vom SGSN zum HLR geschickt, wenn sich ein Teilnehmer am Netzwerk anmeldet, um dessen Authentifizierungsdaten zu erhalten.
- Update Location: Mit dieser Nachricht informiert der SGSN das HLR, dass sich ein Teilnehmer in seinem Versorgungsgebiet angemeldet hat und erfolgreich identifiziert wurde.
- Insert Subscriber Data: Als Antwort auf die Update Location-Nachricht liefert das HLR dem SGSN Informationen zurück, welche Dienste der Teilnehmer im GPRS-Netzwerk verwenden darf (Abb. 7.26).

Das Gp Interface wird verwendet, um GPRS-Netzwerke unterschiedlicher Länder miteinander zu verbinden und somit GPRS International Roaming zu ermöglichen. Die Nutzdaten der Teilnehmer werden über das Gp Interface zwischen dem SGSN im besuchten Netzwerk und dem GGSN im Heimatnetzwerk in gleicher Weise wie über das netzwerkinterne Gn Interface übertragen (Abb. 7.27).

Befindet sich ein deutscher Teilnehmer z. B. in Spanien und nutzt GPRS für die Datenübertragung, werden seine Daten vom SGSN in Spanien an den GGSN im deutschen Heimatnetz übertragen. Von dort aus werden seine Datenpakete dann ins Internet weitergeleitet. Zunächst scheint dies wenig sinnvoll, da für die Daten des Teilnehmers theoretisch nicht nur der SGSN, sondern auch der GGSN in Spanien genutzt werden könnte. Der große Vorteil der Verwendung des Gp Interfaces und des GGSN im Heimatnetzwerk des Teilnehmers ist jedoch, dass keine Einstellungen im Endgerät für das Roaming geändert werden müssen und dass das Heimatnetzwerk die Kontrolle über die Abrechnung der übertragenen Daten behält.

Über das Gp Interface werden nur IP Nutz- und Signalisierungsdaten zwischen SGSNs und GGSNs unterschiedlicher GPRS-Netzwerke ausgetauscht. Um GPRS

Abb. 7.26 Der Gr Protokoll-Stack



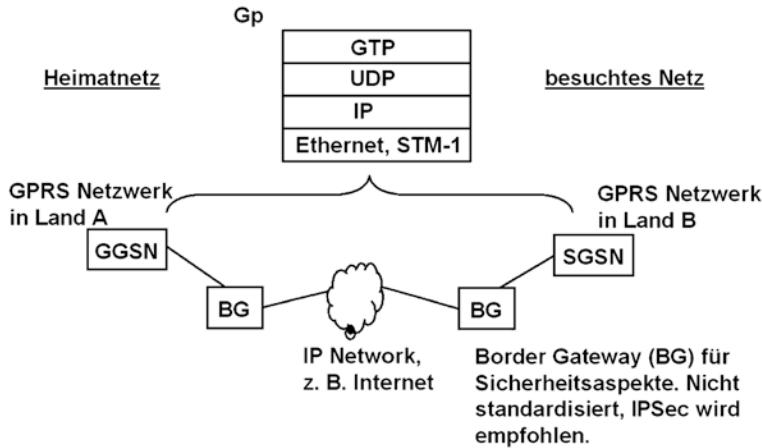


Abb. 7.27 Gp Interface für internationales Roaming

Roaming zu ermöglichen, muss ein SGSN noch zusätzlich über das Gr Interface auch Zugriff auf das HLR im Heimatnetzwerk des Teilnehmers haben.

Das Gs Interface ist ein optionales Interface und verbindet die SGSNs des paketvermittelnden GPRS-Teilnetzwerks mit den MSCs des leitungsvermittelnden GSM-Netzwerkes. Die Vorteile, die sich durch Verwendung dieses Interfaces ergeben, wurden bereits in Abschn. 7.3.6 (Network Operation Mode 1) beschrieben.

7.8 GPRS Mobility und Session Management (GMM/SM)

Neben der Weiterleitung der Nutzdaten zwischen den Teilnehmern und dem Internet sind zwei weitere wesentliche Aufgaben des GPRS-Netzwerkes die Mobilitätsverwaltung der Teilnehmer (Mobility Management) sowie die Kontrolle der Nutzdatenverbindungen (Session Management). Zu diesem Zweck wurden in den GPRS-Standards auf den unterschiedlichen Interfaces Signalisierungsnachrichten und Signalisierungsabläufe definiert. Diese Abläufe werden unter dem Begriff GPRS Mobility Management and Session Management, kurz GMM/SM zusammengefasst.

7.8.1 Mobility Management-Aufgaben

Bevor über ein Endgerät eine Verbindung zu einem externen Netzwerk wie dem Internet aufgebaut werden kann, muss sich das Endgerät zunächst am GPRS-Netzwerk anmelden. Dieser Anmeldevorgang wird GPRS Attach genannt und ist in Abb. 7.28 dargestellt. Das Endgerät beginnt diese Prozedur mit einer Attach Request-Nachricht, die

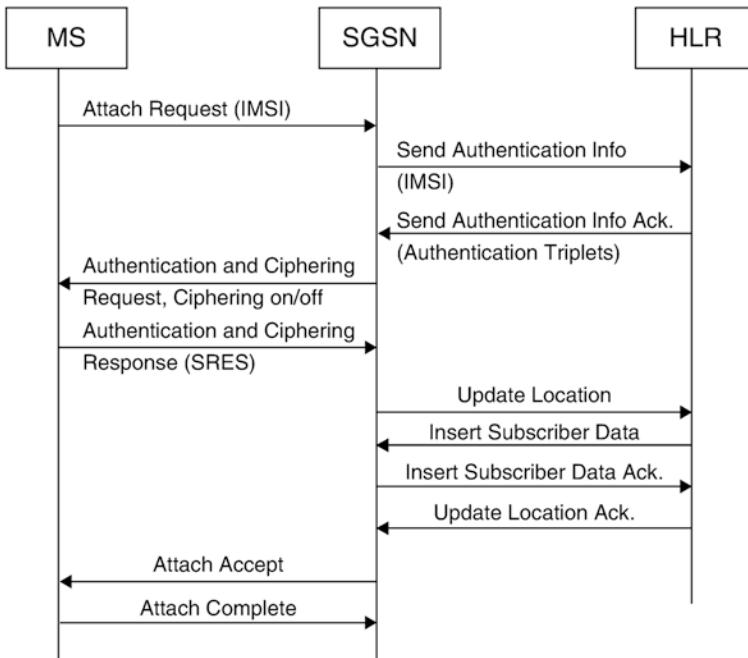


Abb. 7.28 GPRS Attach Message Flow

entweder seine IMSI oder die Packet Temporary Mobile Subscriber Identity (P-TMSI) enthält, die beim letzten Anmeldevorgang oder Routing Area Update vergeben wurde. Die P-TMSI wird auch Temporary Logical Link ID (TLLI) genannt. Somit wird die IMSI nur selten verwendet und die wahre Identität des Teilnehmers ist so nur dem Netzwerk bekannt.

Beim ersten Attach an einem SGSN mit der IMSI kennt der SGSN den Teilnehmer nicht. Deshalb fordert der SGSN daraufhin, wie in Abb. 7.28 gezeigt, die Authentifizierungsdaten mit einer Send Authentication Information-Nachricht beim HLR/Authentication Center (AC) an. Das HLR/AC liefert als Antwort ein oder mehrere Authentication Triplets, deren Erzeugung bereits in Abschn. 6.6.4 beschrieben wurde. Die zurückgegebene Zufallszahl RAND wird danach vom SGSN in einer „Authentication and Ciphering Request“ Nachricht dem Endgerät übergeben. Das Endgerät berechnet dann in der SIM Karte die Antwort (SRES) und antwortet mit einer Authentication and Ciphering Response-Nachricht. Im SGSN wird daraufhin die vom Endgerät erhaltene Antwort mit der SRES des HLR/Authentication Centers verglichen. Stimmt sie überein, ist der Teilnehmer erfolgreich authentifiziert.

Während bei GSM die Verschlüsselung mit einer weiteren Nachricht aktiviert wird, enthält in GPRS schon die Authentication and Ciphering Response-Nachricht diesen Befehl. Wird die Verschlüsselung also gleichzeitig mit der Authentifizierung aktiviert

(optional), werden alle nachfolgenden Signalisierungs- und Nutzdatenpakete verschlüsselt von und zum SGSN übertragen.

Im nächsten Schritt informiert der SGSN das HLR über den Aufenthaltsort des Teilnehmers mit einer Update Location-Nachricht. Das HLR sendet daraufhin die Daten des Teilnehmers mit einer Insert Subscriber Data-Nachricht zum SGSN. Danach wird dem Endgerät die erfolgreiche Anmeldung mit einer Attach Accept-Nachricht bestätigt und das Endgerät beendet den Dialog mit einem Attach Complete.

Abb. 7.29 zeigt Auszüge aus dem Inhalt einer GPRS Attach-Nachricht, die auf dem Gb Interface aufgezeichnet wurde. Die Nachricht enthält unter anderem die beim letzten Location Update oder Attach vergebene TMSI und Informationen über die letzte Position (MCC, MNC, LAC und RAC) des Teilnehmers. Außerdem enthält die Nachricht

```
[...]
Mobility Management: ATTACH REQUEST
  MS Network Capability:
    GPRS encryption algorithm GEA/1: 1 = available
[...]
  -----001   Attach Type : 001bin = GPRS attach
  -100---- GPRS Ciphering Key Sequence Number : 100bin
  DRX Parameter
  01000000  Split PG cycle code : 64 = 64
  -----011  Non-DRX timer: max. 4 sec non-DRX mode after transfer state
  ----0---  SPLIT on CCCH: not supported
  Mobile Identity
  -----100  Type of identity: TMSI
  ----0---
  xxxxxxxx  TMSI: D4CC3EC4h
  Old Routing Area Identification
  xxxxxxxx  Mobile Country Code: 232
  xxxxxxxx  Mobile Network Code: 03
  xxxxxxxx  Location area code: 6F32h
  00000001  Routing area code: 0Fh
  MS Radio Access Capability
  0001---- Access technology type: 1 = GSM E (900 MHz Band)
  Access capabilities
  ---100-- RF power capability: 4h
          A5 bits
  -----1    A5/1: 1 = Encryption algorithm available
  1-----   A5/2: 1 = Encryption algorithm available
  -0-----   A5/3: 0 = Encryption algorithm not available
[...]
  -----1-   ES IND : 1h = early Classmark Sending is implemented
[...]
  Multislot capability
  xxxxxxxx  GPRS multi slot class: 4 (3 downlink + 1 uplink)
  --0----  GPRS extended dynamic allocation: not implemented
  ----1101  Switch-measure-switch value: 0
  1000----  Switch-measure value: 8
  xxxxxxxx  Access technology type: 3 = GSM 1800
  xxxxxxxx  Access capabilities
  001----  RF power capability: 1
  ----1---  ES IND: 1 = early Classmark Sending is implemented
[...]
```

Abb. 7.29 Auszüge aus einer GPRS Attach Request-Nachricht nach 3GPP TS 24.008, 9.4.1

Informationen über die technischen Fähigkeiten des Endgeräts wie z. B. die Multislot-Klasse, welche Frequenzbänder unterstützt werden (900, 1800 MHz,...), etc. Somit ist es möglich, mit der Zeit die GPRS-Fähigkeiten neuer Endgeräte zu erweitern (z. B. bessere Multislot-Klasse) und netzwerkseitig nur Funktionalitäten zu nutzen, die ein Endgerät auch unterstützt.

War der Teilnehmer zuvor bei einem anderen SGSN registriert, ist die Prozedur noch etwas umfangreicher. Vor Abschluss der Prozedur muss dann das HLR mit einer Cancel Location-Nachricht zuerst die Daten im bisherigen SGSN löschen. Erst danach übergibt das HLR die Teilnehmerdaten an den neuen SGSN.

Ist das Gs Interface zwischen MSC und SGSN im Netzwerk vorhanden (NOM 1), kann der GSM Attach und der GPRS Attach in einem Vorgang durchgeführt werden. Dies beschleunigt den Vorgang für das Endgerät und reduziert den Signalisierungsaufwand im Radionetzwerk. Über das Gs Interface gibt der SGSN die Attach-Nachricht auch an die für die Location Area des Teilnehmers zuständige MSC weiter.

Die zweite wichtige Mobility Management-Aufgabe ist der Routing Area Update (RAU). Ähnlich dem GSM Location Update muss dieser immer dann durchgeführt werden, wenn das Endgerät zu einer Zelle wechselt (Cell Update), die zu einer anderen Routing Area gehört. Eine Routing Area ist ein Teilbereich einer GSM Location Area oder kann mit ihr identisch sein. Die Durchführung des Routing Area Updates ist dem GSM Location Update sehr ähnlich (siehe Abschn. 6.8.1). Falls das Gs Interface zwischen MSC und SGSN vorhanden ist, kann der GSM Location Area Update und der GPRS Routing Area Update vom Endgerät gleichzeitig durchgeführt werden. Der SGSN gibt dann die entsprechenden Informationen an die zuständige MSC weiter.

Wechselt ein Endgerät in eine Zelle, die in einer Routing Area eines neuen SGSNs liegt, findet aus Endgerätesicht ein ganz normaler Routing Area Update statt. Der neue SGSN kennt jedoch den Teilnehmer nicht und muss sich erst dessen Authentifizierungs- und Teilnehmerdaten besorgen. Da die Routing Area Update-Nachricht Informationen über die vorherige Routing Area enthält, kann der neue SGSN danach beim alten SGSN diese Informationen anfordern. Dies dient auch gleichzeitig dazu, dass der bisherige SGSN bis auf weiteres alle vom GGSN eingehenden Nutzdatenpakete an den neuen SGSN weiterleitet, um möglichst keine Nutzdaten zu verlieren. Damit der GGSN in Zukunft seine Nutzdaten direkt an den neuen SGSN schickt, informiert der neue SGSN als nächstes den GGSN über den neuen Aufenthaltsort des Teilnehmers. Zum Schluss wird auch das HLR vom neuen Standort des Teilnehmers informiert und die Teilnehmerdaten im alten SGSN gelöscht. Details dieser Prozedur sind in 3GPP TS 23.060 in Abschn. 4.9.1.2.2 beschrieben.

7.8.2 GPRS Session Management

Nachdem sich das Endgerät über die Attach-Prozedur am Netzwerk angemeldet hat, kann nun für die Kommunikation mit dem Internet oder Firmenintranet ein sogenannter

Packet Data Protocol (PDP) Kontext beim Netzwerk beantragt werden. Aus Sicht des Benutzers ist diese Prozedur nötig, um eine IP-Adresse zu erhalten.

Eine paketvermittelnde Verbindung wird in Anlehnung an einen „Voice Call“ auch als „Packet Call“ bezeichnet, da im GPRS-Netz die Paketverbindung ähnlich einer leitungsvermittelten Telefonverbindung explizit auf- und abgebaut wird. Großer Unterschied ist jedoch, dass bei einem Packet Call nur Ressourcen verwendet werden, wenn tatsächlich Daten übertragen werden. Somit ist der PDP-Kontext eines Packet Calls nur eine logische Verbindung, die nur physische Ressourcen benötigt, wenn auch tatsächlich Daten übertragen werden. Auch wenn keine Daten übertragen werden, kann der PDP-Kontext über Stunden oder sogar Tage aktiv bleiben. Dies wird auch als „Always On“ Funktionalität bezeichnet.

Abb. 7.30 zeigt den Ablauf einer PDP Context Activation-Prozedur. Diese wird durch eine PDP Context Activation Request-Nachricht vom Endgerät an den SGSN gestartet. Wichtigster Parameter ist der sogenannte Access Point Name (APN). Der APN dient dem SGSN dazu, den richtigen GGSN (Access Point) für den Übergang ins Internet für den Teilnehmer zu finden. Ein Netzbetreiber hat somit die Möglichkeit, viele unterschiedliche Dienste anzubieten. Dazu gehören zum Beispiel:

- Eine direkte Verbindung mit dem Internet
- Eine direkte Verbindung mit dem Internet für Prepaid Kunden
- Eine direkte IP-Verbindung zu einem Firmennetzwerk

Der SGSN ermittelt mit dem übergebenen APN die IP-Adresse des zu diesem APN gehörenden GGSNs. Für die Namensauflösung in eine IP-Adresse verwendet das GPRS-Netzwerk das Domain Name System (DNS). DNS Server werden auch im Internet ver-

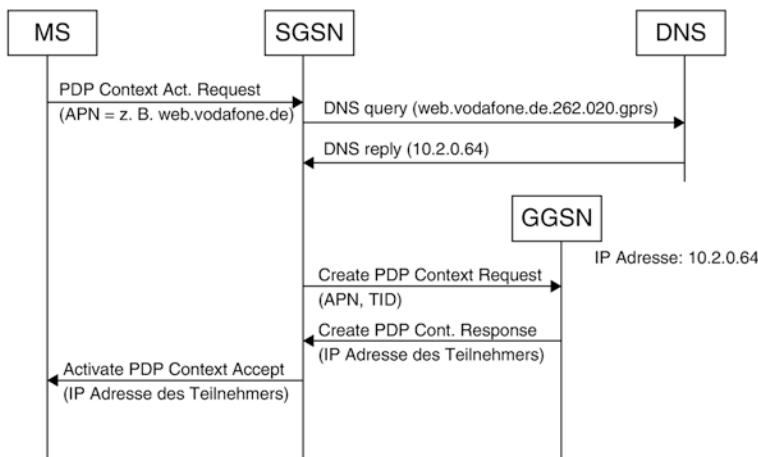


Abb. 7.30 Aufbau eines PDP-Kontext

wendet, um z. B. beim Webbrowsen den Namen einer Website wie z. B. www.spiegel.de in die IP-Adresse des Webservers der Spiegelredaktion umzuwandeln. Um die Adresse des GGSNs zu finden, geht der SGSN in genau gleicher Weise mit einem APN vor. Aus diesem Grund muss sich der Netzbetreiber bei der Vergabe von Namen für APNs auch an die Regeln der DNS Namensgebung halten. Um den APN international eindeutig zu machen, fügt der SGSN an das Ende des APN automatisch den Mobile Country Code (MCC) und den Mobile Network Code (MNC) aus der IMSI des Teilnehmers, sowie die Top Level Domain.gprs hinzu. Übergibt der Teilnehmer z. B. als APN den String „web.vodafone.de“ an das GPRS-Netzwerk, ermittelt der SGSN über eine DNS-Anfrage und den erweiterten APN „web.vodafone.de.mnc002.mcc262.gprs“ die IP-Adresse des zuständigen GGSNs.

Da der APN durch Anfügen des MCC und MNC weltweit eindeutig ist, kann ein Teilnehmer ohne Änderungen seiner GPRS-Einstellungen auch in einem ausländischen Netz roamen. Damit die internationale APN-Namensauflösung erfolgreich ist, müssen alle Domain Name Server der zusammengeschalteten GPRS-Netzwerke verbunden und kaskadiert sein. Weiterhin muss für das GPRS Roaming auch eine SS-7 Signalisierungsverbindung mit dem HLR im Heimatnetzwerk für die Attach-Prozedur vorhanden sein (Gr Interface), sowie eine IP-Verbindung für die Nutzdaten und Signalisierungsdaten zwischen SGSN und GGSN (Gp Interface).

Nachdem die IP-Adresse des für den APN zuständigen GGSNs bekannt ist, leitet der SGSN die PDP Context Activation Anforderung an den GGSN weiter. Teil dieser Nachricht ist der vom Teilnehmer gewünschte APN, sowie seine IMSI. Um später die Nutzdatenpakete des Teilnehmers transparent durch das GPRS-Netzwerk leiten zu können (tunneln), vergibt der SGSN eine sogenannte Tunnel ID (TID) für diesen PDP-Kontext. Diese ist ebenfalls Teil der Nachricht an den GGSN. Die TID wird dabei aus der IMSI des Teilnehmers und einem zwei Stellen langen Network Subsystem Access Point Identifier (NSAPI) zusammengesetzt. Der NSAPI ist notwendig, da ein Teilnehmer mehrere PDP-Kontexte gleichzeitig aufgebaut haben kann. In der Praxis wird dies aber bisher nur von wenigen Netzbetreibern eingesetzt, um z. B. den Internet Zugang von der Kommunikation mit Netzbetreiber spezifischen Diensten wie z. B. MMS logisch zu trennen.

Anhand der APN überprüft der GGSN, mit welchem Netzwerk der Teilnehmer verbunden werden soll. Optional gibt es noch die Möglichkeit, einen Benutzernamen und ein Passwort zwischen Teilnehmer und GGSN in der PDP Context Activation-Nachricht auszutauschen. Dies wird in der Praxis von manchen Netzwerkbetreibern verwendet, was die Konfiguration des Endgeräts unnötig kompliziert. Stimmt auch der GGSN dem Verbindungswunsch zu, vergibt er für den PDP-Kontext eine IP-Adresse und schickt diese in einer PDP Context Activation Response-Nachricht an den SGSN zurück. Außerdem speichert der GGSN für diese Verbindung folgende Informationen:

- TID des Teilnehmers
- IP-Adresse des SGSN für den Austausch von Nutzdaten

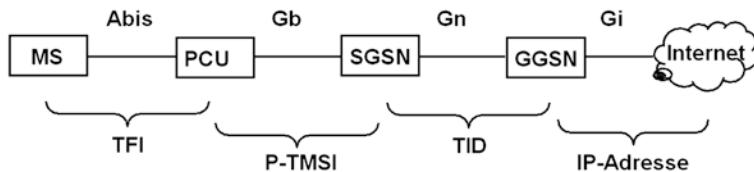


Abb. 7.31 Identifikation der Teilnehmerpakete im GPRS-Netzwerk

- IP-Adresse des SGSN für den Austausch von GPRS Signalisierungsdaten
- Die für den Teilnehmer vergebene IP-Adresse

Nach Erhalt der PDP Context Activation Response-Nachricht speichert der SGSN die IP-Adresse des GGSNs in seinem Eintrag für den neuen PDP-Kontext, da fortan alle Datenpakete des Teilnehmers an diese IP-Adresse weitergeleitet werden. Im letzten Schritt schickt der SGSN eine PDP Context Activation Accept-Nachricht an das Endgerät zurück und übergibt darin die vom GGSN zugewiesene IP-Adresse.

In den verschiedenen Netzwerkabschnitten werden die Datenpakete eines Teilnehmers aufgrund der unterschiedlichen Netzwerkprotokolle und Paketgrößen auch unterschiedlich identifiziert. Auf der Luftschnittstelle mit seinen kleinen Datenpaketen von 456 Bit = 57 Bytes abzüglich Fehlerkorrektur, wird der Teilnehmer mit dem 3 Bit Temporary Flow Identifier (TFI) adressiert. Im Radio-Netzwerk wird der Teilnehmer mit der P-TMSI/TLLI identifiziert und im Kernnetzwerk mit der GPRS Tunnel ID (TID). Nur im externen Netzwerk wie dem Internet wird die zugeteilte IP-Adresse des Teilnehmers für das Routing der Datenpakete verwendet. Abb. 7.31 zeigt diese unterschiedliche Teilnehmeridentifizierung im Überblick.

7.9 Session Management aus Anwendersicht

Aus Anwendersicht wird ein PDP-Kontext immer dann aufgebaut, wenn Daten aus dem Internet übertragen werden sollen. Smartphones in den Anfangszeiten von GPRS bauten einen PDP-Kontext nur beim Start eines Programms auf, das über das Internet Daten austauschen wollte. Nach Beenden des Programms wurde dann der PDP-Kontext wieder abgebaut. Dies hat sich jedoch bei Smartphones wesentlich gewandelt, da heute viele Applikation gleichzeitig laufen und auf eine ständige Internetverbindung angewiesen sind.

Embedded-Geräte die heute ein mobiles (GPRS) Datenmodul eingebaut haben, kommunizieren mit diesem z. B. über eine USB Verbindung, auf der eine serielle Schnittstelle simuliert wird. Über diese virtuelle serielle Schnittstelle wird auch heute noch das in den 1980er-Jahren für leitungsvermittelnde Analogmodems entwickelte „AT-Kommandointerface“ verwendet. Der nächste Abschnitt zeigt nun zunächst, wie

der Modemstack normalerweise verwendet wurde, um eine Internetverbindung über ein Festnetzmodem oder eine leitungsvermittelte GSM-Verbindung aufzubauen. Im nächsten Schritt wird dann gezeigt, wie sich eine Modemverbindung von einer GPRS-Verbindung unterscheidet und welche Konfigurationseinstellungen für GPRS deshalb auf dem Embedded Gerät benötigt werden.

7.9.1 Leitungsvermittelter Verbindungsauflaufbau

Ein Modem bietet für die Verbindungsaufnahme mit einer Gegenstelle eine textbasierte Kommandoschnittstelle an, die AT-Interface genannt wird. Um zum Beispiel eine Telefonnummer zu wählen, wird dem Modem das Kommando ATD zusammen mit der Telefonnummer (z. B. ATD 0.899.011.782) übergeben. Das Modem wählt daraufhin diese Nummer und stellt eine Datenverbindung mit dem Dial In Server-Modem des Internet Service Providers (ISP) her. War der Verbindungsauflaufbau erfolgreich, sendet das Modem eine CONNECT-Nachricht mit der ausgehandelten Übertragungsgeschwindigkeit an das Datenendgerät zurück (z. B. CONNECT 38.400). Daraufhin wechselt das Modem aus dem Kommandomodus in den Übertragungsmodus und leitet alle Daten von nun an transparent weiter.

Um Datenpakete über diese transparente serielle Verbindung zu übertragen, verwendet das DFÜ-Netzwerk das Point to Point Protocol (PPP). Der PPP Client ist dabei das Endgerät, der PPP Server der Dial In Server des Internet Service Providers (Abb. 7.32).

Nach dem Verbindungsauflaufbau wird mit dem PPP-Protokoll zunächst der Teilnehmer authentifiziert. Nach erfolgreicher Authentifizierung übermittelt der PPP Server dann dem Teilnehmer alle für die nachfolgende IP Übertragung nötigen Parameter. Dazu gehört insbesondere die IP-Adresse für das Endgerät, sowie die IP-Adresse des DNS Servers für die Namensauflösung. In der danach beginnenden Datenübertragungsphase

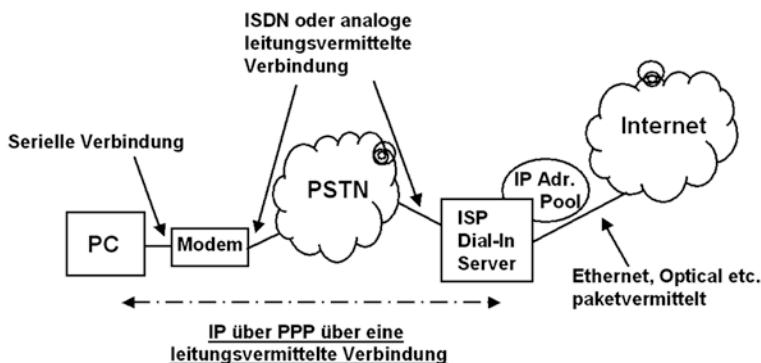


Abb. 7.32 Internetverbindung per Modem und PPP

ist es die Aufgabe des PPP Protokolls, IP-Pakete über die transparente Verbindung zu übertragen. Um den Anfang und das Ende jedes Paketes auf der anderen Seite auch korrekt erkennen zu können, fügt das PPP-Protokoll eine Start- und Endekennung, sowie einen Header an die vom IP Layer erhaltenen Pakete an.

7.9.2 GPRS-Verbindungsauflaufbau

GPRS unterscheidet sich von einer leitungsvermittelten Verbindung in folgenden Punkten:

- GPRS ist bereits eine paketvermittelte Verbindung zum Internet, es gibt keine leitungsvermittelte Verbindung mehr zum ISP Dial In Server.
- Es gibt keine Telefonnummer, die beim Verbindungsauflaufbau gewählt werden müsste.
- Es gibt keinen PPP Server, zu dem sich ein Client verbinden könnte. Im GPRS-Netzwerk vergibt der GGSN die IP-Adressen der Teilnehmer während der PDP Context Activation-Prozedur.

Aus diesen Gründen ist die für eine leitungsvermittelte Verbindung vorgestellte Prozedur für die Kontaktaufnahme mit dem Internet eigentlich wenig geeignet. Um aber keine spezielle Software für die Interneteinwahl per GPRS für Datenendgeräte entwickeln zu müssen, wurde das Verfahren wie folgt für GPRS-Verbindungen angepasst:

Um existierende PPP Client Software auf Endgeräten weiterverwenden zu können, wird auch für eine GPRS-Verbindung ein PPP Server benötigt. Die Software für den PPP Server wurde dazu direkt in das Datenmoduls integriert und bildet aus Sicht des Datenendgeräts die Schnittstelle zum Internet. Statt also eine PPP Verbindung zum Internet Service Provider herzustellen, endet die PPP Verbindung nun bereits im Datenmodul. Abb. 7.33 zeigt diese Konfiguration.

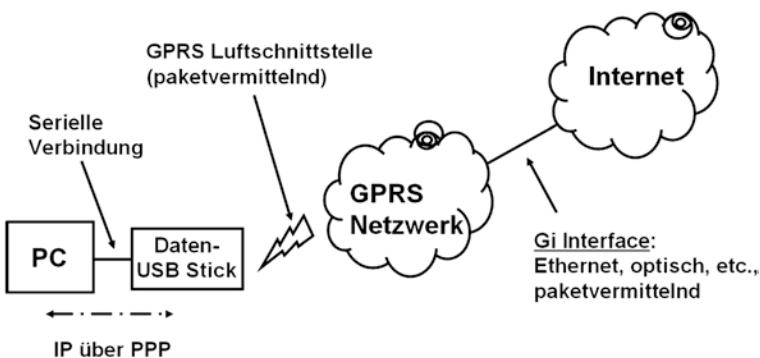


Abb. 7.33 Internetverbindung per GPRS und PPP

Der PPP Server im Datenmodul übersetzt die PPP Verbindungsaufnahme in eine Activate PDP Context Request-Nachricht und sendet diese an das GPRS-Netzwerk. Die IP-Adresse wird dann wie zuvor in Abb. 7.30 gezeigt zugeteilt.

Nach erfolgreicher Verbindungsaufnahme schickt der GGSN über die Activate PDP Context Accept-Nachricht eine IP-Adresse an das Datenmodul zurück. Von dort wird die IP-Adresse dem Datenendgerät per PPP zurückgegeben und der Verbindungsaufbau ist abgeschlossen.

Während der Verbindung (dem Packet Call) werden im Datenmodul die über das PPP-Protokoll eingehenden IP-Pakete in kleine GPRS-Datenpakete aufgeteilt und zum Netzwerk gesendet. In der umgekehrten Richtung werden die vom Netzwerk eingehenden kleinen GPRS-Datenpakete vom Datenmodul wieder in komplett IP-Pakete zusammengesetzt und danach über die serielle PPP Verbindung zum Datenendgerät weitergegeben.

Um dieser geänderten Konfiguration Rechnung zu tragen, sind folgende Einstellungen im Endgerät notwendig: Aus Sicht des Endgeräts ist das externe Datenmodul ein gewöhnliches Modem. Deshalb muss zunächst ein Standardmodemtreiber für die Schnittstelle konfiguriert werden, über die das Datenmodul angesprochen werden kann. Da der benötigte Standardmodemtreiber bei den unterschiedlichen Betriebssystemen mitgeliefert wird, ist der Anwender nicht auf einen speziellen Modemtreiber des Datenmodulherstellers angewiesen. Benötigt wird jedoch ein Treiber, der eine virtuelle serielle Schnittstelle über USB für das Endgerät zur Verfügung stellt.

Vor dem Aufbau der Verbindung muss dem Datenmodul der APN mitgeteilt werden, über die eine GPRS-Internetverbindung aufgebaut werden soll. Damit im Endgerät keine Softwareänderungen nötig sind, entschied man sich bei der Standardisierung dazu, den APN vor dem eigentlichen Verbindungsaufbau mit einem zusätzlichen AT-Kommando zu übergeben. Dies geschieht mit dem AT+CGDCONT Kommando. Um zum Beispiel eine IP-Verbindung über das GPRS-Netzwerk mit dem APN „internet.t-d1.de“ aufzubauen, ist folgendes AT-Kommando notwendig: „AT+CGDCONT=1, „IP“, „internet.t-d1.de““. Ist das Endgerät z. B. ein PC, wird dieses Kommando, wie in Abb. 7.34 gezeigt, in den „Erweiterten Einstellungen“ der Modemkonfiguration eingetragen.

Ein weiterer Unterschied zwischen einer leitungsvermittelten Internetverbindung und einer GPRS-Internetverbindung ist die Tatsache, dass für GPRS keine Telefonnummer gewählt werden muss. Statt der Telefonnummer wird deshalb der String *99***1# übergeben. Dieser String wird über das ATD-Kommando zum Datenmodul übergeben (ATD *99***1#). Nach Erhalt dieses Kommandos startet das Datenmodul daraufhin den PPP Server und verwendet den im AT+CGDCONT übergebenen APN für die PDP Context Activation-Prozedur. War die PDP Context Activation erfolgreich, liefert der PPP Server die in der PDP Context Activation Accept-Nachricht enthaltene IP-Adresse für das Datenendgerät zurück, sowie die IP-Adresse des DNS Servers für die Umwandlung von Domain-Namen (z. B. www.cm-networks.de) in IP Addressen. Ab diesem Zeitpunkt ist dann die Internetverbindung hergestellt.

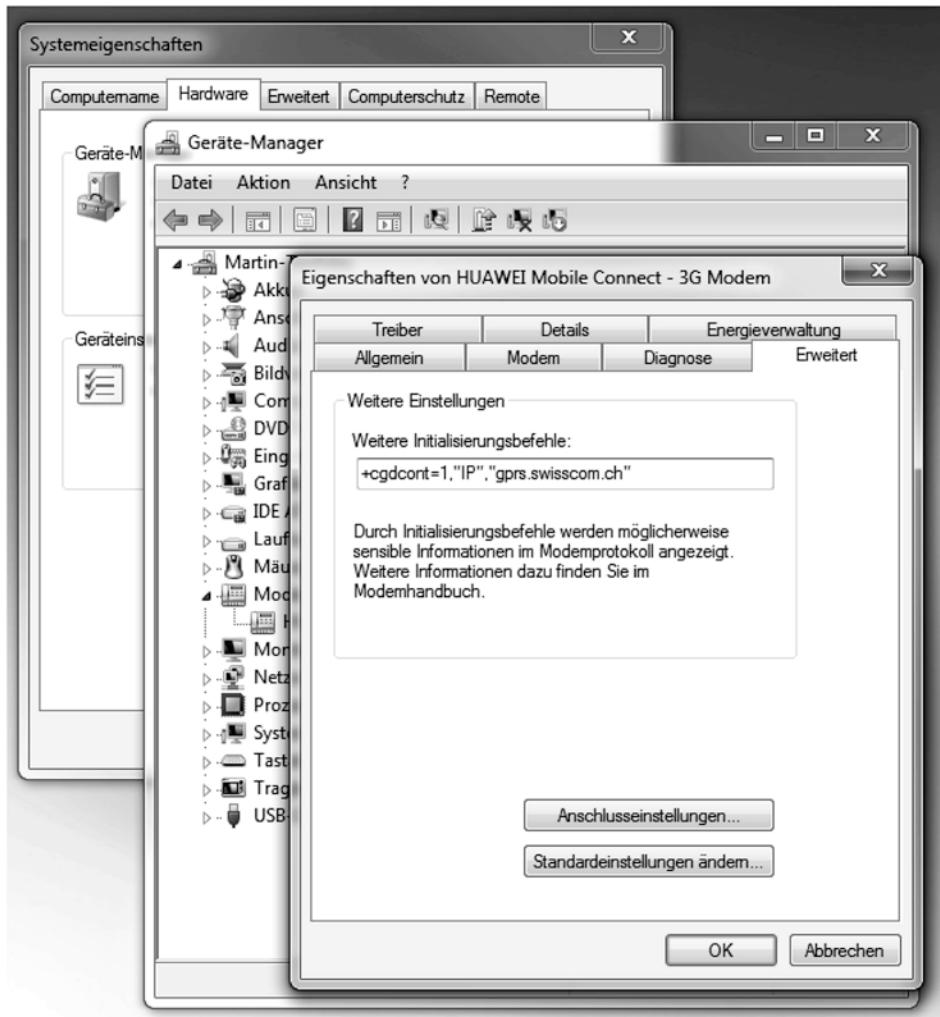


Abb. 7.34 APN als erweiterte Optionen bei der Modemkonfiguration

7.10 Fragen und Aufgaben

1. Welche Unterschiede gibt es zwischen leitungsvermittelter Datenübertragung und paketorientierter Datenübertragung?
2. Welche Vorteile bietet die GPRS-Datenübertragung gegenüber der früher üblichen GSM-Datenübertragung?
3. Warum gibt es unterschiedliche Coding Schemes?
4. Wie unterscheidet sich der GPRS Ready State vom GPRS Standby State?

5. Führt das Netzwerk bei GPRS einen Handover durch, wenn während eines Zellwechsels Daten übertragen werden?
6. Welche neuen Netzwerkelemente wurden mit GPRS eingeführt und welche grundsätzlichen Aufgaben haben diese?
7. Was ist ein Temporary Block Flow?
8. Welche Vorgänge finden bei einem Inter-SGSN Routing Area Update (IRAU) statt?
9. Warum kommt das IP-Protokoll auf dem Gn Interface zweimal im Protokollstack vor?
10. Wie wird erreicht, dass beim internationalen Roaming für GPRS im Ausland keine Einstellungen im Endgerät geändert werden müssen?
11. Was ist der Unterschied zwischen einem GPRS Attach und einer PDP Context Activation?
12. Welche Rolle spielt der Access Point Name (APN) bei der PDP Context Activation-Prozedur?

Lösungen sind auf der Website zum Buch unter <http://www.cm-networks.de> zu finden.

Stichwortverzeichnis

- 5G Anker, 160
5G NR Carrier Aggregation, 123
5G NR Standalone Architektur, 176
5G Standalone, 119, 176
802.11e-Standard, 330
802.11f-Standard, 272, 279
802.11n-Standard, 267, 292
802.11-Standard, 265
802.1x-Standard, 268, 313
8PSK-Prinzip, 474
- A**
- A3-Algorithmus, 414
A5-Algorithmus, 441
A8-Algorithmus, 441
AAS (Active Antenna Arrays), 135
Abis Interface, 427
Absolute Radio Frequency Channel Number, 167
Access Grant Channel, 426, 480
Access Management Function, 177
Access Point, 270
Access Point Name, 184, 494
Access Stratum, 161
Access Transfer Control Function, 238
Access Transfer Gateway, 238
ACK Frame, 281
ACL (Asynchronous Connection-Less), 347
Active Antenna Arrays, 134
Adaptive Frequency Hopping, 340, 343
Adaptive Multi Rate, 435
Adaptive Multi-Rate Narrowband, 227
Ad-hoc Mode, 269
Advanced Encryption Standard, 314
- AES (Advanced Encryption Standard), 314, 324
AFH (Adaptive Frequency Hopping), 340, 343, 356, 380
AGCH (Access Grant Channel), 426, 480
AID (Association ID), 280
Air Interface, 420
Always On-Modus, 469, 500
AMF (Access Management Function), 177
AMR-Algorithmus (Adaptive Multi Rate-Algorithmus), 435
AMR-NB (Adaptive Multi-Rate Narrowband), 227
AMR-WB (AMR Wideband Codec), 436
AMR Wideband Codec, 227, 436
Anklopfen, 412
ANM (Answer Message), 400
ANR (Automatic Neighbor Relation), 10, 74
Answer Message, 400
APDU (Application Protocol Data Units), 383, 459
APN (Access Point Name), 184, 494
Application Protocol Data Units, 383, 459
Application Server, 216
ARFCN (Absolute Radio Frequency Channel Number), 167
AS (Access Stratum), 161
AS (Application Server), 216
Asserted Identities, 232
Association ID, 279, 280
Asynchronous Connection-Less, 347
ATCF (Access Transfer Control Function), 238
ATGW (Access Transfer Gateway), 238
AUSF (Authentication Server Function), 178
Authentication Center, 413

- Authentication Server Function, 178
 Authentication Triplets, 413
 Authentifizierungsalgorithmus, 414
 Automatic Neighbor Relation, 10, 74
- B**
- Bandbreite, 468
 Bandwidth Part, 141
 Baseband-Prozessor, 452
 Baseband Unit, 120
 Base Station Controller, 428
 Base Station Subsystem, 269, 403, 417
 Base Station Subsystem Mobile Application Part, 401
 Base Transceiver Station, 419
 Basic Service, 411
 Basic Service Set, 269
 Basic Service Set ID, 273
 BBU (Baseband Unit), 120
 BCCH (Broadcast Common Control Channel), 425, 480
 Beacon Frames, 273
 Beamforming, 301
 Bearer Independent Call Control, 408
 Bearer Independent Core Network, 394
 BICC, 403
 Billing, 407
 Billing Record, 407, 448
 Block Acknowledgement, 334
 Broadcast Common Control Channel, 425, 480
 BSC (Base Station Controller), 428
 BSS (Base Station Subsystem), 269, 403, 417
 BSSMAP (Base Station Subsystem Mobile Application Part), 401
 BTS (Base Transceiver Station), 419
 Burst, 420
 BWP (Bandwidth Part), 141
- C**
- Call Detail Records, 487
 Call Hold, 412
 Call Waiting, 412
 Campus Netzwerk, 116
 CAP (XML Configuration Access Protocol), 236
 Carrier Aggregation, 64
 CCE (Control Channel Element), 27, 143
- CCK (Complementary Code Keying), 288
 CDR (Call Detail Records), 487
 Cell-ID, 445
 Cellular Internet of Things, 101
 Cell Update, 483, 499
 Channel Bonding, 302
 Channel Request, 428
 Channel State Information Reference Signale, 135
 CiIoT (Cellular Internet of Things), 101
 Ciphering-Algorithmus, 441
 Ciphering Key, 372, 441
 Clear to Send, 282
 Cloud Native, 199
 Coding Scheme, 473
 Combination Key, 368
 Combined Location Update, 478
 Common Channel, 423
 Complementary Code Keying, 288
 Component Carrier, 66
 Connection-Active-Zustand, 355
 Connection-Hold-Zustand, 355
 Connection-Park-Zustand, 356
 Connection-Sniff, 355
 Container, 199
 Control Channel Element, 27, 143
 Control Region Set, 143
 Convolutional Coder, 290, 439, 473
 CORESET (Control Region Set), 143
 CS-Fallback, 77
 CSI-RS (Channel State Information Reference Signale), 135
 CSMA/CA, 283
 CSMA/CD, 283
 CTS (Clear to Send), 282
 Cyclic Prefix, 19
- D**
- Data Network Name, 184
 DCF (Distributed Coordination Function), 283
 DCI (Downlink Control Information), 144
 DCNR (Dual Connectivity-New Radio Restriction), 170
 Dedicated Bearer, 223
 Dedicated Channel, 422
 Dedicated File, 458
 Dedicated Radio Bearer, 216
 Dedicated Traffic Channel, 139

- Delivery TIM, 281
Demodulation Reference Signal, 139
DF (Dedicated File), 458
DHCP (Dynamic Host Configuration Protocol), 271
DIFS (Distributed Coordination Function Interframe Space), 281
Digitaler Signalprozessor, 454
Direct Link Protocol, 268, 270
Direct Sequence Spread Spectrum, 287
Direct Transfer Application Part, 401
Discontinuous Reception-Zeit, 41, 55
Discontinuous Transmission, 442
Distributed Coordination Function, 283
Distributed Coordination Function Interframe Space, 281
Distributed Virtual Resource Block, 24
Distribution System, 279
DLP (Direct Link Protocol), 268, 270
DMRS (Demodulation Reference Signal), 139
DNN (Data Network Name), 184
DNS (Domain Name System), 500
Domain Name System, 500
Downlink Control Information, 144
Downlink Shared Channel, 139
DRX (Discontinuous Reception), 41, 55, 57
DSP (digitaler Signalprozessor), 454
DSS (Dynamic Spectrum Sharing), 118, 122, 151
DSSS (Direct Sequence Spread Spectrum), 287
DTAP (Direct Transfer Application Part), 401
DTCH (Dedicated Traffic Channel), 139
DTIM (Delivery TIM), 281
DTMF (Dual Tone Multi Frequency), 233, 397
DTX (Discontinuous Transmission), 442
Dual Connectivity-New Radio Restriction, 170
Dual-Tone Multi-Frequency, 233, 397
DVRB (Distributed Virtual Resource Block), 24
Dynamic Host Configuration Protocol, 271
Dynamic Spectrum Sharing, 118, 122, 151
- E**
Early-Media, 231
eCPRI (Evolved Common Public Radio Interface), 125
EDCA (Enhanced Distributed Channel Access), 331
- EDGE (Enhanced Data Rates for GSM Evolution), 474
EDR (Enhanced Data Rate), 342, 351
eDRX (Extended Idle Mode Discontinuous Reception), 109
EEPROM, 457
EF (Elementary File), 458
EFR (Enhanced Full Rate Codec), 435
EGPRS, 468, 474
Elementary File, 458
Embedded Multitasking Betriebssystem, 453
EN-DC Handover Prozedur, 171
Enhanced Data Rate, 340, 342, 351
Enhanced Data Rates for GSM Evolution, 474
Enhanced Distributed Channel Access, 331
Enhanced Full Rate Codec, 435
Enhanced Voice Services, 227
eNodeB, 6
Entfernung, 432
ePDG (evolved Packet Data Gateway), 254
eSCO, 340, 350, 356, 380
ESS (Extended Service Set), 271
ETSI (European Telecommunication Standards Institute), 395
European Telecommunication Standards Institute, 395
Evolved Common Public Radio Interface, 125
evolved Packet Data Gateway (ePDG), 254
EVS (Enhanced Voice Services), 227
Extended Idle Mode Discontinuous Reception, 109
Extended Service Set, 271
- F**
FACCH (Fast Associated Control Channel), 423
Faltungskodierer, 291, 439
Fast Assoiated Control Channel, 423
Fast Fourier Transformation, 19
FCCH (Frequency Correction Channel), 425
FDMA (Frequency Division Multiple Access), 420
FEC (Forward Error Correction), 348, 373
FFT (Fast Fourier Transformation), 19
FHS (Frequency Hop Synchronization), 352, 353
FHSS (Frequency Hopping Spread Spectrum), 343

- Final Block Indicator, 489
 Flusskontrolle, 485
 Forward Error Correction, 348, 373
 FR (Full Rate Codec), 435
 FR1 (Frequency Range 1), 118, 126
 FR2 (Frequency Range 2), 118, 126, 148
 Frame, 420
 Frame Aggregation, 293
 Frame Relay, 486
 Frequency Correction Channel, 425
 Frequency Division Multiple Access, 420
 Frequency Hopping Spread Spectrum, 343
 Frequency Hop Synchronization, 352, 353
 Frequency Range 1, 118, 126
 Frequency Range 2, 118, 126, 148
 Frequenzmultiplex, 420
 Fronthaul Interface, 125
 Full Rate Codec, 435
- G**
 Gateway GPRS Support Node, 487
 Gaussian Frequency Shift Keying, 342
 Gb Interface, 492
 General Object Exchange, 377
 General Object Exchange Profile, 377
 GFSK (Gaussian Frequency Shift Keying), 343
 GGSN (Gateway GPRS Support Node), 487
 Gi Interface, 494
 Globally Unique Temporary ID, 47
 GMM/SM (GPRS Mobility Management and Session Management), 487, 496
 G-MSC, 447
 gNB, 122
 Gn Interface, 492
 GOEP (General Object Exchange Profile), 377
 Gp Interface, 495
 GPRS Mobility Management, 487
 GPRS Mobility Management and Session Management, 487, 496
 GPRS Tunneling Protocol, 493
 Gr Interface, 494
 Gs Interface, 496, 499
 GSM (Global System for Mobile Communications), 401, 417
 GSM for Railways, 418
 GSM-R (GSM for Railways), 418
 GTP (GPRS Tunneling Protocol), 493
 Guard-Intervall, 294
- Guard Time, 421
 GUTI (Globally Unique Temporary ID), 47
- H**
 Half Rate Codec, 435
 Halten, 412
 HARQ (Hybrid Automatic Repeat Request), 140
 HARQ (Hybrid Automatic Retransmission reQuest), 27
 HARQ Feedback, 140
 HCI (Host Controller Interface), 357
 Heterogeneous Network, 124
 HLR (Home Location Register), 401, 410
 Home Location Register, 401, 410
 Host Controller Interface, 357
 HR/DSSS, 288
 HR (Half Rate Codec), 435
 Hybrid Automatic Repeat Request, 140
 Hybrid Automatic Retransmission reQuest, 27
- I**
 IAM (Initial Address Message), 399
 IAPP (Inter Access Point Protocol), 279
 Idle State, 481
 I-FFT (Inverse Fast Fourier Transformation), 18
 IMEI (International Mobile Equipment Identifier), 180
 IMSI (International Mobile Subscriber Identity), 179, 410, 456
 IN (Intelligent Network), 404
 Independent BSS, 269
 Infrastructure BSS, 270
 Initial Address Message, 399
 Inquiry, 353
 Inquiry Scan, 354
 Intelligent Network, 404
 Inter Access Point Protocol, 279
 Inter-MSC Handover, 451
 International Mobile Equipment Identifier, 180
 International Mobile Subscriber Identity, 179, 410, 456
 International Telecommunication Union, 395
 Internet Protokoll, 469
 Interrogating-CSCF, 216
 Intra BSC Handover, 450
 Inverse Fast Fourier Transformation, 18

- IP (Internet Protokoll), 469
IP Multimedia Subsystem, 215
IP-Short-Message-Gateway, 235
ISM-Frequenzband, 342
ISM-Kanäle, 274
ITU (International Telecommunication Union), 395
- K**
Kanalkodierer, 438
Kc (Ciphering Key), 372, 441
Ki, 413
Komprimierung von Sprachdaten, 434
Kubernetes, 200
- L**
L2CAP Layer, 360
LAA (License Assisted Access), 67
LAC (Location Area Code), 425, 498
LAPD (Link Access Protocol D-Channel), 427
Leistungsklasse, 344, 431
License Assisted Access, 67
Link Access Protocol D-Channel, 427
Link Controller, 353
Link Key, 367
Link Manager, 356
LLC Header, 285
Localized Virtual Resource Blocks, 24
Location Area, 445
Location Area Code, 425, 498
Location Area ID, 445
Location Area Update, 444
Logical Link Control and Adaptation Protocol, 360
LTE-Band, 5
Luftschnittstelle, 420
LVRB (Localized Virtual Resource Blocks), 24
- M**
M3UA, 402
Machine Type Communication, 101
Main File, 458
MAP (Mobile Application Part), 401, 495
Massive-MIMO, 133
Master-Slave Role Switch, 346
Maximum Ratio Combining, 301
- MCC (Mobile Country Code), 425, 498, 501
MCS (Modulation and Coding Schemes), 474
Media Gateway, 408
Message Transfer Part, 399
Message Waiting Flag, 416
MF (Main File), 458
Mikrokontrollersystem, 456
MIMO (Multiple Input Multiple Output), 30, 295
Mini-Slot Scheduli, 130
MME (Mobility Management Entity), 10
MMTel (Multimedia Telephony), 216
mmWave, 117, 148
MNC (Mobile Network Code), 425, 501
Mobile Application Part, 401, 495
Mobile Country Code, 425, 498, 501
Mobile Network Code, 425, 498, 501
Mobile Number Portability, 411
Mobile Station, 401, 452
Mobile Subscriber ISDN Number, 410
Mobile Switching Center, 399
Mobility Management, 405
Mobility Management Entity, 10
Modulation and Coding Schemes, 474
MS (Mobile Station), 401
MSC (Mobile Switching Center), 399
MSC-Server, 408
MSISDN (Mobile Subscriber ISDN Number), 410
MTP (Message Transfer Part), 399
Multimedia Telephony, 216
Multiple Input Multiple Output-Datenübertragung, 30, 294
Multislot, 470
- N**
N26-Schnittstelle, 193
Nachbarzellen, 420
NAS (Non-Access Stratum), 10, 161
Network Allocation Vector, 284
Network Operation Mode, 477
Network Slicin, 204
Network Subsystem, 404
Network Subsystem Access Point Identifier, 501
Netzmonitor, 444
NIDD (Non-IP Data Delivery), 111
NOM (Network Operation Mode), 477

Non-Access Stratum, 10, 161
 Non-IP Data Delivery, 111
 Non-Standalone Architecture, 116
 NR Carrier Aggregation, 123
 NSA (Non-Standalone Architecture), 116
 NSAPI (Network Subsystem Access Point Identifier), 501
 NSS (Network Subsystem), 404
 Numerology, 127

O

OBEX (General Object Exchange), 377
 OFDMA (Orthogonal Frequency Division Multiple Access), 18
 Option 2, 122
 Option 3x, 122
 Orthogonal Frequency Division Multiple Access, 18

P

PACCH (Packet Associated Control Channel), 479, 490
 Packet Associated Control Channel, 479, 490
 Packet Bursting, 334
 Packet Call, 500
 Packet Control Unit, 484
 Packet Data Convergence Protocol, 35
 Packet Data Network Gateway, 12
 Packet Data Protocol, 500
 Packet Data Traffic Channel, 470, 479
 Packet Filter Set, 187
 Packet Temporary Mobile Subscriber Identity, 497, 502
 Packet Timing Advance Control Channel, 479, 485
 Paging, 354, 445
 Paging Channel, 425
 Paging Nachricht, 428
 Pairing, 367
 PAPR (Peak to Average Power Ratio), 20
 PCC (Primary Component Carrier), 168
 PCF (Policy Control Function), 185
 PCH (Paging Channel), 425
 PCI (Physical Cell ID), 25, 166
 PCM (Pulse Code Modulated), 407
 P-CSCF (Proxy-CSCF), 216
 PCU (Packet Control Unit), 484

PDCP (Packet Data Convergence Protocol), 35
 PDN-Gateway, 12
 PDP (Packet Data Protocol), 499
 PDP Context Activation, 487, 500
 PDSCH (Downlink Shared Channel), 139
 PDTCH (Packet Data Traffic Channel), 470, 479
 Peak to Average Power Ratio, 20
 PFS (Packet Filter Set), 187
 Physical Cell ID, 25, 166
 Physical Layer Convergence Protocol, 287, 289, 291
 Physical Resource Block, 131, 139
 Physical Uplink Shared Channel, 140
 Piconetz, 344
 PIN, 368
 PLCP (Physical Layer Convergence Protocol), 287, 289, 291
 PLMN (Public Land Mobile Network), 405
 PMI (Pre-coding Matrix Indicator), 136
 Point to Point Protocol, 503
 Policy Control Function, 185
 Power Class, 344
 Power Save Mode, 102
 Power Save Multi Poll, 299
 Power-Saving Mode, 279
 PPP (Point to Point Protocol), 503
 PRB (Physical Resource Block), 131, 139
 Pre-coding Matrix Indicator, 136
 Precondition, 224
 Pre-Shared Key, 314
 Primary Component Carrier, 168
 Profil, 375
 Protocol Service Multiplexer, 361
 Provide Roaming Number, 447
 Proxy-CSCF, 216
 PSM (Protocol Service Multiplexer), 361
 PSMP (Power Save Multi Poll), 299
 PS-Poll Frames, 280
 PSTN (Public Standard Telephone Network), 404
 PTCCH (Packet Timing Advance Control Channel), 479, 485
 P-TMSI (Packet Temporary Mobile Subscriber Identity), 502
 Public Land Mobile Network, 405
 Public Standard Telephone Network, 404
 Pulse Code Modulated, 407
 Punktierung, 473
 PUSCH (Physical Uplink Shared Channel), 140

Q

QAM (Quadrature Amplitude Modulation), 21, 23
QCI (Quality of Service Identifier), 223
QoS (Quality of Service), 216, 268, 330
Quad-Band Mobiltelefone, 418
Quadrature Amplitude Modulation, 21, 23
Quadrature Phase Shift Keying, 23
Quality of Service (WLAN), 268
Quality of Service, 216, 330
Quality of Service Identifier, 223
Quantisierer, 407

R

RAB (Radio Access Bearers), 156
RAC, 498
RACH (Random Access Channel), 425, 480
Radio Access Bearers, 156
Radio Resource Control, 160
RAND, 384, 457, 497
Random Access, 29
Random Access Channel, 425, 480
RAN Notification Area, 201
RAU (Routing Area Update), 483, 499
Ready State, 482
Ready to Send, 282
Realtime Transport Protocol, 214, 223
Reassociation, 278
Received Signal Strength Indication, 353
Reduced Inter Frame Space, 299
Reframing, 151
Reference Signal Received Power, 204
Reference Signal Received Quality, 166, 204
REG (Resource Element Group), 143
Registration Notification Events, 219
REL (Release Message), 400
Relay MSC, 451
Release Complete Message, 400
Release Message, 400
Remote Radio Unit, 120
Resource Element Group, 143
RFCOMM, 364
RIFS (Reduced Inter Frame Space), 299
RISC-Prozessor, 452
RLC/MAC, 488
RLC (Release Complete Message), 400
RNA (RAN Notification Area), 201
Robust Header Compression, 226

RoHC (Robust Header Compression), 226
Routing Area Update, 483, 499
RRC (Radio Resource Control), 160
RRC-Connected, 156
RRC-IDle, 156
RRU (Remote Radio Unit), 120
RSRP (Reference Signal Received Power), 204
RSRQ (Reference Signal Received Quality), 166, 204
RSSI (Received Signal Strength Indication), 353
RTP (Realtime Transport Protocol), 214, 223
RTS (Ready to Send), 282

S

S1-Control Plane, 9
S5-Schnittstelle, 12
SA (5G Standalone), 119, 176
SACCH (Slow Associated Control Channel), 423
SAFER + (Secure And Fast Encryption Routine), 367
SCC (Secondary Component Carrier), 168
SCC-AS (Service Centralization and Continuity Application Server), 240
SCEF (Service Capability Exposure Function), 111
SC-FDMA (Single Carrier-Frequency Division Multiple Access), 20
SCH (Synchronization Channel), 425
SCO (Synchronous Connection Oriented), 349, 356
S-CSCF (Serving Call Session Control Function), 215
SCTP (Stream Control Transmission Protocol), 9, 402
SDCCH (Standalone Dedicated Control Channel), 425
SDP (Session Description Protocol), 214, 224
Secondary Component Carrier, 168
Secure And Fast Encryption Routine, 367
Security Mode 1, 374
Self-Organizing Networks, 74
Service Capability Exposure Function, 111
Service Centralization and Continuity Application Server (SCC-AS), 240
Service Set ID, 270
Serving Call Session Control Function, 215

- Serving-Gateway, 11
Serving GPRS Support Node, 485
Session Description Protocol, 214, 224, 229
Session Management, 486
Session Management Function, 177
SGSN (Serving GPRS Support Node), 485
S-GW (Serving-Gateway), 11
Short Interframe Space, 281
Short Message Service, 411
Short Message Service Center, 415
SIB (System Information Broadcast), 202
SIFS (Short Interframe Space), 281
Signaling Gateway, 403, 408
Signaling Radio Bearer, 161
Signal to Interference and Noise Ratio, 166
Signed Response, 371, 384, 457, 497
SIM Application Toolkit, 458
SIM Karte, 456
Single Carrier-Frequency Division Multiple Access, 20
Single Radio Voice Call Continuity, 238
Single-RAN, 73
SINR (Signal to Interference and Noise Ratio), 166
SIP (Session Initiation Protocol), 209
SIP Precondition, 224
SIP Proxy, 210
SIP Registrar, 211
SIP URI, 212
Slot Konfiguration, 138
Slow Associated Control Channel, 423
SMF (Session Management Function), 177
SMS (Short Message Service), 411
SMSC (Short Message Service Center), 415
Sounding Reference Signale, 136
Split Bearer, 157
Sprachdaten
 Komprimierung, 434
SRB-1 (Signaling Radio Bearer), 161
SRES (Signed Response), 371, 384, 457, 497
SRS (Sounding Reference Signale), 136
SRVCC (Single Radio Voice Call Continuity), 238
SSB (Synchronization Signal Block), 132
SSB IDs (Synchronization Signal Block Beam IDs), 133
SSID (Service Set ID), 270
Standalone Dedicated Control Channel, 425
Standby State, 483
Stealing Flags, 422
STM (Synchronous Transfer Mode), 396
Stream Control Transmission Protocol, 9, 402
Stromsparmodus, 279
Subcarrier, 19, 128
Subcarrier Spacing, 167
Subscriber Concealed Identity, 189
Subscription Concealed Identity, 180
Subscription Permanent Identifier, 179
Subsequent Handback, 451
Subsequent Inter MSC Handover, 451
SUCI (Subscriber Concealed Identity), 180, 189
SUPI (Subscription Permanent Identifier), 179
Supplementary Services, 411
Switching Matrix, 392
Synchronisation, 43
Synchronization Channel, 425
Synchronization Signal Block, 132
Synchronization Signal Block Beam IDs, 133
Synchronous Connection Oriented, 349, 356
Synchronous Transfer Mode, 396
SYS_INFO 13, 480
SYS_INFO, 425, 479
System Information Broadcast, 202
System On a Chip, 455
- T**
TAC (Tracking Area Code), 45
Tail, 422
TBF (Temporary Block Flow), 488
TCAP (Transaction Capability Application Part), 401
TCH (Traffic Channel), 423
TDD (Time Division Duplex), 126
TDMA (Time Division Multiple Access), 420
TDMA Frame, 420
TEID (Tunnel Endpoint ID), 49
Telephony Application Server, 219
Temporal Key Integrity Protocol, 314, 323
Temporary Block Flow, 488
Temporary Flow Identifier, 488, 502
Temporary Logical Link ID, 497, 502
Temporary Mobile Subscriber Identity, 425, 445
Terminating-Access Domain Selection, 242
TFI (Temporary Flow Identifier), 488, 502
TIM (Traffic Indication Map), 280

- Time Division Duplex, 126
Time Division Multiple Access, 420
Timeslot, 396, 420
Timing Advance, 432, 479
TKIP (Temporal Key Integrity Protocol), 314, 323
TLLI (Temporary Logical Link ID), 497, 502
TMSI (Temporary Mobile Subscriber Identity), 425, 445
Tonwahl, 397
Tracking Area, 56
Tracking Area Code, 45
Traffic Chanel, 423
Traffic Flow Templates, 223
Traffic Indication Map, 279
Training Sequence, 422
Transaction Capability Application Part, 401
Transcoding and Rate Adaptation Unit, 434
TRAU (Transcoding and Rate Adaptation Unit), 434
Tunnel Endpoint ID, 49
- U**
UART-Schnittstelle, 358, 364
UDM (Universal Data Management), 178
UDR (Unified Data Repository), 178
Um-Interface, 420
Unacknowledged Mode Data Radio Bearer, 222
Unified Data Repository, 178
Universal Data Management, 178
Universally Unique ID, 362
Universal Serial Bus, 358
UPF (User Plane Function), 178
Uplink State Flag, 473, 489
USB (Universal Serial Bus), 358
User Agent, 210
- User Plane, 485
User Plane Function, 178
USF (Uplink State Flag), 473, 489
UUID (Universally Unique ID), 362
- V**
Verbindungsmatrix, 392
Visitor Location Register, 409
VLR (Visitor Location Register), 409
Voice Activity Detection, 442
Voice over LTE, 209
Voice over Wifi, 253
VoLTE (Voice over LTE), 209
VoLTE Notrufe, 243
VoLTE Roaming, 245
VoLTE S8-Home Routing, 248
VoWifi (Voice over Wifi), 253
- W**
Wi-Fi Protected Setup, 324
Wireless Bridging, 273
Wireless LAN, 265
Wireless Protected Access, 313
WPA (Wireless Protected Access), 313
WPS (Wi-Fi Protected Setup), 324
- X**
X2-Schnittstelle, 10
XML Configuration Access Protocol, 236
- Z**
Zeitmultiplex, 420
Zeitschlitz, 396, 420