

MOBILE BACKHAUL

MOBILE BACKHAUL

EDITORS

Esa Metsälä

Juha Salmelin

Nokia Siemens Networks, Finland



A John Wiley & Sons, Ltd., Publication

This edition first published 2012
© 2012 John Wiley and Sons, Ltd

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

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

The right of the author to be identified as the author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

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

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

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

Library of Congress Cataloging-in-Publication Data

Mobile backhaul / editors, Juha Salmelin, Esa Markus Metsala.
p. cm.

Includes bibliographical references and index.
ISBN 978-1-119-97420-8 (cloth)

1. Mobile communication systems. I. Salmelin, Juha. II. Metsala, Esa Markus.
TK5103.2.M577 2012
621.39'81—dc23

2011051682

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

ISBN: 978-1-119-97420-8

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

Contents

Foreword	xv
Acknowledgements	xvii
List of Abbreviations	xix
List of Contributors	xxxi
1 Introduction	1
<i>Esa Metsälä, Juha Salmelin and Erik Salo</i>	
1.1 Why Read This Book	1
1.2 What is ‘Mobile Backhaul’	2
1.3 Targets and Scope of the Book	3
1.4 Organization of the Book	3
PART I MOBILE AND PACKET NETWORKS	
2 Mobile Backhaul and the New Packet Era	7
<i>Erik Salo and Juha Salmelin</i>	
2.1 Backhaul Network, Tiers and Costs	7
2.1.1 <i>Backhaul Network Tiers</i>	7
2.1.2 <i>Backhaul Network Costs Distribution</i>	8
2.2 Legacy Backhaul Networks	9
2.2.1 <i>Backhaul Basic Technologies</i>	9
2.2.2 <i>Backhaul Topology</i>	10
2.3 Drivers for the MBH Network Change	10
2.3.1 <i>Mobile Service Developments and Traffic Growth</i>	13
2.3.2 <i>Mobile Network Developments</i>	16
2.3.3 <i>Backhaul Cost-Efficiency Improvements</i>	18
2.3.4 <i>Lower Operational Costs</i>	19
2.3.5 <i>Developments in General Transport</i>	21
2.4 Packet Based Backhaul Networks	21
2.4.1 <i>Physical Network and Topology</i>	22
2.4.2 <i>Logical Network and Protocol Layers</i>	22

2.5	Making Transition to Packet Technology Networks	22
2.5.1	<i>Transition Strategies for Packet-Based Backhaul</i>	23
2.5.2	<i>Implementing Transition and Network Evolution</i>	27
3	3GPP Mobile Systems	29
<i>Esa Metsälä</i>		
3.1	3GPP	29
3.1.1	<i>Radio Technologies and Backhaul</i>	29
3.1.2	<i>Organization</i>	30
3.1.3	<i>Specifications</i>	31
3.1.4	<i>Releases</i>	32
3.2	2G	33
3.2.1	<i>Circuit Switched Traffic</i>	33
3.2.2	<i>Packet Switched Traffic</i>	36
3.2.3	<i>Abis</i>	37
3.3	3G	38
3.3.1	<i>Circuit Switched Traffic</i>	40
3.3.2	<i>Packet Switched Traffic</i>	41
3.3.3	<i>3G Air Interface Channels</i>	42
3.3.4	<i>FP, MAC and RLC Protocols</i>	43
3.3.5	<i>HSDPA (HS-DSCH) and HSUPA (E-DCH)</i>	45
3.3.6	<i>Iub</i>	48
3.3.7	<i>Iur</i>	50
3.3.8	<i>Iu-cs</i>	50
3.3.9	<i>Iu-ps</i>	51
3.3.10	<i>GTP-U Protocol</i>	52
3.4	LTE	54
3.4.1	<i>Architecture</i>	54
3.4.2	<i>Packet Switched Traffic</i>	56
3.4.3	<i>Air Interface</i>	58
3.4.4	<i>SI</i>	58
3.4.5	<i>X2</i>	59
3.4.6	<i>Bearers</i>	60
3.4.7	<i>Mobility Management</i>	61
3.4.8	<i>Interworking with 2G and 3G</i>	63
3.4.9	<i>Voice Support</i>	63
3.4.10	<i>Self Configuration and Self-Optimization</i>	64
3.5	Summary	64
	References	65
4	Packet Networks	68
<i>Esa Metsälä</i>		
4.1	Mobile Backhaul Application	68
4.1.1	<i>Backhaul Service</i>	68
4.1.2	<i>Access, Aggregation and Core</i>	70
4.1.3	<i>3GPP Guidance for the Backhaul</i>	71
4.1.4	<i>Networking and Backhaul</i>	72

4.2	Standardization	73
4.2.1	<i>IEEE</i>	73
4.2.2	<i>IETF</i>	74
4.2.3	<i>ISO</i>	74
4.2.4	<i>ITU-T</i>	74
4.2.5	<i>MEF</i>	75
4.2.6	<i>IP/MPLS Forum</i>	75
4.3	Physical Interfaces	76
4.3.1	<i>High Data Rates</i>	76
4.3.2	<i>Ethernet Ports</i>	77
4.3.3	<i>E1/T1/JT1</i>	77
4.3.4	<i>SDH/Sonet</i>	79
4.4	PPP and ML-PPP	80
4.4.1	<i>PPP over E1/T1/JT1</i>	80
4.4.2	<i>ML-PPP</i>	81
4.4.3	<i>PPP over Sonet/SDH</i>	83
4.5	Ethernet and Carrier Ethernet	83
4.5.1	<i>Carrier Ethernet</i>	84
4.5.2	<i>Ethernet and Ethernet Bridging</i>	85
4.5.3	<i>Ethernet Link Aggregation</i>	87
4.5.4	<i>VLANs</i>	87
4.5.5	<i>Class of Service</i>	88
4.5.6	<i>VLAN Example</i>	88
4.5.7	<i>Ethernet OAM</i>	89
4.5.8	<i>Provider Bridging</i>	91
4.5.9	<i>Provider Backbone Bridging</i>	92
4.5.10	<i>MPLS Based Carrier Ethernet</i>	92
4.6	IP and Transport Layer Protocols	92
4.6.1	<i>IP</i>	93
4.6.2	<i>IP Addresses and Address Assignment</i>	96
4.6.3	<i>Forwarding</i>	99
4.6.4	<i>Routing Protocols</i>	100
4.6.5	<i>Differentiated Services</i>	101
4.6.6	<i>Address Resolution Protocol</i>	102
4.6.7	<i>ICMP</i>	102
4.6.8	<i>UDP</i>	103
4.6.9	<i>RTP</i>	104
4.6.10	<i>TCP</i>	105
4.6.11	<i>SCTP</i>	107
4.6.12	<i>IPv6</i>	108
4.7	MPLS/IP Applications	109
4.7.1	<i>MPLS Architecture</i>	110
4.7.2	<i>Label Distribution Protocol</i>	111
4.7.3	<i>BGP</i>	111
4.7.4	<i>MPLS Ping</i>	113
4.7.5	<i>MPLS L3 VPN and MP-BGP</i>	113
4.7.6	<i>Pseudowire Emulation Edge to Edge</i>	116

4.7.7	<i>MPLS L2 VPN–VPLS</i>	118
4.7.8	<i>MPLS-TE</i>	121
4.7.9	<i>MPLS-TP</i>	123
4.8	Summary	123
	References	124
5	Backhaul Transport Technologies	128
<i>Jouko Kapanen, Jyri Putkonen and Juha Salmelin</i>		
5.1	Transport Systems	128
5.1.1	<i>OSI-Model</i>	128
5.1.2	<i>Access Schemes</i>	129
5.1.3	<i>Plesiochronous Digital Hierarchy (PDH)</i>	130
5.1.4	<i>Synchronous Digital Hierarchy (SDH)</i>	131
5.1.5	<i>SDH Protection</i>	134
5.1.6	<i>Optical Transport Hierarchy (OTH)</i>	135
5.1.7	<i>Next Generation SDH (NG-SDH)</i>	136
5.1.8	<i>Asynchronous Transfer Mode (ATM)</i>	137
5.1.9	<i>Hybrid TDM/Packet</i>	137
5.2	Wireless Backhaul Technology	138
5.2.1	<i>Radio Wave Propagation</i>	138
5.2.2	<i>Frequencies and Capacities</i>	141
5.2.3	<i>Network Topologies</i>	144
5.2.4	<i>Availability and Resiliency</i>	145
5.2.5	<i>Performance</i>	146
5.2.6	<i>Other Wireless Technologies</i>	148
5.3	Wire-Line Backhaul Technology	148
5.3.1	<i>DSL Technologies</i>	148
5.3.2	<i>Optical Technology</i>	150
5.3.3	<i>Ethernet Interfaces</i>	153
5.3.4	<i>Ethernet in the First Mile</i>	154
5.3.5	<i>DOCSIS</i>	154
5.4	Aggregation and Backbone Tiers	155
5.5	Leased Line Services for Mobile Backhaul	156
5.5.1	<i>Ethernet Services and SLA's (MEF)</i>	157
5.5.2	<i>Leased Ethernet Service Offering</i>	162
5.5.3	<i>IP as a Backhaul Service</i>	162
5.6	Summary	163
	References	163

PART II MOBILE BACKHAUL FUNCTIONALITY

6	Synchronization	167
<i>Antti Pietiläinen and Juha Salmelin</i>		
6.1	Cellular Networks Synchronization Requirements	167
6.1.1	<i>Frequency Accuracy</i>	167

6.1.2	<i>Time Accuracy</i>	168
6.2	Frequency Synchronization in TDM Networks	169
6.2.1	<i>Synchronization Architecture in TDM Networks</i>	169
6.2.2	<i>PDH</i>	170
6.2.3	<i>SDH/SONET</i>	171
6.2.4	<i>ATM</i>	172
6.2.5	<i>OTN</i>	172
6.3	Frequency Synchronization in Packet Networks	172
6.3.1	<i>ACR (Adaptive Clock Recovery)</i>	173
6.3.2	<i>NTP</i>	173
6.3.3	<i>PTP Protocol</i>	174
6.3.4	<i>ITU PTP Telecom Profile for Frequency Synchronization</i>	177
6.3.5	<i>Synchronous Ethernet</i>	179
6.3.6	<i>Chaining Different Synchronization Technologies</i>	180
6.3.7	<i>Summary of ITU Recommendations Related to Frequency Synchronization in Packet Networks</i>	180
6.3.8	<i>TICTOC</i>	181
6.4	Synchronization Metrics for TDM and Synchronous Ethernet	182
6.4.1	<i>Stability Metric MTIE</i>	182
6.4.2	<i>Relationship between TDM Wander Specification and Base Station Clock Accuracy</i>	184
6.4.3	<i>TDEV</i>	185
6.5	Packet Synchronization Fundamentals and Metrics	187
6.5.1	<i>The Principles of Packet Timing for Frequency Synchronization</i>	187
6.5.2	<i>Packet Delay Metrics for Frequency Synchronization</i>	192
6.5.3	<i>Two-way Messaging</i>	198
6.5.4	<i>Delay Jumps</i>	198
6.5.5	<i>Testing Packet Timing Slaves</i>	198
6.6	Rules of Thumb for Packet Timing Network Implementation	199
6.7	Time Synchronization	201
6.7.1	<i>GNSS Systems</i>	201
6.7.2	<i>PTP for Time Synchronization</i>	202
6.8	Conclusions	202
	References	203
7	Resilience	204
	<i>Esa Metsälä</i>	
7.1	Introduction	204
7.1.1	<i>Restoration and Protection</i>	204
7.1.2	<i>Recovery</i>	205
7.1.3	<i>Availability</i>	206
7.1.4	<i>MTBF and MTTR</i>	207

7.1.5	<i>Increasing Availability</i>	207
7.1.6	<i>Network Failures</i>	209
7.1.7	<i>Human Errors</i>	209
7.2	Native Ethernet and Resilience	210
7.2.1	<i>Ethernet Bridging</i>	210
7.2.2	<i>Spanning Tree Operation</i>	211
7.3	Carrier Grade Ethernet	214
7.3.1	<i>Carrier Ethernet</i>	214
7.3.2	<i>MEF Services</i>	214
7.3.3	<i>Ethernet OAM</i>	215
7.4	IP Layer	216
7.4.1	<i>VRRP</i>	216
7.4.2	<i>Load Sharing</i>	217
7.4.3	<i>Routing Protocols</i>	217
7.4.4	<i>OSPF</i>	218
7.4.5	<i>BFD</i>	222
7.4.6	<i>Further Topics</i>	223
7.4.7	<i>Loop Free Alternates</i>	224
7.5	MPLS Resilience	224
7.5.1	<i>Label Allocation</i>	224
7.5.2	<i>LDP Sessions</i>	226
7.5.3	<i>IP MPLS VPN</i>	226
7.5.4	<i>VPLS</i>	227
7.5.5	<i>MPLS TE and Fast Reroute</i>	228
7.5.6	<i>MPLS OAM</i>	229
7.5.7	<i>MPLS-TP</i>	229
7.5.8	<i>GMPLS Control Plane</i>	230
7.6	Resilience in the BTS Access	231
7.6.1	<i>BTS and BTS Site</i>	231
7.6.2	<i>BTS Access</i>	232
7.6.3	<i>IP Addressing</i>	232
7.6.4	<i>Active-Passive Ports</i>	234
7.6.5	<i>IP Load Sharing</i>	235
7.6.6	<i>Ethernet Link Aggregation</i>	236
7.6.7	<i>OSPF in the Access</i>	236
7.6.8	<i>Static Routes</i>	239
7.6.9	<i>First Hop Gateway Redundancy</i>	240
7.6.10	<i>Microwave Access Links</i>	240
7.6.11	<i>Attachment to a MEF Service</i>	241
7.7	Resilience in the Controllers and the Core Interface	244
7.7.1	<i>BSC and RNC and Their Site Solutions</i>	244
7.7.2	<i>VRRP Example</i>	244
7.7.3	<i>Signaling Resilience with SCTP Multihoming</i>	244
7.7.4	<i>Use of Multiple Core Network Nodes</i>	246
7.8	Summary	247
	References	248

8 QoS	250
<i>Thomas Deiß, Jouko Kapanen, Esa Metsälä and Csaba Vulkán</i>	
8.1 End User Service, Radio Network Layers and the Transport Layer Service	250
8.1.1 <i>Transport Layer Service</i>	251
8.1.2 <i>End-to-End QoS</i>	251
8.1.3 <i>Need for Backhaul QoS</i>	252
8.1.4 <i>QoS Alignment with Radio and Backhaul</i>	254
8.2 TCP and UDP as End User Transport Layer Protocols	255
8.2.1 <i>UDP</i>	256
8.2.2 <i>TCP</i>	256
8.2.3 <i>TCP Congestion Control</i>	257
8.2.4 <i>TCP Over Wireless</i>	262
8.3 DSCP, Traffic Class, and Priority Bits	263
8.3.1 <i>Differentiated Services</i>	263
8.3.2 <i>IPv6</i>	265
8.3.3 <i>Per-Hop Behaviours</i>	265
8.3.4 <i>Recommended Use of DSCPs and Treatment Aggregates</i>	266
8.3.5 <i>DSCP in IP Tunnels</i>	268
8.3.6 <i>Use of DSCPs for Mobile Backhaul</i>	268
8.3.7 <i>MPLS Traffic Class</i>	270
8.3.8 <i>IEEE802.1Q Priority Bits</i>	270
8.3.9 <i>VLANs</i>	273
8.3.10 <i>QoS with MEF Services</i>	273
8.4 Ingress and Egress Functions	275
8.4.1 <i>Ingress Classification and Policing</i>	275
8.4.2 <i>Single-Rate Two Color Policer</i>	276
8.4.3 <i>Two-Rate Three Color Policer</i>	276
8.4.4 <i>Egress Scheduling, Queue Management, and Shaping</i>	276
8.4.5 <i>Strict Priority Scheduler</i>	277
8.4.6 <i>Weighted Round Robin Scheduler</i>	277
8.4.7 <i>Weighted Fair Queuing</i>	277
8.4.8 <i>Combined Schedulers</i>	278
8.4.9 <i>Buffering</i>	279
8.4.10 <i>Tail Drop</i>	279
8.4.11 <i>Active Queue Management</i>	280
8.4.12 <i>Shaping</i>	280
8.5 2G	281
8.5.1 <i>Native PCM-Based Abis</i>	281
8.5.2 <i>Abis Over Pseudowire</i>	281
8.5.3 <i>Abis Example</i>	281
8.6 3G/HSPA	282
8.6.1 <i>Bearers and Their Attributes</i>	282
8.6.2 <i>Iub</i>	283
8.6.3 <i>Iub Example</i>	285
8.6.4 <i>Congestion Control in MBH</i>	288

8.6.5	<i>Congestion Control in HSPA Systems</i>	288
8.6.6	<i>HSDPA Congestion Control</i>	289
8.6.7	<i>HSUPA Congestion Control</i>	291
8.6.8	<i>Co-existence of Radio Networks</i>	292
8.7	LTE	293
8.7.1	<i>QoS Architecture</i>	293
8.7.2	<i>Packet Flows and Bearers</i>	294
8.7.3	<i>QoS Parameters</i>	296
8.7.4	<i>Admission Control</i>	297
8.7.5	<i>S1 Interface</i>	298
8.7.6	<i>S1 Example</i>	298
8.8	Summary	300
	References	301
9	Security	303
<i>Esa Metsälä and José Manuel Tapia Pérez</i>		
9.1	Security in 3GPP Mobile Networks	303
9.1.1	<i>Network Domain Security</i>	305
9.1.2	<i>2G</i>	308
9.1.3	<i>Abis, A and Gb</i>	308
9.1.4	<i>3G</i>	309
9.1.5	<i>Iub</i>	310
9.1.6	<i>Iu-cs, Iu-ps and Iur Interfaces</i>	310
9.1.7	<i>LTE</i>	311
9.1.8	<i>S1 and X2 Interfaces</i>	311
9.1.9	<i>Management Traffic</i>	312
9.2	Protection of the Backhaul	313
9.2.1	<i>Cryptographic Protection Compared to Other Protection</i>	313
9.2.2	<i>Leased Service and A Self-Deployed Backhaul</i>	313
9.2.3	<i>Traffic Separation</i>	314
9.2.4	<i>Ethernet Services</i>	314
9.2.5	<i>IEEE 802.1x and IEEE802.1ae</i>	316
9.2.6	<i>MEF</i>	316
9.3	IP Layer Protection	316
9.3.1	<i>IPsec</i>	316
9.3.2	<i>IPsec SA</i>	317
9.3.3	<i>IPsec ESP</i>	317
9.3.4	<i>IPsec AH</i>	318
9.3.5	<i>IKE Protocol</i>	320
9.3.6	<i>Anti-Replay Protection</i>	322
9.3.7	<i>Network Element Authentication</i>	324
9.3.8	<i>Firewalls and Access Control Lists</i>	329
9.3.9	<i>Network Control Protocols Protection</i>	330
9.4	IP Sec VPN Deployment	331
9.4.1	<i>Cell and Hub Site Solutions</i>	331
9.4.2	<i>IPsec Profiles</i>	332
9.4.3	<i>VPN Resilience</i>	333

9.4.4	<i>Fragmentation</i>	336
9.4.5	<i>IPsec and Quality of Service</i>	337
9.4.6	<i>LTE S1 and X2 Study Case</i>	340
9.5	Summary	344
	References	344
10	Packet Backhaul Solutions	346
	<i>Erik Salo and Juha Salmelin</i>	
10.1	<i>Creating a Packet Based MBH Solution</i>	346
10.2	<i>MBH Solution Starting Points</i>	347
10.2.1	<i>Hard Starting Points</i>	348
10.2.2	<i>Soft Starting Points</i>	348
10.3	<i>MBH Optimization Considerations</i>	349
10.3.1	<i>Economic Optimization</i>	349
10.3.2	<i>Technical Optimization</i>	350
10.3.3	<i>Optimization for a Particular Operator</i>	350
10.3.4	<i>Optimization for a Certain Region</i>	351
10.3.5	<i>Optimization for Flexibility</i>	351
10.3.6	<i>Optimization of Implementation</i>	351
10.4	<i>MBH Solution Alternatives</i>	352
10.4.1	<i>Enhancing SDH/Sonet Networks with NG-SDH/MSPP Equipment</i>	352
10.4.2	<i>Enhancing SDH/Sonet Networks with a Packet Overlay</i>	353
10.4.3	<i>Fully Packet Based Networks for MBH Backbone and Aggregation</i>	356
10.4.4	<i>Building Fully Packet Based MBH Access Network for New Base Stations</i>	357
10.4.5	<i>Building Fully Packet Based MBH Access Networks Area by Area</i>	360
10.4.6	<i>Other Possible Approaches/Strategies</i>	360
10.5	<i>Outsourcing the MBH Network or Parts of it</i>	360
10.5.1	<i>Economic Considerations</i>	361
10.5.2	<i>Strategic and Organizational Considerations</i>	361
10.5.3	<i>Technical Issues</i>	362
10.6	<i>Selecting MBH Access Solution for a Particular Case</i>	363
10.6.1	<i>MBH Solution for LTE in a Dense Urban Area (in a Developed Environment)</i>	364
10.6.2	<i>MBH Solution for Suburban Area for 3G + LTE (in a Developed Environment)</i>	366
10.6.3	<i>MBH Solution in a Rural Area for a New 3G Network</i>	367
10.7	<i>From the Selected MBH Solution to Detailed Network Plans</i>	368
10.8	Summary	369
11	Summary	370
	<i>Esa Metsälä and Juha Salmelin</i>	
Index		373

Foreword

In the last few years we have seen an explosion of traffic on mobile networks as more and more of our communication has become mobile. Today, many of the world's most advanced mobile networks struggle to meet the performance and cost requirements placed on them and they typically respond by investing heavily in technologies such as HSPA+ and LTE to increase the capacity of the air interface. What cannot be ignored in the end-to-end architecture of mobile networks is the transport that plays a major role in determining the overall performance and cost of such networks. It is in this context that hardly any title could be more topical than *Mobile Backhaul*.

The natural cycle of technology evolution and rapid replacement of legacy networks by more capable and efficient packet networks further complicate the transport strategies employed by service providers. As the benefits of past investments compete with the capacity and efficiency of new investments, the result is an environment where multiple technologies co-exist.

In addition, packet based communication poses technical challenges that are not straight forward to overcome. Not all packet technologies can be applied to mobile backhaul and a careful examination of the underlying technology is required to ensure the integrity of the overall system and its ability to meet specific requirements. Two relevant examples are quality of service and security. IP networks behave very differently in these two domains when compared with the legacy networks based on E1's. Additionally, mobile networks create specific issues for the backhaul. A relevant example here is the need for synchronization in a mobile environment in order to enable handover of users from one base station to another and prevent interference, however, IP networks were not originally designed with these requirements in mind and need modifications to handle such requirements appropriately. Similarly, in an IP environment the delivery of all packets is not guaranteed and this lack of guarantee poses challenges for mobile communication. Other technical and non-technical challenges also exist in building backhaul networks which when taken together, lead us to conclude that a common rule for building advanced backhaul networks capable of adequately handling many simultaneous requirements is impossible to find. At the end, every network has to be specified and deployed as a unique solution.

The book you're holding now does a great job in discussing these topics and reviewing and assessing technologies and possible solutions that are available either in the mobile network or in the packet network domains. Many innovations are known to be the result of combining multiple disciplines and this is exactly what the authors and editors of this book have done.

Hossein Moiin
Chief Technology and Strategy Officer
Nokia Siemens Networks
Espoo, Finland

Acknowledgements

The editors would like to acknowledge the great contribution of authors from our colleagues at Nokia Siemens Networks: Thomas Deiß, Jouko Kapanen, José Manuel Tapia Pérez, Antti Pietiläinen, Jyri Putkonen, Csaba Vulkán and, last but not least, Erik Salo, ‘the great old transport guru’ who has already fully served his official work career.

Especially great thanks to Esa Törmä, the other grand old man in the mobile backhaul area for his insight and thinking, which has served us for laying the foundation and structure for many of the topics we have addressed in this book.

For specific review comments and suggestions we would like to thank Heikki Almay, Damian Dalglish, Joachim Eckstein, Carl Eklund, Timo Liuska, Sanna Mäenpää, Olli Pekka Mäkinen, Jukka Peltola, Mohammedneja Rahmato, Konstantin Shemyak, Antti Toskala, Jouko Törmänen, Eugen Wallmeier and Roland Wölker.

Special thanks to Harri Holma and Antti Toskala for a very useful guideline, example, and hints on how to proceed with writing a technical book and for informing of the practicalities involved in this type of a project.

Also we like to thank the team at John Wiley & Sons and the co-operators for an easy editing process, and for the flexibility which we believe we used up to its full quota; and especially Mark Hammond, Sandra Grayson, Richard Davies, Sophia Travis, Prachi Sinha Sahay, Prakash Naorem and Sara Barnes.

We appreciate the patience and support of our families and our authors’ families during the writing periods, which often extended into late night and weekend time.

Yet, despite all the hours consumed, we very much enjoyed writing this book. We are grateful for comments and suggestions for improvements or changes that could be implemented in forthcoming editions of this book. The feedback is welcome at editors’ email addresses: esa.metsala@nsn.com and juha.salmelin@nsn.com.

List of Abbreviations

10GE	10 gigabit Ethernet
1GE	Gigabit Ethernet (IEEE 802.3 standard interface)
1x	(CDMA2000) 1 times Radio Transmission Technology
2G	2 nd generation mobile system, GSM
3G	3 rd generation mobile system, WCDMA
3GPP	3rd Generation Partnership Project
64-QAM	64-level Quadrature Amplitude Modulation
AAL2	ATM Adaptation Layer 2
AAL2SIG	AAL2 Signaling
AAL5	ATM Adaptation Layer 5
AC	Attachment Circuit
AC	Admission Control
ACH	Associated Channel Header
ACK	Acknowledgment
ACR	Adaptive Clock Recovery
ACR	Adaptive clock recovery
ADM	Add-drop Multiplexing
ADSL	Asymmetric Digital Subscriber Line
AF	Assured Forwarding
A-GPS	Assisted GPS
AH	Authentication Header
AIMD	Additive Increase Multiplicative Decrease
AIS	Alarm Indication Signal
AKA	Authentication and Key Agreement
AM	Acknowledged Mode
AMBR	Aggregate Maximum Bit Rate
AMR	Adaptive Multi-Rate
ANSI	American National Standards Institute
AP	Access Point

APS	Automatic Protection Switching
ARIB	The Association of Radio Industries and Businesses
ARP	Allocation and Retention Priority
ARP	Address Resolution Protocol
ARQ	Automatic Repeat Request
AS	Autonomous System
ATIS	The Alliance for Telecommunications Industry Solutions
ATM	Asynchronous Transfer Mode
ATM	Asynchronous Transfer Mode
AU	Administrative Unit
BC	Boundary Clock
BCH	Broadcast Channel
BCP	Best Current Practice
BE	Best Effort
BFD	Bidirectional Forwarding Detection
BGP	Border Gateway Protocol
BPDU	Bridge Protocol Data Unit
BS	Base station
BSC	Base Station Controller
BSS	Base Station Subsystem
BTS	Base Station
BW	Bandwidth
BWP	Bandwidth Profile
C/I	Carrier/Interference
CA	Certificate Authority
CAPEX	Capital Expenditure
CBR	Constant Bit Rate
CBS	Committed Burst Size
CBWFQ	Class-Based Weighted Fair Queuing
CC	Congestion Control
CC	Connectivity Check
CC	Cable Cut (metric)
CCM	Connectivity Check Message
CDMA	Code Division Multiple Access
CE	Customer Edge
CEPT	European Conference of Postal and Telecommunications Administrations
CES	Circuit emulation service
CESoP	Circuit Emulation Service over Packet

CESoPSN	Structure-aware time division multiplexed circuit emulation service over packet switched network
CF	Coupling Flag
CFM	Connectivity Fault Management
CFN	Connection Frame Number
CHAP	Challenge Handshake Authentication Protocol
CIR	Committed Information Rate
CM	Color Mode
CMP	Certificate Management Protocol
CN	Core Network
CoS	Class of Service
CP	Control Plane
CRC	Cyclic Redundancy Check
CRL	Certificate Revocation List
CRNC	Controlling RNC
CS	Circuit switched
CS	Class Selector
CWDM	Coarse Wavelength Division Multiplexing
CWND	Congestion Window
DCH	Dedicated Channel
DEI	Drop Eligible Indicator
Delay_Req	Delay request (message used in PTP)
Delay_Resp	Delay response (message used in PTP)
DHCP	Dynamic Host Configuration Protocol
DiffServ, DS	Differentiated Services
DL	Downlink
DoS	Denial Of Service
DPD	Dead Peer Detection
DRNC	Drift RNC
DSCP	DS Code Point
DSL	Digital Subscriber Line
DWDM	Dense Wavelength Division Multiplexing
E1	Basic bit rate of European PDH; 2,048 Mbit/s
E-BGP	External BGP
EBS	Excess Burst Size
ECM	EPS Connection Management
ECMP	Equal Cost Multipath
ECN	Explicit Congestion Notification
E-DCH	Enhanced DCH

EDGE	Enhanced Data rates for Global Evolution
EEC	Synchronous Ethernet equipment clock
EF	Expedited Forwarding
EGP	Exterior Gateway Protocol
EIGRP	Enhanced Interior Gateway Protocol
EIR	Excess Information Rate
E-LSP	Explicitly TC (Traffic Class)-encoded PSC (PHB Scheduling Class) LSP
EMM	EPS Mobility Management
eNB, eNodeB	E-UTRAN NodeB
EoC	Ethernet over Copper
EPL	Ethernet Private Line
EPLAN	Ethernet Privat LAN
EP-LAN	Ethernet Private LAN
EPS	Evolved Packet System
EP-Tree	Ethernet Private Tree
E-RAB	E-UTRAN Radio Access Bearer
ESMC	Ethernet synchronization messaging channel
ESP	Encapsulating Security Payload
Eth	Ethernet
ETSI	European Telecommunications Standards Institute
E-UTRAN	Evolved UTRAN
EVC	Ethernet Virtual Connection
EV-DO	Evolution-Data Optimized
EVPL	Ethernet Virtual Private Line
EVPLAN	Ethernet Virtual Private LAN
EVP-LAN	Ethernet Virtual Private LAN
EVP-Tree	Ethernet Virtual Private Tree
FACH	Forward Access Channel
FCS	Frame Check Sequence
FD	Frame Delay
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
FDV	Frame Delay Variation
FEC	Forwarding Equivalence Class
FLR	Frame Loss Rate
FLR	Frame Loss Rate
FP	Frame Protocol
FRR	Fast Reroute
FSN	Frame Sequence Number

FTP	File Transfer Protocol
G-ACH	Generic Associated Channel
GAL	General Associated Label
GBR	Guaranteed Bit Rate
GERAN	GPRS/Edge Radio Access Network
GFP	General Framing Procedure
GGSN	Gateway GPRS Support Node
GLONASS	Globalnaya navigatsionnaya sputnikovaya sistema or Global Navigation Satellite System
GM	Grandmaster clock (used in PTP)
GMPLS	Generalized MPLS
GNSS	Global navigation satellite system
GPRS	General Packet Radio Service
GPS	Global positioning system
GSM	Global System for Mobile communications, originally Groupe Spécial Mobile
GTP-C	GPRS Tunneling Protocol Control Plane
GTP-U	GPRS Tunneling protocol User Plane
GW	Gateway
H-ARQ	Hybrid Automatic Repeat Request
H-BWP	Hierarchical Bandwidth Profile
HDLC	High Level Data Link Control
HRM-1	Hypothetical reference model 1
HSCSD	High Speed Circuit Switched Data
HSDPA	High Speed Downlink Packet Access
HS-DSCH	High Speed Downlink Shared Channel
HSPA	High Speed Packet Access
HSS	Home Subscriber Server
HSUPA	High Speed Uplink Packet Access
HTTP	Hypertext Transfer Protocol
HW	Hardware
IANA	Internet Assigned Numbers Authority
I-BGP	Internal BGP
ICMP	Internet Control Message Protocol
IE	Information Element
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet engineering task force
IGP	Interior Gateway Protocol
IKE	Internet Key Exchange

IMS	IP Multimedia Subsystem
IMSI	International Mobile Subscriber Identity
IP	Internet Protocol
IP-CAN	IP Connectivity Access Network
IS-95	Interim standard 95
IS-IS	Intermediate System - Intermediate System
ISO	International Organization for Standardization
ITU	International Telecommunications Union
L1	Layer 1 (in the OSI protocol stack)
L2	Layer 2 (in the OSI protocol stack)
L2VPN	Layer 2 virtual private network
L3	Layer 3 (in the OSI protocol stack)
L4	Layer 4 (in the OSI protocol stack)
LACP	Link Aggregation Control Protocol
LAG	Link Aggregation Group
LAN	Local Area Network
LCP	Link Control Protocol
LDP	Label Distribution Protocol
LED	Light Emitting Diode
LLC	Logical Link Control
LLQ	Low Latency Queuing
L-LSP	Label-only-inferred PSC (PHB Scheduling Class) LSP
LOS	Loss of Signal; Line of Sight
LSA	Link State Advertisement
LSDB	Link State Database
LSP	Label Switched Path
LSR	Label Switch Router
LTE	Long Term Evolution (3GPP mobile network standard)
M	Master
M3UA	Message Transfer Part 3 (MTP3) User Adaptation Layer
MAC	Media Access Control
MAFE	Maximum average frequency error
MATIE	Maximum average time interval error
MBH	Mobile Backhaul
MBMS	Multimedia broadcast multicast service
MBR	Maximum Bit Rate
MDEF	Modified Allan Deviation
MDEV	Modified Allan deviation
MEF	Metro Ethernet Forum

MEP	Maintenance End Point
MGW	Media Gateway
MIB	Management information base
MIMO	Multiple Input Multiple Output
MIP	Maintenance Intermediate Point
ML-PPP	Multi Link-Point-to-Point Protocol
MME	Mobility Management Entity
MP	Management Plane
MP-BGP	Multi Protocol BGP
MPLS	Multi Protocol Label Switching
MPLS-TE	MPLS Traffic Engineering
MPLS-TP	MPLS Transport Profile
MRU	Maximum Receive Unit
MSP	Multiplex Section Protection
MSPP	Multiservice Provisioning Platform
MSS	Maximum Segment Size
MSTP	Multiple Spanning Tree Protocol
MTBF	Mean Time Between Failure
MTIE	Maximum Time Interval Error
MTTR	Mean Time To Repair
MTU	Maximum Transfer Unit
MWR	Microwave Radio
NAT	Network Address Translation
NDS	Network Domain Security
NE	Network Element
NG-SDH	Next Generation Synchronous Digital Hierarchy
NIC	Network Interface Card
NID	Network interface device
NLRI	Network Level Reachability Information
NMS	Network Management System
N-PE	Network-PE (Provider Edge)
NPV	Net Present Value
NRI	Network Resource Identifier
NTP	Network time protocol
O&M	Operation and Maintenance
OAM	Operation, Administration and Maintenance
OC1	Optical carrier 1, 51.84 Mbit/s
OC3	Optical carrier 3, 155.52 Mbit/s
OCSP	Online Certificate Status Protocol

OFDM	Orthogonal frequency-division multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OH	Overhead
OPEX	Operational Expenditure
opex	Operating expenditures
OSI	Open System Interworking
OSPF	Open Shortest Path First
OSSP	Organization specific slow protocol
OTDOA	Observed time difference of arrival
OTH	Optical Transport Hierarchy
OTN	Optical transport network
OUI	Organizationally Unique Identifier
OWAMP	One-way Active Measurement Protocol
PAP	Password Authentication Protocol
PB	Provider Bridging
PBB	Provider Backbone Bridging
PCC	Policy and Charging Control
PCEF	Policy and Charging Enforcement Function
PCH	Paging Channel
PCM	Pulse code modulation
PCP	Priority Code Point
PCRF	Policy and Charging Rules Function
PDCP	Packet Data Convergence Protocol
PDH	Plesiochronous Digital Hierarchy
PDN	Packet Data Network
PDU	Protocol Data Unit
PDV	Packet delay variation
PE	Provider Edge
PEC-B	Combined packet slave clock and packet master clock
PEC-M	Packet master clock
PEC-S	Packet slave clock
PFS	Perfect Forward Secrecy
P-GW	PDN Gateway (PDN GW)
PHB	Per-Hop Behaviour
PHP	Penultimate Hop Popping
PKI	Public Key Infrastructure
pktfilteredMTIE	Packet-filtered MTIE (maximum time interval error)
PLL	Phase-locked loop
PM	Performance Monitoring

PMIP	Proxy Mobile IP
PMP	Point-to-Multipoint
PMTUD	Path MTU Discovery
PON	Passive optical network
PoS	Packet over SONET
POTS	Plain Old Telephone Service
ppb	Parts per billion
ppm	Parts per million
PPP	Point-to-Point Protocol
pps	Packets per second, pulse per second
PRC	Primary reference clock
PRTC	Primary reference time clock
PS	Packet Switched
PSK	Pre-Shared Key
PSN	Packet Switched Network
PSTN	Public Switched Telephone Network
PTP	Precision time protocol
PTP	Point-to-Point
PW	Pseudowire
PWE	Pseudowire Emulation
PWE3	Pseudo Wire Emulation Edge-to-Edge
QCI	QoS Class Identifier
QoE	Quality of Experience
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
R&D	Research and Development
RACH	Random Access Channel
RAN	Radio Access Network
RB	Radio Bearer
RED	Random Early Detection
RF	Radio frequency
RFC	Request For Comments
RFQ	Request for Quotation
RIP	Routing Information Protocol
RLC	Radio Link Control
RNC	Radio Network Controller
RR	Route Reflector
RRM	Radio Resource Management
RSTP	Rapid Spanning Tree Protocol

RSVP	Resource ReserVation Protocol
RSVP-TE	Resource ReserVation Protocol-Traffic Engineering
RTCP	RTP Control Protocol
RTO	Retransmission Timeout Timer
RTP	Real Time Transport Protocol
RTT	Round Trip Time
S	Slave
SA	Security Association
SACK	Selective Acknowledgement
SAE	System Architecture Evolution
SAK	Secure Association Key
SAToP	Structure-agnostic time division multiplexing over packet
SCCP	Signalling Connection Control Part
SC-FDMA	Single Carrier FDMA
SCTP	Stream Control Transmission Protocol
SDF	Service Data Flow
SDH	Synchronous Digital Hierarchy
SEC	SDH equipment clock, SDH equipment slave clock
SEG	Security Gateway
SFN	Single-frequency network
SFP	Small Form Pluggable
SGSN	Serving GPRS Support Node
S-GW	Serving Gateway
SIM	Subscriber Identity Module
SLA	Service Level Agreement
SLS	Service Level Specification
SON	Self Optimizing Networks
SONET	Synchronous Optical Network
SPF	Shortest Path First
SPI	Scheduling Priority Indicator
SRB	Signaling Radio Bearer
SRNC	Serving RNC
STM	Synchronous transport module
SW	Software
SW	Switch
Sync	Synchronization (message used in PTP)
T1	Basic bit rate of US & Japanse PDH; 1,544 Mbit/s
TC	Traffic Class
TC	Transparent clock

TCO	Total costs of Ownership
TCP	Transmission Control Protocol
TDD	Time Division Duplex
TDEV	Time deviation
TDM	Time Division Multiplexing
TDMA	Time Division Multiple Access
TD-SCDMA	Time division synchronous code division multiple access
TE	Traffic Engineering
TE	Terminal Equipment
TEID	Tunnel Endpoint Identifier
THP	Traffic Handling Priority
TICTOC	Timing over IP connection and transfer of clock
TLS	Transport Layer Security
TLV	Type-Length-Value
ToP	Timing over Packet
TOS	Type of Service
TTI	Transport Time Interval
TTL	Time To Live
TWAMP	Two-way Active Measurement Protocol
UDP	User Datagram Protocol
UE	User Equipment (mobile terminal)
UL	Uplink
UNI	User-Network Interface
UP	User Plane
U-PE	User-PE (Provider Edge)
UTRAN	Universal Terrestrial Radio Access Network
WAN	Wide Area Network
VC	Virtual Channel
VC	Virtual Container
WCDMA	Wideband CDMA
WDM	Wavelength Division Multiplexing
VDSL	Very-high-bit-rate digital subscriber line
WFQ	Weighted Fair Queuing
WiMAX	Worldwide Interoperability for Microwave Access
VLAN	Virtual local area network
VLR	Visitor Location Register
VoD	Video on Demand
VoIP	Voice over IP
VoLTE	Voice over LTE

VPLS	Virtual Private LAN Service
VPN	Virtual Private Network
VPWS	Virtual Private Wire Service
WRED	Weighted RED
VRF	Virtual Routing and Forwarding
WRR	Weighted Round Robin
VRRP	Virtual Router Redundancy Protocol
VSI	Virtual Switch Instance
X2	Interface in a LTE network
XPIC	Cross Polarisation Interference Cancellation

List of Contributors

Thomas Deiß

Nokia Siemens Networks
Düsseldorf, Germany

Jouko Kapanen

Nokia Siemens Networks
Espoo, Finland

Esa Metsälä

Nokia Siemens Networks
Espoo, Finland

José Manuel Tapia Pérez

Nokia Siemens Networks
Espoo, Finland

Antti Pietiläinen

Nokia Siemens Networks
Espoo, Finland

Jyri Putkonen

Nokia Siemens Networks
Espoo, Finland

Juha Salmelin

Nokia Siemens Networks
Espoo, Finland

Erik Salo

Independent Consultant
Espoo, Finland

Csaba Vulkán

Nokia Siemens Networks
Budapest, Hungary

1

Introduction

Esa Metsälä, Juha Salmelin and Erik Salo

1.1 Why Read This Book

Several textbooks exist either on mobile networks or on (packet) networking, but separately, and they usually consider their subjects in isolation from each other. However, no mobile network exists without a related transport network connecting the elements; and also, mobile networks are becoming more and more important ‘customers’ for many kinds of packet networks.

Therefore this book is about considering these two domains together, and about looking at mobile network and backhaul network interactions, and how these two domains should take each other into account, particularly in the new era of (fully) packet-based transport solutions.

Mobile backhaul, as shown in Figure 1.1, is at the intersection of a mobile network and a transport network. Some aspects are more closely related to the radio network. Another area originates from the transport and networking side.

Usually mobile networks, radio interfaces, and other radio related topics are discussed within a circle of radio communication experts, without considering so much the other parts of the whole network. As an example, the 3GPP view of the transport connection between any two mobile network elements is a single, straight line. This very high level of abstraction serves focusing on the mobile network specific issues. However, when transport connections in real life are more complex, and in the era of packet networks very much more complex, different types of issues start to appear; functionalities and especially performance of the mobile network are impacted. The influence can in some cases be significant.

Correspondingly, it is not that easy for a networking expert to delve into the details of mobile network – even the fundamental concepts may be hidden into a number of mobile network standards. Also, 3GPP mobile networks, and their radio interfaces are not at all like a wireless LAN: their protocols are not based on Ethernet (which one could feel familiar with), and there is more of protocol layering and a division of functionality between mobile elements; simply put, they are more complex.

Therefore, if your background is in radio communication and mobile network, you will benefit from having an understanding of how the backhaul is built and how it influences the

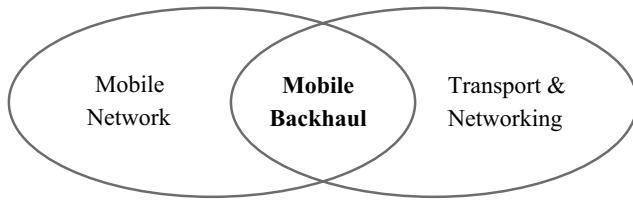


Figure 1.1 Mobile backhaul.

actual behavior and performance of the mobile network. For example, end-user bit rates are not limited only by the radio interface, but also by backhaul links, and thus it is becoming more important to understand and take into account the backhaul solutions applied.

On the other hand, if you are an IP and networking expert, you can use your competence more effectively when you understand more of the internal workings of the radio network side, and the basic requirements of radio networks relating to connections. Even if mobile backhaul is not the main driver for the networking industry, innovative solutions are needed to provide economic connections for advanced mobile networks and their services and data volumes, for example to cope with the high peak rates of HSPA + and LTE networks.

1.2 What is ‘Mobile Backhaul’

Figure 1.1 already provides the first answer: mobile backhaul unites a mobile network with transport/packet networks. Some elements and functions of the mobile network are within the scope, and the rest of functionality and characteristics are coming from the transport and packet networking side. Ultimately it is the mobile network that serves the end-users, however, the deployment and design of the mobile backhaul impacts not only the mobile element interfaces, but also contributes to mobile network’s overall operation and performance.

Mobile networks themselves are already well established in very many parts of the world, and mobile networks continue to expand, covering wider and wider areas of the globe. They also develop at a rapid rate and offer more and more services, including many kinds of wideband services, and enable higher and higher bit rates between the terminals and the network. This means that especially the data traffic is growing very fast in many mobile networks. Therefore well-working inter-element connections are necessary for the mobile networks to operate properly, and the role of supporting transport and packet networks is increasing. These transport and packet networks serving mobile networks are called ‘mobile backhaul networks’, or often just ‘mobile backhaul’ (MBH), as they connect a large number of base station sites to a limited number of centralized sites (see Figure 1.2).

The MBH networks are presently experiencing a big change, as the growth of mobile data traffic and development of packet transport technologies and equipment has created a strong push to use packet-based MBH solutions, both to increase feasible data throughputs and to improve the cost-efficiency of MBH networks.

Backhaul networks have always been an important part of the overall mobile network business case; *connections to the base station sites are important*, as their number is very high. These ‘last mile links’ (or ‘first mile links’, depending on your point of view) influence significantly the overall network costs. Now when network capacities increase and cell sizes

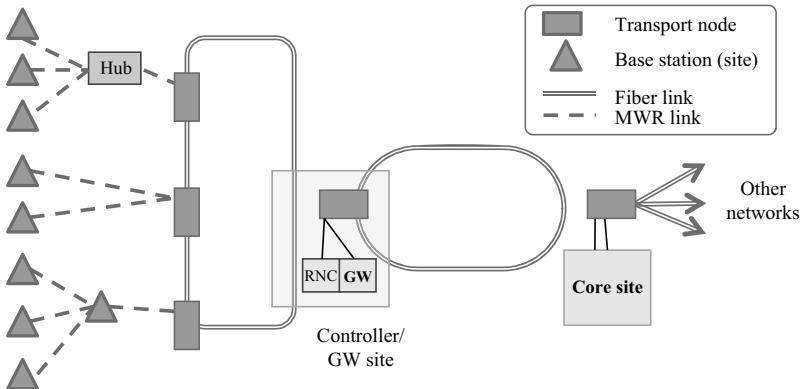


Figure 1.2 Mobile backhaul (MBH) network connects the sites of a mobile network.

decrease, transport share of the overall network costs tends to increase. Packet-based transport solutions help here, in keeping transport cost increase at a reasonable level.

1.3 Targets and Scope of the Book

This book is intended to give an overview of different aspects of mobile backhaul networks, and also provide a more detailed discussion on protocols, functionalities and technologies on both the radio network side and on the backhaul and networking technologies.

By nature, some terminology will be more 3GPP and radio-oriented, while a part comes from the networking world.

The book covers the mobile backhaul networks from the base station sites up to the core sites; however, it puts more emphasis on network segments closer to the base station sites (access tier, see Figure 2.2), as these parts have more mobile specific characteristics and are also economically the most important part of the backhaul.

Upper MBH network tiers often serve a combination of mobile and fixed traffic, are more built based on fixed traffic requirements and have a smaller influence on the mobile network economy; however, it is important to take into account also their impact on mobile network performance. Backhauls for indoor solutions are often a mix of fixed and mobile traffic and are beyond the scope of this book.

Technically, the book covers networking (or transport and transmission) related functionalities. Radio network protocols and key functionalities are reviewed as the radio network is the client layer for the mobile backhaul. While reading the backhaul oriented chapters it is useful to keep the basic mobile network architecture and operation in mind as this helps to identify interactions that are of a more subtle nature.

1.4 Organization of the Book

The body of the book is organized in two parts.

Part I considers networks as entities, from needs and change drivers to network transitions and from mobile systems to packet networking and implementation aspects. Part II studies key

functionalities in MBH: Synchronization, Resilience, Quality of Service, and Security, with the aim of going deeper into each of these topics.

The first chapter in Part I, Chapter 2, provides an introduction to the backhaul networks, to the needs and economic aspects of transport in mobile networks, and discusses the drivers for the packet-based MBH solutions as well as some transition issues. Part I continues with Chapter 3 describing mobile systems standardized by 3GPP. The emphasis is on logical interfaces and the related protocol stacks for the transport and for the end user service delivery. Radio network key functionality is introduced as well.

Chapter 4 in turn provides an overview of packet networks and networking technologies and protocols especially for readers who are already more familiar with the radio network technologies. Chapter 4 also discusses how the packet technologies are used in implementing a backhaul service for the radio network layer.

The last chapter of Part I, Chapter 5, discusses transport technologies and systems used in MBH networks, their main characteristics and briefly the services available for outsourcing MBH functionality; the focus in Chapter 5 is on the systems needed and used in the MBH access tier.

Part II starts with Chapter 6 discussing an important and very mobile network specific transport topic, namely provision of synchronization for mobile base stations over the transport network – a topic of increasing importance when we move towards packet-based MBH networks.

Chapter 7 addresses resilience. When moving to the packet network, carrier grade resilience is needed. Failure types in the packet network differ from those experienced in TDM networks, which easily causes a concern unless the topic is addressed. Recovery after a failure in the packet network as well typically relies on different methods than TDM does.

Chapter 8 is about quality of service (QoS) in the backhaul, focusing on QoS needs of all traffic types existing in the backhaul. Also, the role of transport in the overall end-to-end quality is discussed. QoS is one of the topics which directly and concretely links the radio network layer with the backhaul layers. It is often the first topic mentioned when discussing common areas between the radio network and backhaul experts.

And then Chapter 9 discusses security in the MBH networks and various networking solutions. With packet-based mobile backhaul, new types of threats emerge, and these need to be addressed. Cryptographic protection with IPsec is one of the tools for protecting the backhaul.

Chapter 10 provides an overview on how a packet-based MBH solution for a particular mobile network case is found and put together, including some examples of possible MBH solutions (solution types) for specific mobile network development cases.

Chapter 11 then contains a brief summary of the book.

Part I

Mobile and Packet Networks

2

Mobile Backhaul and the New Packet Era

Erik Salo and Juha Salmelin

2.1 Backhaul Network, Tiers and Costs

Mobile backhaul (MBH) networks serve mobile networks by providing connections between mobile network elements located in different geographical sites; no mobile network exists without a related MBH network. The main task of a MBH network is to connect a very large number of mobile network base station sites to a relatively small number of central sites where the mobile network core elements are located.

Basically a MBH network transfers transparently mobile system internal traffic and signaling between the mobile system elements. Even if mobile traffic is not interpreted by the MBH network, its properties affect mobile traffic in several ways, and thus the MBH network has a significant influence on the mobile network end-to-end quality. It is important to take these dependencies and influence fully into account when designing mobile networks which are optimized in respect of total cost and end-to-end performance.

To give some idea of the MBH network overall structure, one can note that a mobile network typically contains thousands, sometimes even tens of thousands, of base station sites (cell sites) while the number of core sites varies from 2 to 10 sites in smaller networks and is up to some tens in very large networks. In addition there is often, depending on the geographical area covered and on the operator network strategy, a number of intermediate sites, where some concentrating mobile network elements are located (typically BSCs, RNCs and/or various gateways) – these sites are often called simply controller or gateway sites.

2.1.1 Backhaul Network Tiers

Because of the high number of base station sites backhaul networks use transport layer traffic concentration on several points along the path from a base station site to the core site. Real physical transport links are shared across as many base station sites as possible.

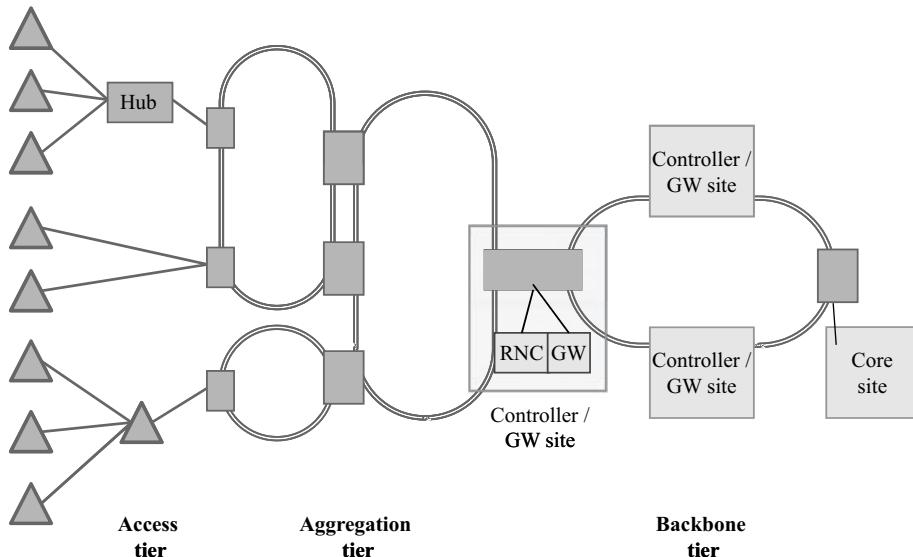


Figure 2.1 Backhaul network tiers (tier naming as used in this book).

Due to different network planning and optimization criteria in central and peripheral parts of the backhaul networks, it is useful to divide these networks into different domains or tiers. Mobile operators use varying partitioning and naming practices for these network tiers; in this book we speak about **access, aggregation and backbone network tiers** as shown in Figure 2.1.

2.1.2 Backhaul Network Costs Distribution

Generally transport capacities are highest in the backbone tier, but the number of links and nodes there is not very high. Backbone networks usually carry combined fixed and mobile traffic using very high capacity links and nodes where the cost per transferred bit is the lowest in the whole network. Thus, because of the low cost per bit and limited number of nodes and links, backbone's share of the backhaul total costs is usually relatively modest.

The opposite is true for the MBH access tier – link capacities are smaller (in relation to other backhaul tiers), but the number of links is very big, corresponding to the number of geographically separate base station sites. In addition, these links often serve just the mobile network, without cost sharing with other services. Thus the access tier contribution to the backhaul costs is high, typically significantly higher than the cost of other tiers. In addition, access tier share of the costs tends to grow bigger when the mobile network cells become smaller.

Aggregation tier falls between these both in link capacities and in the number of links. The aggregation tier is also quite often shared by mobile and fixed traffic, and then the link and node capacities are quite high also in this tier – and costs are shared between the services. The total cost of the MBH aggregation tier is usually much smaller than that of the access tier but still bigger than costs of the backbone tier, due to the higher number of links.

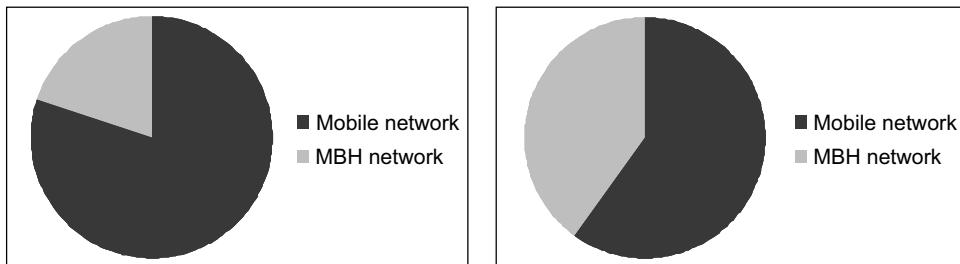


Figure 2.2 Typical MBH's share in mobile operator total network related costs; on the left for 2G/3G networks covering urban and suburban areas, on the right for mobile networks either covering large rural areas or having high numbers of small cells (in practice most networks are somewhere in between).

In all, backhaul transport forms a significant part of mobile operators' costs. Depending on the area, size and density of the mobile network, and organization of the MBH network and its maintenance, the MBH network's share can be 10...40% of the entire network related mobile operator costs (Figure 2.2).

Therefore cost efficiency is very important in all MBH networks – cost optimization is important in the backbone tier, more important in the aggregation tier and most important in the access tier; often the access tier represents more than 70...80% of all the backhaul network costs (Figure 2.3).

Presently, growth of mobile traffic, especially mobile data traffic, and also the increase in number of cell sites in high traffic areas tend to further increase cost of the backhaul networks. Transition to packet-based transport and other means to improve cost-efficiency of the MBH networks are discussed later in this chapter.

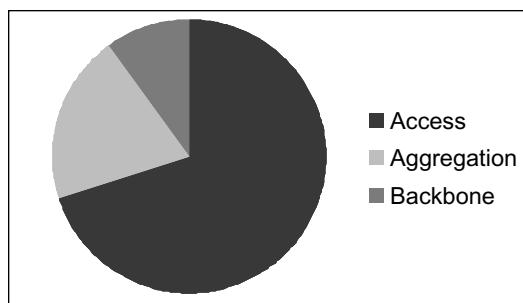


Figure 2.3 Typical cost distribution among MBH network tiers.

2.2 Legacy Backhaul Networks

2.2.1 Backhaul Basic Technologies

Most existing backhaul networks have been built to serve 2nd generation mobile networks (e.g. GSM or CDMA networks), or to serve combined 2nd and 3rd generation mobile networks (e.g. for the GSM and WCDMA networks in the same area). In addition there are a few backhaul networks serving just 3rd generation mobile networks.

These MBH networks are often built based on TDM transport technologies, i.e. mainly based on PDH and SDH/Sonet transport equipment, especially on lower tiers of the networks. Backhaul connections with these technologies are of fixed bandwidth, and they can be changed only by node reconfigurations locally or, with newer equipment, remotely by a network management system. Transport capacities in these networks are 1.5 or 2 Mbit/s or their multiples, 8 Mbit/s, 34 and 45 Mbit/s, and in upper tiers of the networks 155 Mbit/s and its multiples (620 Mbit/s, 2.5 Gbit/s etc) – the TDM bit rates are discussed in more detail in Chapter 5.

Some backhaul networks built specifically for 3rd generation mobile networks (WCDMA) also contain some ATM equipment e.g. for traffic concentration, especially on the aggregation and backbone network tiers. This equipment can improve capacity utilization efficiency compared to TDM, but on the other hand it adds another network layer to be managed and maintained, thus increasing network operational costs. ATM equipment has similar physical interfaces as PDH and SDH equipment (e.g. bit rates of 1.5 or 2, 34 or 45, and 155 and 620 Mbit/s).

More recently built or upgraded mobile backhaul networks also contain packet technologies, especially in the backbone tier and in the aggregation tier, but sometimes also in the access tier. Growing data traffic in mobile networks has increased the use of packet transport also in older backhaul networks, and in the near future growth of mobile data will make transition to the packet-based MBH networks necessary in many more mobile networks, if not in all of them, as will be discussed later in this chapter.

2.2.2 *Backhaul Topology*

As the 2nd and 3rd generation mobile networks do not have direct interconnections between the base stations, logical topology (traffic topology) of the backhaul transport networks is in these networks always a pure star: traffic goes directly from each base station to its controller.

Physical topology of a backhaul transport network usually is, however, very different – it is based on economic optimization of transport links and nodes as well as on the need to have resilience at least on the upper layers of the network. Physical topology is also very much the result of network history and gradual evolution, as with time more and more base station sites have been added to the network.

Often, the physical topology has tree and chain shapes near the base stations, as economy requires traffic concentration and minimization of the length of the individual (non-shared) access links. In upper network tiers the physical topology is based more on rings, as existence of alternative transport paths is here more important for the overall network resilience: failures affect a greater number of base stations. On upper tiers also the costs can be divided for a large number of base stations, making alternative routes economically more feasible. An example of MBH network physical topology (in the lower network tiers) is shown in Figure 2.4.

2.3 Drivers for the MBH Network Change

Higher and More Bursty Transport Load

Major traffic load for the 2nd generation mobile networks was, and is, voice – and that is often carried in a low bit rate format. For example, in GSM networks only 16 kbit/s is typically used per voice channel between the base station and the controller sites. Thus transport capacity required per base station site in these networks is very modest, in rural areas often only a

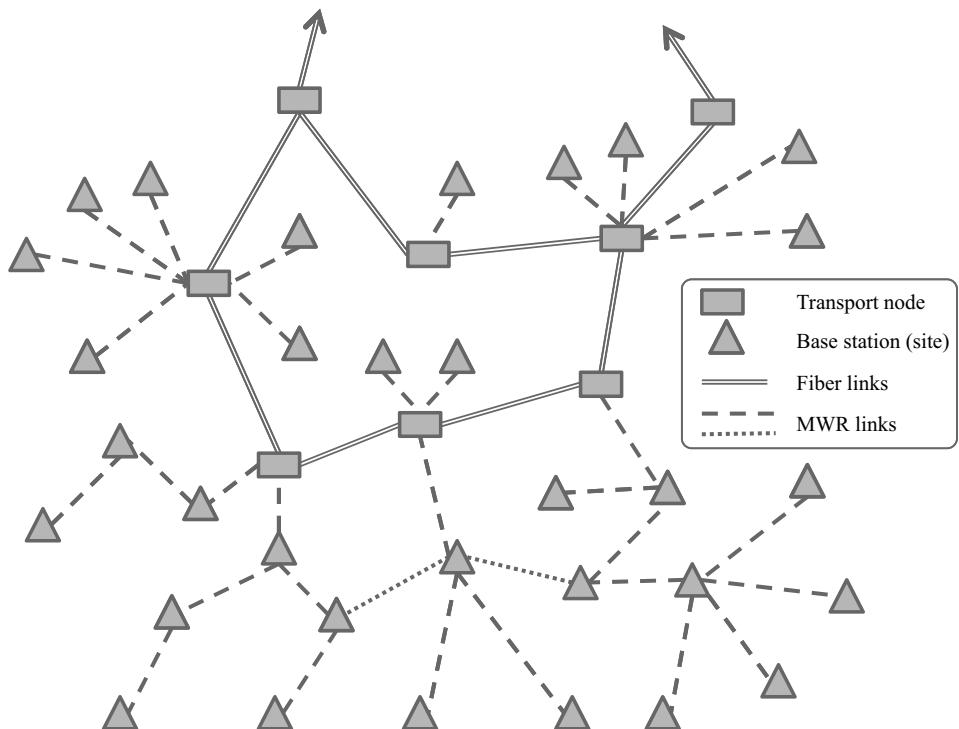


Figure 2.4 An example of MBH physical topology (access and lower aggregation tiers).

fraction of a 2 Mbit/s link, and also in many city areas just a few 2 Mbit/s lines are needed ($n \times 2 \text{ Mbit/s}$, $n = 1 \dots 4$). In this type of network it is often possible to build such a high backhaul transport capacity that it is possible to carry all the traffic base stations are able to provide.

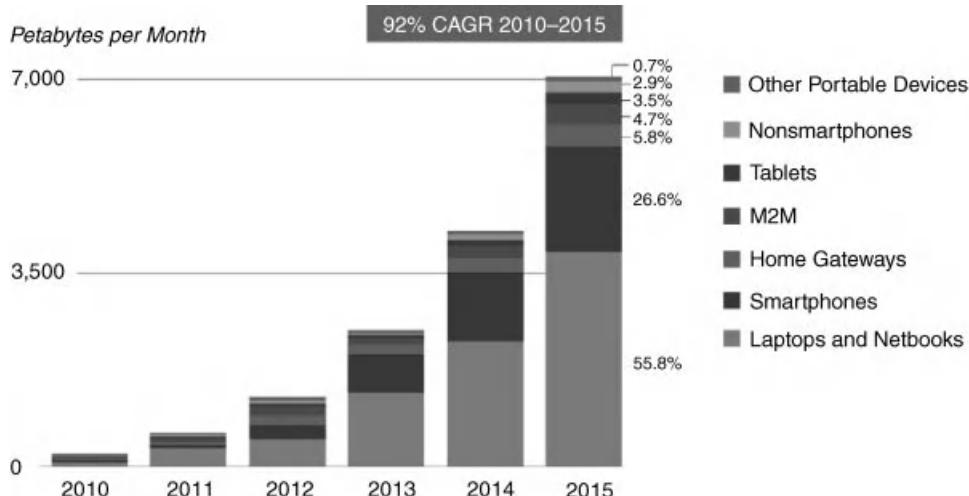
However, backhaul capacity needs are changing dramatically, and has already changed a lot: mobile data is growing fast and is already the major application in many 3rd generation mobile networks. And the growth of data traffic is just accelerating (see Table 2.1). This is supported by the further increase of mobile network capacities with new technologies and network generations, such as HSPA and HSPA + and LTE.

Increasingly, this new capacity is used for data intensive services and for video (especially where flat tariffs are applied), and the role of laptop PCs and new types of terminals is important here (see Figure 2.5). Additionally, often the instantaneous peak rates increase much more than the average traffic volume.

All this means that much greater total traffic will be present also in the backhaul links. Thus the transmission capacity required in the backhaul links for a base station site, especially in urban environments, is increasing strongly, and the actual load is much burstier than earlier. More and more often the MBH throughput will be critical for the actual mobile service capabilities, and for end-user experienced mobile service quality. If the backhaul capacities are not properly dimensioned initially, and later expanded in time, the backhaul network can become a bottleneck for the service quality and even for its reliable delivery.

Table 2.1 Examples of mobile data growth in 2010 (Source: Cisco VNI, 2011).

Region	Mobile Operator and Content Provider Examples
Korea	<ul style="list-style-type: none"> From mid-2009 to mid-2010, KT recorded a 344% increase in 3G mobile data traffic, SK Telecom's traffic grew 232%, and LG's traffic grew 114%. KT expects a 49-fold increase in mobile device traffic from 2009 to 2012, but plans to offload 40 percent of this traffic.
Japan	<ul style="list-style-type: none"> Softbank's mobile traffic grew 260% from 1Q 2009 to 1Q 2010, according to estimates by HSBC. KDDI expects mobile data traffic to grow 15-fold by 2015. NTT DoCoMo's mobile data traffic grew 60% from year to year.
China	<ul style="list-style-type: none"> China Unicom's 3G traffic increased 62% in a single quarter from Q1 to Q2 of 2010.
France	<ul style="list-style-type: none"> SFR's mobile data traffic has tripled each year since 2008.
Italy	<ul style="list-style-type: none"> Telecom Italia delivered 15 times more mobile data traffic in 2010 than in 2007.
Europe	<ul style="list-style-type: none"> Vodafone's European mobile data traffic increased 115% from 1Q 2009 to 2Q 2009, and 88% from 2Q 2009 to 2Q 2010. TeliaSonera expects mobile data traffic to double each year for the next 5 years.
United States	<ul style="list-style-type: none"> AT&T reports that traffic grew 30-fold from 3Q 2009 to 3Q 2010.
Global	<ul style="list-style-type: none"> Google reports that the number of YouTube videos delivered to mobile devices tripled in 2010, reaching 200 million video views per day.

**Figure 2.5** Example of mobile traffic forecasts (Source: Cisco VNI, 2011).

More Cells Sites

Often there is a need, in addition to increasing capacities of individual cells, to increase the number of cells to accommodate the fast growing traffic, especially in city centers, business districts and other high traffic areas (hot spots). These smaller urban cells with high traffic potential need short distance high capacity transport links, with high flexibility in network

building in these urban environments. Also lower backhaul costs than for earlier ‘big cell sites’ are now required (see next points).

Revenue Per Bit Decreasing

Another trend, happening at the same time as traffic grows, is the reduced revenue per bit of the mobile services. While many new services and traffic types mean ten or hundred times higher data volumes, end user ability and willingness to pay more for these new services is limited. The worst case is with the above mentioned flat rate services where even a strong traffic increase does not create related additional revenues. Thus even if the operator revenues for all the services increase, revenue per delivered bit can be just one tenth or even 1/100 of what it used to be. Thus cost efficiency is essential in the backhaul transport solutions and this is clearly becoming more and more important.

Lower Operational Costs

A third trend, related to the overall cost efficiency, is the need to operate all networks with smaller costs and more optimized organizations. In mobile backhaul networks this puts strong requirements for network simplification and automation of network operations to the widest extent possible. For example, self-healing properties of the networks reduce the need for immediate actions and thus contribute to reducing network maintenance costs (in addition to improving the service quality).

Network simplification in turn requires, for example, that the number of different technologies used within the network is reduced – as wide a use as possible of similar technologies is preferable. This creates scale benefits in equipment purchasing and savings especially in the operation and maintenance of the network, e.g. type of different skills needed is reduced. Also the effort needed for network planning is reduced when there are fewer different technologies and network layers.

Developments in General Transport

A significant change driver is also the development going on in general transport networks. The costs of packet-based transport solutions are now clearly lower than those of similar capacities with TDM or ATM technologies, and often power consumption of new packet-based equipment is also significantly lower than that of legacy equipment for similar capacities. In addition, the R&D efforts in the industry are focused on packet-based equipment, and their technical performance, power-efficiency and cost-efficiency continue to improve. Thus there is a strong need to benefit from these developments in the mobile backhaul solutions as well.

The developments of general transport networks also mean that leased line offerings are changing: in the longer term packet-based leased connections will have much lower price tags and more offerings and later also better geographical coverage than conventional leased lines, especially in case of higher capacity connections.

2.3.1 Mobile Service Developments and Traffic Growth

2.3.1.1 Traffic Forecasts Needed for Proper MBH Design

Mobile services developments drive traffic changes and determine what also needs to be carried over the MBH network. Evolution in the use of various mobile services has been extensively discussed in the literature, and is a hot topic in many research reports and

newsletters. Several types of forecasts on expected mobile services usage and on expected revenues per service are presented. These kinds of forecasts are obviously essential from the mobile operator business case point of view.

However, from the MBH network design point of view, detailed distribution of mobile network usage for various types of mobile services is not so fundamental. Instead the total amount of traffic expected from each cell is essential, as well as the distribution of this traffic among the major traffic classes (e.g. voice/data, real time/non-real time, delay critical/not delay sensitive etc). This information is fundamental for the dimensioning of the MBH network, and it is also very important for some of the technical requirements of MBH, like for the delays allowed within a MBH network.

Therefore, even if exact service forecasts are usually not needed for the MBH network planning, total mobile traffic forecasts are essential, as well as forecasts for a coarse division of this traffic into various traffic classes. This is an input that is necessarily needed for an economic design of the MBH network; building a whole MBH network according to the maximum capabilities of new mobile systems is very rarely economically justified, as in many areas and in many cell sites real mobile traffic will only slowly (if ever) reach those mobile systems technical limits. Obviously there are also cell sites where these limits will be reached much more quickly; therefore mobile traffic forecasts should be available separately for different type of areas. And it can be emphasized that these traffic forecasts should be exactly for the mobile operator and mobile network for which the MBH network will be designed – more general traffic forecasts made, e.g. for a country, do not necessarily lead to an optimal MBH design.

It shall, however, be also noted that during a forecast period quite rapid changes can happen, often increasing expected traffic volumes. This can happen locally, caused by some new buildings or shops or service points within the base station coverage area, or more network wide due to some new services that get rapid popularity. This requires some margins in the backhaul dimensioning or flexibility for rapid upgrades, whichever is more economic in each case.

2.3.1.2 Traffic Peak Rates

The peak bit rates created or needed by various mobile services are obviously very important for the MBH design, as these rates must be supported also in all parts of a MBH connection – otherwise in reality those rates will not be available for the end-user. Many mobile services can adapt to the (transport) bit rate available, and can work over slower or faster connections, but clearly with different end-user experience; for example, many applications use TCP family protocols and increase bit rate until the network limit is reached. Certain mobile services are also adaptive but need a minimum bit rate to work well, e.g. many video services. And then there are also fixed bit rate services, e.g. video distribution to several users simultaneously. In all cases the whole MBH connection needs to support at least the single end-user peak bit rate.

A question about these service peak rates is very often a commercial question – due to competitive reasons, mobile operators want to promise certain available peak bit rates. Then this promised peak bit rate becomes the network design target instead of actual service related requirements or forecasts. From the MBH network point of view, it can have very significant cost implications whether the same peak rate is promised over all the mobile network area or not; for example, if it is promised also in all suburban and rural areas, all the MBH links in these areas must have a minimum capacity greater than the promised peak rate.

2.3.1.3 Average Service Bit Rate

Usage of broadband services is crucial from the total mobile network traffic volume point of view. Most important high volume service in many mobile networks is simple broadband access to internet; that can generate very high traffic volumes. When there are no fees based on usage (i.e. flat fee tariff) or other contractual or technical limitations, monthly cumulative traffic volumes can become very big in the case of ‘heavy net users’.

However, from the MBH point of view, most important is the average loading created during a typical ‘busy hour’, i.e. when a high number of users are simultaneously actively using their network access. Such a number can be much more difficult to forecast than the peak rates discussed above, as it depends on the expected number of simultaneous (broadband) users in a cell and on how coincident their traffic actually is. In the case of very bursty traffic (e.g. typical access to web pages) the probability of simultaneous peak loads is relatively low, but in the case of more continuous bit streams (e.g. large file downloads or video services) the probability of simultaneous high bit rates becomes rather high.

In spite of the difficulties in forecasting a ‘busy hour’ average load or an average bit rate of all users in a cell, at least some coarse forecasts should be attempted. Building a MBH network based on the maximum average bit rates which a mobile system can support is usually unnecessarily expensive – this particularly applies in the early phases of new mobile networks able to support high instantaneous bit rates (e.g. in early phases of a LTE network).

2.3.1.4 Traffic Distribution Into Classes

Distribution of traffic into classes with different requirements is also important, as the real time services are more demanding on connection capacity than other services; constant bit rate services are especially demanding. For the MBH design quite coarse distribution is satisfactory, mainly an estimate is needed about the share or volume of non-adaptive services; an example of a traffic forecast made for MBH planning is shown in Figure 2.6. These, especially some types of video services, need to be taken into account with a higher average loading than the more flexible services.

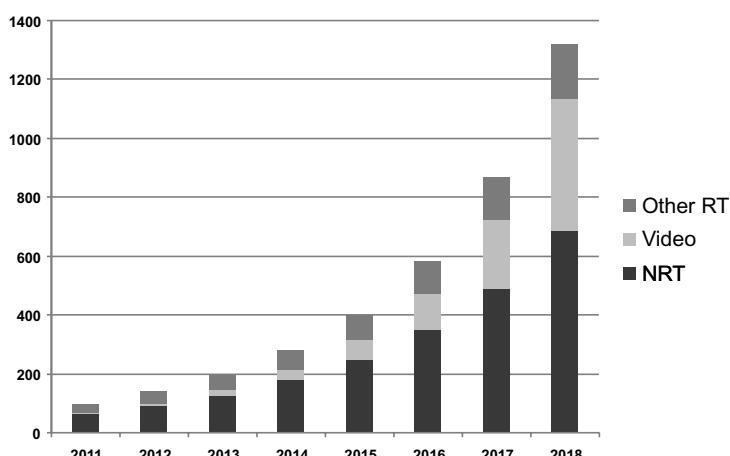


Figure 2.6 Example of mobile traffic forecast made for MBH planning.

2.3.2 *Mobile Network Developments*

Mobile service developments discussed above significantly change the traffic that needs to be carried over the backhaul network. In addition, there are mobile network changes, especially architectural changes (described more in detail in Chapter 3), which influence the backhaul network tasks and requirements. Some of the mobile network developments which are more important from the MBH network point of view are briefly discussed here.

2.3.2.1 More Native Packet Traffic

Total traffic volume expected from a base station site is strongly increasing with new mobile services, as discussed earlier, especially in urban base station sites. In addition, increasing portions of this traffic will be natively packet-based traffic, as new broadband services are all packet based. From the MBH point of view, packet traffic is also increasing because the new base stations have packet interfaces either for the majority of traffic (data traffic) or for all traffic (e.g. LTE base stations).

2.3.2.2 Flatter Mobile Network Architecture

New mobile networks, e.g. LTE, will have different mobile network architecture: there are no more controllers between the base stations and core elements. Therefore the transport domain becomes more unified, logical connections are directly from the base station to the core, and there will be more concentration points and other network nodes based on transport equipment.

On the other hand, some mobile core network elements may be distributed, placed closer to the base stations in the physical network, to increase mobile network capacity and to reduce delays. In such cases different logical connections may end at different sites, making transport configuration more complex.

2.3.2.3 Base Stations May be Connected to Several Core Sites

Another change made possible in the new architecture is that base stations may be connected to different core nodes and sites for redundancy reasons, even more than two different sites. Logical connections are then arranged from the base station site to several core sites, and traffic may move between those connections quite quickly, or even suddenly in the case of a link or a node failure.

2.3.2.4 Direct Links Between New Base Stations

Still a relatively new character having significant backhaul transport influence is that in the new architectures there are also logical connections directly between the base stations (e.g. in LTE architecture so called X2 interface, see Figure 2.7). In the transport domain such links can be implemented in different ways, either just logical paths using the same physical structure as before, or by adding new ‘transverse’ links to shorten the length of such connections. However, there are strong economic limitations to adding too many additional physical links into a

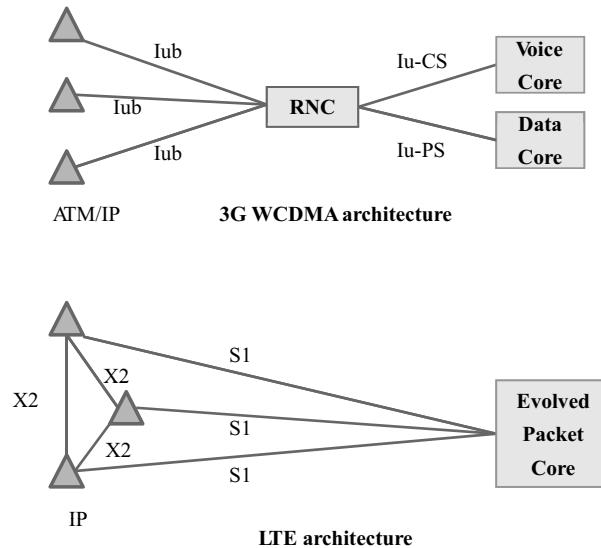


Figure 2.7 An example of mobile network architecture change, 3G WCDMA architecture vs. new LTE architecture (Note: new X2 interfaces between the base stations).

backhaul network, and the practical solution is a compromise between ‘traditional’ tree and chain topology (often already built topology) and a more meshed topology (more direct physical links).

2.3.2.5 More Urban (Hot Spot) Cell Sites

In addition to increasing capacities of individual cells, there is a need to increase the number of cells to accommodate the fast growing traffic, especially in city centers, business districts and other high traffic areas (hot spots). These smaller urban cells with high traffic potential need high capacity short distance transport links. Also lower backhaul costs than for earlier ‘big cells’ are required.

2.3.2.6 Connecting Temporary Cells

Special events, like mass sport competitions or big outdoor music events, are likely to concentrate a lot of people into a small area and many of them expect to have their normal net connections during the event, especially during pauses. This may require arrangement of additional cells to increase the mobile network capacity for the duration of the event. Temporary cells may also be required after different kinds of catastrophes (e.g. earthquakes and floods).

These temporary cells need to have a connection into the regular network, i.e. to be connected to the permanent MBH network at a suitable node point. This in turn means that the backhaul network needs to have enough flexibility to make such connections possible and also that it must be capable of carrying the additional traffic offered.

2.3.3 Backhaul Cost-Efficiency Improvements

Both the service developments (influencing traffic volumes and types) and the mobile network developments (architectural changes) put their requirements on the MBH network design and implementation. In addition, there is a strong need to reduce the costs of backhaul networks, both relating to the investment costs for a certain capacity as well as considering the operational costs of the whole backhaul network.

2.3.3.1 Lower Costs for Higher Backhaul Capacities

High capacities required for the backhaul connections, especially in heavy traffic areas, and the need to keep network investments under control are together a big challenge for the design and dimensioning of backhaul networks.

Packet-based transport equipment usually has significantly lower costs for a certain transport capacity, i.e. lower costs per bit/s. This applies both for optical transport and microwave transport and even more to separate ‘traffic concentration’ nodes needed in the backhaul networks (e.g. packet switching equipment vs. TDM cross-connect nodes). This lower-cost-per-bit of packet-based transport equipment partly helps in avoiding a situation where network costs linearly follow the capacity requirements.

Another method is to try to increase sharing of physical links, either within the operator’s own network or by sharing physical transport with another operator(s). A transport link of, say 1 Gbit/s capacity, is significantly cheaper than two separate links of half the capacity. Sharing within the network can be influenced by the transport topology design; sharing with other networks is obviously a matter of finding similar transport needs and then negotiations and agreements. Sharing can be especially useful in access links where a small number of cells do not necessarily use the link capacity efficiently but it must be dimensioned according to the peak capacity.

2.3.3.2 Better Ability to Handle Highly Bursty Traffic

The increasing peak-to-average ratio of the user traffic is also a challenge for an economic implementation of backhaul transport connections. The high peak rates easily speak for correspondingly high transport capacities; however, traffic burstiness can also be utilized for transport efficiency improvements in packet networks where there are no hard capacity allocations per user. The shorter the peaks are, the less is the probability of simultaneous demands, and thus a higher amount of statistical multiplexing can be assumed. Possible statistical gains also increase when the number of expected end users of a link increases – thus in a backhaul network higher statistical gains can be assumed above first traffic aggregation points.

2.3.3.3 Optimized Backhaul Dimensioning

Two previous points together mean that an effective tool to combat investment cost escalation is a good *backhaul design and dimensioning strategy*: the backhaul network capabilities and capacities are developed according to real service and business-based needs and not so much according to the theoretical maximum throughputs of the base stations.

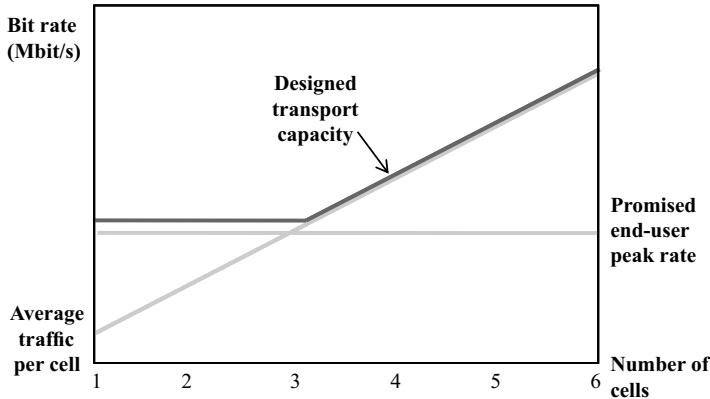


Figure 2.8 Backhaul dimensioning principle (in the access tier).

In practice several backhaul network dimensioning approaches are possible, but the cost efficient ones often combine the ability to provide the promised bit rate to any single user and the expected average traffic of all users in the cells; the principle is shown in Figure 2.8. So the backhaul link capacities are not dimensioned according to the summed up end-user capacities, but based on expected usage profiles and expected number of users per link. Some, and often significant, statistical gain can be assumed already for the first link of the base station site, as there are typically several cells served from the same site.

2.3.4 Lower Operational Costs

Network operational costs are easily as significant for the mobile operator business case as the investments needed to create the network, and this applies also for the MBH networks.

Operational costs are related to network site running costs (e.g. equipment power consumption), network maintenance and fault repairs, to network re-configurations (when connecting new links and nodes and for traffic redistribution) and also to network planning and administration. All of these depend on the MBH network size and its complexity, and how its administration, planning and operation and maintenance (O&M) are organized. Even more, the relative amount of operational costs depends on the use of own facilities or leased connections and outsourced transport services – all lease fees are typically counted as network operational costs.¹

In general the MBH network share of a mobile operator's total operational costs depends very much on the area, size and density of the mobile network. Maintenance costs depend more than other costs on the geographical area covered, as the maintenance is more expensive in sparsely populated areas, areas with long distances between the network nodes and from the maintenance centers and manned service points. Transport's share of the operational costs can be kept under control by aiming at as simplified MBH network as possible and by using network automation on a wide scale.

¹ Therefore the backhaul operational costs between mobile operators and networks cannot be compared at all without knowing how much of the backhaul is based on leased lines and outsourcing in general.

2.3.4.1 Network Simplification

Network simplification means that the target MBH network has a simple clear topology, that is easy to understand and that helps in localization of faults and performance problems. Clear topology is important both in the physical network and in the logical network (i.e. in configuration of transport paths). For example, alternative paths between the network nodes are used only where connections are important enough and reliability requires redundancy. From the operational point of view also the use of all kinds of transport overlay structures should be avoided as much as feasible, and limited mainly to transition periods in the network. Overlay structures easily increase the effort needed for network maintenance, connection configurations and network planning, and overlay structure using different technologies requires keeping competences to manage all of them. In addition, personnel capable and skilled in the maintenance of the older generation equipment are becoming with time more scarce, and costs are increased in finding and keeping such personnel in house.

2.3.4.2 Similar Technologies

Use of similar technologies and similar kind of equipment in the MBH network clearly contributes favorably to network simplification. Thus from the operational point of view the target network should contain as few different technologies as possible, and often a fully packet-based MBH network is a reasonable target, at least in the longer term. Therefore, removing some network layers can significantly reduce operational effort and thus costs (e.g. removing the ATM layer used for WCDMA traffic, and/or reducing use of TDM links when the network is upgraded).

Also, within packet-based transport technologies a mobile operator may benefit by creating a strategy for focusing on the use of only two or three types of packet technologies as widely as possible, especially in the MBH network parts built by the mobile operator itself. Even further simplification is possible by using just a few types of equipment in the MBH network, and the network O&M may benefit a lot from this kind of simplification. However, such an approach needs to be carefully judged against possible trade-offs in future equipment purchasing (it may limit vendor competition in the following network expansion phases).

2.3.4.3 New Equipment with Better Performance and Management

Newer generation equipment generally has better power-efficiency, i.e. it uses less power for a similar amount of data, and this obviously reduces the bill to be paid for the electricity; and when a site requires air conditioning for keeping internal temperatures within an operating range of equipment, there is double benefit, as lower power consumption also means less heat production.

Newer equipment also has generally better and more extensive remote management capabilities so that site visits to perform some operations locally are less likely – fewer site visits are obviously a big contributor to lower maintenance costs.

2.3.4.4 Network Automation

Network wide automation is another way of reducing operational costs. Good network O&M tools with automated routine functions increase productivity of the network management

teams, as less manual work is needed in the most common tasks and effort can be focused more on network optimization. For example, creation of new logical connections within the MBH network, or modification of existing ones, can be carried out with little manual work.

Also, automation within the MBH network itself reduces costs; for example, effective protection of connections (automatic switching to alternative paths) gives more time for maintenance teams, and reduces the likelihood that expensive night or weekend repairs are needed. It is worth noting, however, that extensive automation within the network nodes can in some cases mean trade-off with the network simplification, so that judgment is needed in priority of the goals.

2.3.5 *Developments in General Transport*

Last but not at all least, a significant change driver is developments going on in general in transport networks. The costs of packet-based transport solutions have become much lower than those of similar capacities with TDM (or ATM) technologies, as the R&D efforts of the industry have been concentrated already for some time on packet-based technologies. And more recently also the sales volumes of the packet-based transport equipment has significantly increased, meaning volume and scale benefits in their production and distribution. At the same time, technical performance of the packet-based transport solutions has been significantly improved, including the power consumption (power per bit), and equipment performance.

Thus, there is a strong push to benefit from these developments in the mobile backhaul solutions as well. Packet-based technologies have proved their superiority first in the MBH backbone networks where the traffic volumes are the highest, but today their superiority is clear also in aggregation networks. This development has continued and nowadays also affects MBH access networks where packet technologies offer higher cost-efficiency for high volume and increasingly bursty mobile traffic.

The development of general transport networks also means that leasing offerings are changing: packet-based connections will be much more widely available and will have lower price tags. There will also be available packet connections of different quality classes and more or less guaranteed bandwidths or throughputs. On the other hand, these network parts are cost optimized for fixed traffic, and technology and feature selections made for that purpose. Thus any mobile specific requirements typically mean increased price for the leased connections, or that such connections will be more difficult to obtain.

2.4 **Packet Based Backhaul Networks**

The previous sections discuss many drivers strongly pushing existing MBH networks to change, and new MBH networks to be different; they will be based on packet technologies and networking. In a fully packet-based MBH network there are no more PDH and SDH/Sonet nor ATM equipment, but it consists of packet switches and routers (layer 2 and 3 devices) connected directly to each other by physical (layer 1) connections, for example by optical fiber or microwave links. Packet networks and networking will be discussed more in Chapter 4 and technologies and equipment used in Chapter 5; below are some general observations.

2.4.1 Physical Network and Topology

Network basic structure and topology will often look quite similar to that of the earlier MBH networks (see section 2.2.2), as the basic economics of the physical network (layer 1) optimization do not change much – physical links are still shared as much as possible for several base station sites and the un-shared final links to base station sites are kept as short as possible.

The nodes used in the traffic merging and concentration points are obviously different in the packet-based MBH networks – they are, for example, Ethernet switches and IP/MPLS routers, i.e. layer 2 and layer 3 devices in packet networking terminology; some examples on possible (access) network solutions are presented in Chapter 10.

In general access tiers are more often implemented based on layer 2 solutions (e.g. Ethernet switching) while in backbone tiers more often layer 3 or IP/MPLS solutions are used; in the aggregation tier both solutions can be feasible. In each individual case selection of the packet technology to be applied is made based on economic comparisons in that environment and network topology, with expected traffic profiles and traffic growth.

2.4.2 Logical Network and Protocol Layers

Logical structure and protocols used in the MBH packet networks are discussed more in Chapter 4. However, it may be worth noting already here that even if a MBH network is based on packet technology, it is a MBH internal solution and separate from the packet technology used for the end-user traffic. This may be best seen looking on a simplified protocol picture in Figure 2.9: the end-user IP traffic is tunneled over the MBH network and remains a few protocol layers above the transport IP and transport L2/L1 protocols. Therefore, for example, the IP addresses used in the MBH network are its internal issue, and have no relation to the end-user traffic and addresses used in that.

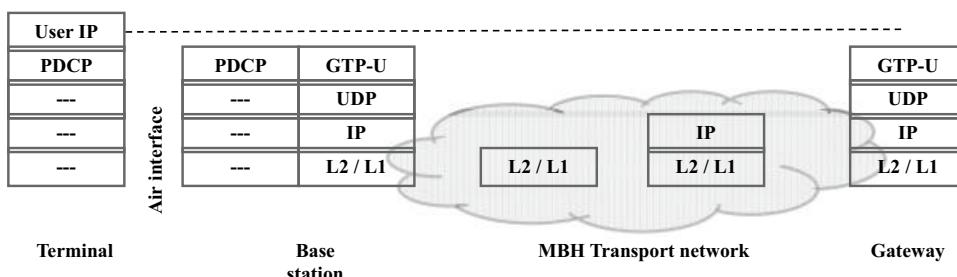


Figure 2.9 Relation of end-user IP and transport IP in a packet-based MBH network (just the principle, here LTE user plane protocols used as an example).

2.5 Making Transition to Packet Technology Networks

The new packet-based MBH networks will be different: main changes are in network capacities and implementation technologies. MBH capacity changes are directly linked to the mobile traffic volume increases, as well as to changes in traffic characteristics. Technology changes in MBH networks are related to different cost-efficiency of the legacy and packet-based transport technologies in handling high traffic volumes and bursty traffic, but they are

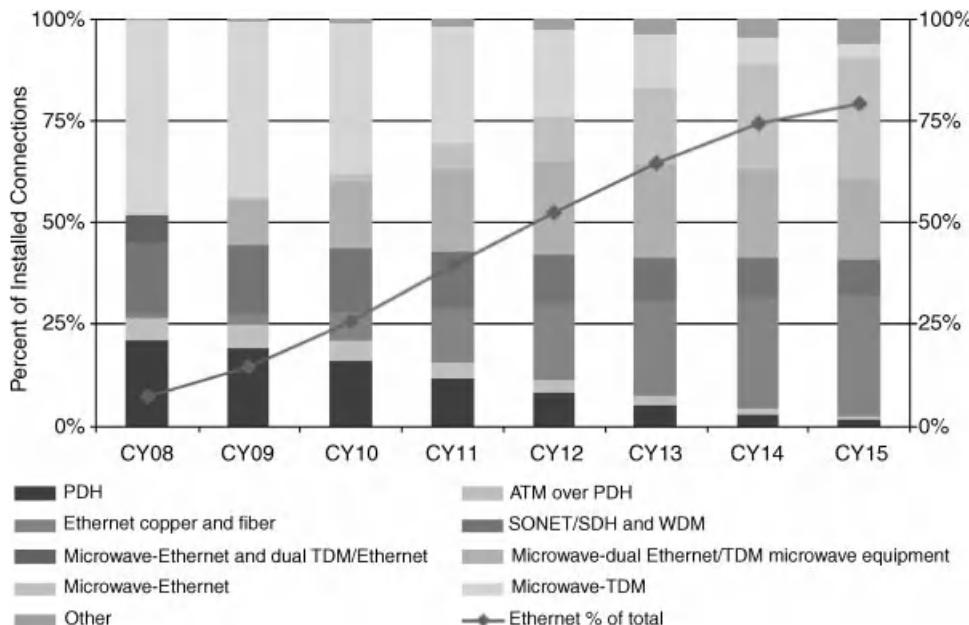


Figure 2.10 Forecast about change to packet-based technologies in MBH networks (from Infonetics Research Inc, 2011).

driven also by the change of interfaces in mobile network elements – more and more of the mobile traffic (at the transport interfaces) is natively packet based. These factors together mean a very strong push for packet-based MBH network solutions, and in the longer term towards fully packet-based MBH networks.

The change in many networks already going on, and according to the forecasts it will even accelerate in the near future. This means that more and more of base station connections will be packet based, often using Ethernet as the lower layer solution. One view about the global change in MBH technical implementations is shown in Figure 2.10.

However, it is important to note also that the different geographical regions and countries and even mobile operators are in different situations: they have different technological starting points, differences in traffic growth as well as significant differences in available funds for change, and that all results in very different speeds in the MBH technology change.

2.5.1 Transition Strategies for Packet-Based Backhaul

There are several possible paths to follow when moving from a small (or medium) capacity MBH network to the higher capacity one needed with extensive use of mobile data and other broadband services. A general backhaul transition strategy is helpful for making detailed transition plans.

An obvious starting point for the MBH transition strategy is the existing backhaul network, including own facilities and existing leasing agreements and other out-sourced services. The first things to look at when formulating the MBH transition approach is the mobile operator

business targets and the mobile network development plans based on those. The MBH network needs to be able to serve the mobile network in all its development phases and in each phase to provide enough capacity for the mobile traffic that is planned for.

In all MBH development plans one has also to note how costs of a wholly different magnitude are related to building new physical links compared to changing the technology (i.e. typically transport equipment) on existing physical links. This applies especially for links based on cables, i.e. on fiber-based transport links (as well as for still used copper links).

Building new physical links usually means not just new cables, but above all digging the ground, and related costs are often very high. Digging costs can easily be higher than the transport equipment costs already for a distance of one kilometer (in rural areas) or even for a link of 100...200 meters (dense urban area). Thus new physical links (e.g. optical cable links) are added into a MBH network only based on careful economic consideration; obviously, if there are possibilities to share the new links with fixed services or with another operator, new physical links can more easily be justified.

2.5.1.1 A Target MBH Network

The transition strategy can be formulated starting from the target network: what kind of MBH network is the longer term (or medium term) goal, what technologies it uses, what kind of topology it has and how is the connectivity arranged. This target network will take into account the above discussed issues of cost optimization, especially operational cost optimization, but it also needs to have enough flexibility, as the business and mobile network plans often change during the implementation phases.

Based on the target MBH network, mobile network development phases can be drafted, starting from the present existing MBH network.

In most cases the MBH target network, in addition to being of higher capacity than the existing one, will use different technologies, and the technology transition(s) needs to be considered in the evolution steps together with the capacity expansions.

2.5.1.2 Serving Existing and New Mobile Systems

A challenge in many (or most) areas is that the MBH network needs, in all its development phases, to support both base stations of older generations (still kept in service for years) and new base stations of the very latest generation, and often also some base stations between these generations.

For example, there will be many base station sites with 2nd and 3rd generation (e.g. GSM and WCDMA) base stations operational for several years where new high capacity base stations of the latest generation (e.g. LTE) are added. The former ones have TDM or ATM interfaces and may not be upgradable for a packet interface, while the newer ones often only have packet-based interfaces (e.g. Ethernet at lower layers). A MBH network obviously needs to provide simultaneously transport connections for all of these base stations.

2.5.1.3 Different Type of Areas

The MBH evolution and transition strategy can be, and in practice very often needs to be, different for different types of areas.

Capacity expansion needs will be highest and occur first in city centers and business districts and in those areas major MBH expansions are needed first. In these areas the likelihood of new filling base station sites (with only new technology base station, e.g. LTE) is highest, and these new sites require new high capacity and low cost MBH links to connect to the rest of the network. These links are obvious candidates for immediately using the packet technology of the planned target network. Thus the transition in these areas starts at the new sites; the high capacity needed may also justify changing all MBH links in that area for the new technology. As change in the upper tiers is often more easily done, and the traffic increase anyway requires new transport solutions, transition in this kind of area may be economically justified more or less simultaneously in the whole MBH network.

On the other hand, in areas with expected smaller traffic increase (and only few if any new base station sites), new technology base stations are coming later. Then the change of base station interface types and correspondingly the MBH technology is purely a question of cost optimization, and the MBH access network transition may be done in a later phase, especially if big investments are required in early phases in other areas. In such low-growth areas MBH technology transition typically happens first in the backbone and then in the aggregation tiers, and later in access tier.

2.5.1.4 Overlay or Replacement

For each area where mobile network is upgraded and expanded, a major decision in the MBH strategy is whether to build a (temporary) MBH overlay network for that area, or to replace the existing MBH equipment in that area wholly by new equipment. Generally speaking, when a new transport overlay is built, it is easier to optimize it for the new base stations and their requirements and traffic profile, but the whole MBH network becomes more complex and there is more equipment to manage and maintain. On the other hand, the replacement approach often leads to a simpler MBH network, but the challenges are: timing of change-over on various sites, managing the (moving) interface between the new and old MBH domains and support of older generation base stations. Economically, selection depends on the case and local circumstances; either of these approaches can lead to a lower investment cost solution, but the overlay tends to result in higher operational costs, at least for a period of time.

Overlay MBH network is usually based on the technology selected for the target network, and its capacity is dimensioned to accommodate traffic from the new base stations, at least in the first phase, in addition to all the traffic of the existing base stations. Actual moving the traffic of the older generation base stations over to the new transport network is postponed until the interfacing and other requirements are fulfilled. It may also be necessary to keep the existing MBH network in operation until all the backhaul requirements, e.g. provision of base station synchronization (see Chapter 6), are satisfactorily fulfilled within the new network. An additional benefit of keeping the legacy network up and running for longer than absolutely necessary is added resiliency during the extended parallel operation – the legacy MBH network together with the older generation base station form a back-up network which can carry at least some traffic in case of problems with the new network; this increased resilience obviously comes with a cost – higher OPEX of operating and maintaining parallel networks.

In the replacement strategy all transport in a certain area is changed, and all existing and new base stations are immediately supported by the new MBH network. Thus all the backhaul requirements, including synchronization, need to be fulfilled immediately from

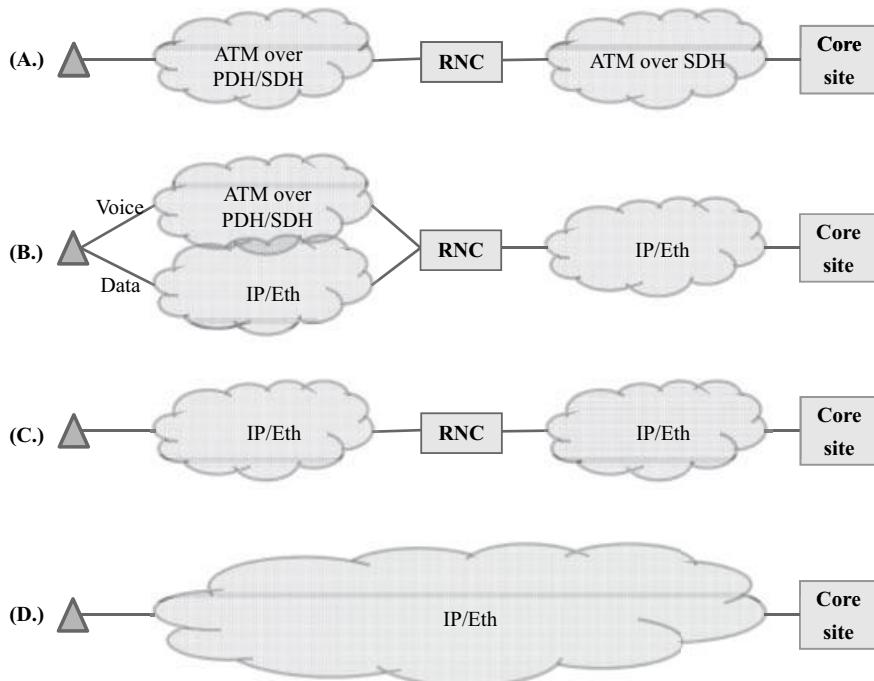


Figure 2.11 An example of the MBH development:

- A. 3G WCDMA network with ATM over TDM (PDH/SDH) transport
- B. Backbone fully packet based, access parallel TDM and packet transport
- C. Fully packet based MBH for 3G WCDMA
- D. Packet based MBH for the LTE architecture.

the change-over day, and some additional equipment may be required at the base station sites to connect the older generation base stations (e.g. layer 2 or 3 devices supporting pseudo-wire emulation for TDM and/or ATM interfaces). Replacement areas and dates need to be carefully selected so that interface between the existing technology MBH network and the new MBH network can be managed without too high costs. A possible network evolution scenario is shown in Figure 2.11.

2.5.1.5 Leased Lines and Out-Sourced Transport Services

In all transition strategies temporary or more permanent use of leased connections and out-sourced transport services need to be considered, as it may help in managing the transition phase. Conventional TDM leased lines may be used to keep the TDM part of the MBH network as a working entity during a transition period. And new leased packet-based connections (or out-sourced transport services) can be used to create a full coverage for the packet-based MBH network faster than would be possible by relying only on one's own facilities.

As always, use of other operators' services in the MBH network is more likely in the backbone and the aggregation tiers where the availability of these services is much better than

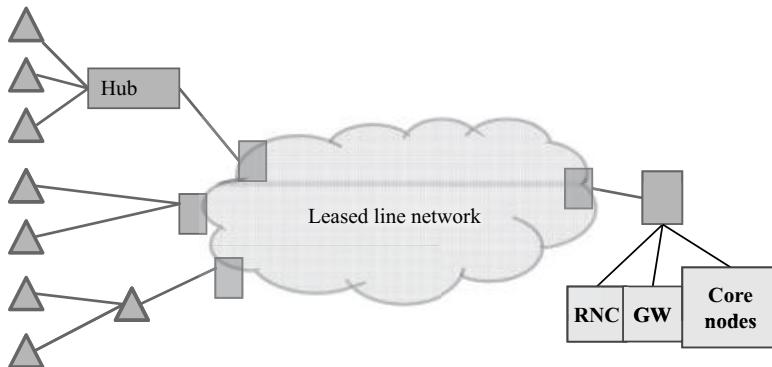


Figure 2.12 An example of a combined MBH solution:

the MBH access tier is based on own transport facilities (e.g. on microwave radio links) and the aggregation and backbone tiers are based on leased lines (*Note: RNC's and 'intermediate gateways'* are here located on the core site, as often is the case when both aggregation and backbone tiers are leased).

in the access links, and also cost level is often lower due to more competition in offerings; a typical case is shown in Figure 2.12.

2.5.2 Implementing Transition and Network Evolution

Practical implementation of the MBH network transition and evolution starts from the selected network strategy and from the existing MBH network. Good planning is an important first step, which to a large extent is responsible for smooth and successful network transitions.

2.5.2.1 Planning the Changes

Proper plans are needed for each network expansion phase, and detailed plans are needed for each area and each site, to make major MBH changes doable and to guarantee proper network operation during the transition steps and after them.

All plans have to be detailed enough and include accurate timing of various steps: addition of new equipment, creating connections over the new links (possibly first for new traffic only), then taking them into service, (later) moving all traffic away from the old equipment, and finally taking old equipment out of service and removing it from the sites.

2.5.2.2 Addition of New Equipment and Links

Addition of new equipment needs careful consideration on small sites, especially on small base station sites, and on leased sites. In those cases the available space is often limited (or additional space is expensive), and thus careful planning of equipment lay-out and installation is needed. In all sites practical things, like cabling, power supplies and temperature control, need to get enough attention. Even if a fast transition is planned, the existing and new

equipment often have to be operational in parallel for some time, and e.g. power supply and heat removal must be sufficient also during this period.

Before making the new equipment and links operational, they are taken under the network control and supervision, either within the existing network management system (if the new equipment is compatible with it) or a new management system is taken into service together with the new equipment. In both cases the network management links also need to be planned, and implemented in an early phase.

Finally, the new MBH connections, before they are taken into service, are tested and proper operation is verified according to the test plans.

2.5.2.3 Parallel Working of Old and New

When the new links are in service, the MBH networks often include different types of connections, also between the same sites, and network operations may for a period need more attention. Also, a different kind of know-how is often needed, especially when the older equipment and links are TDM (or ATM) based and the new parallel equipment and links are based on packet technologies.

When the new and existing equipment are under a different network management system, some additional care and administrative work is needed to keep the overall picture of the MBH network up-to-date and to consider possible mutual influences of changes in the parallel systems.

2.5.2.4 Simplifying the Network

The final step in the transition is typically a simplification of the MBH network. In this phase all the traffic is changed over the new links; sometimes interworking units (e.g. pseudo-wire units for TDM over packet) are needed to enable this. This step usually means significant complexity reduction in network management and in network operations, and thus helps to reduce the MBH operational expenses, as discussed earlier.

3

3GPP Mobile Systems

Esa Metsälä

The focus of this chapter is the mobile system. As 3GPP mobile systems, 2G, 3G, and LTE are covered. The intention is to review the functional division of the mobile system, by introducing mobile backhaul related characteristics. These characteristics mostly relate to the radio access network, so there is less emphasis on the core network in this text. An obvious topic is the logical interface definitions in specifications – what protocols has 3GPP standardized for the backhaul?

The goal of the mobile system as a whole is in delivering a service to the user over the air interface. For the mobile backhaul, it is thus useful to study the protocols and elements that are involved in this delivery at the radio network layer.

Section 3.1 consists of an introduction, which may help in reading the subsequent sections. 2G system (with evolution) is discussed in 3.2, 3G system (with HSPA and further evolution) in 3.3, and LTE in 3.4. Focus is on HSPA and LTE.

3.1 3GPP

3.1.1 Radio Technologies and Backhaul

Since the introduction of HSDPA in 3GPP rel-5, and into the networks since the mid 2000s, mobile backhaul became a topic of increased importance. Introduction of LTE is a continuation on the same theme. Enhancement of the air interface and mobile network capacities have proven not to be always easily met on the backhaul side. Expanding the transport network capacity can be both time-consuming and costly.

Existing TDM networks may (and will) continue to be used, but they lack the capacity and flexibility of packet networks. An expansion into mobile broadband requires packet technologies, which is a topic for Chapter 4 of this book.

While the present networks continue to serve voice as well as data users, new radio technologies are introduced. Multiple radio networks are operated in parallel. The systems may use their own

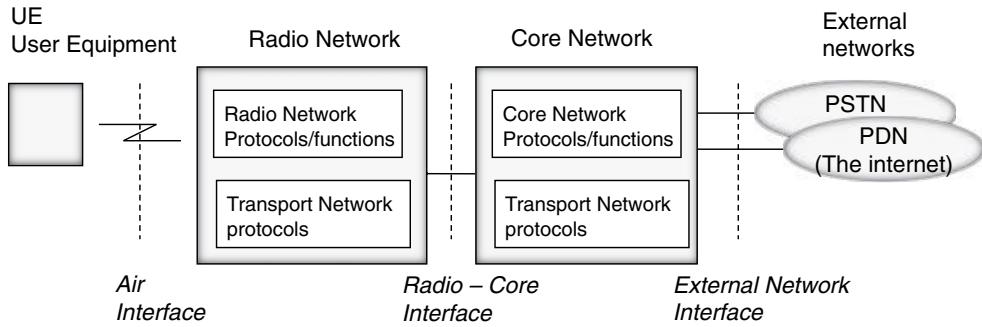


Figure 3.1 Mobile system. Based on [1].

separate backhaul, but for cost efficiency, a single packet network is a consideration. This requires that the common packet network can meet the needs of all of the systems.

3GPP radio networks, 2G, 3G, and LTE, have their own needs for the mobile backhaul for delay, loss, QoS, synchronization, security, resilience and so on. Requirements originate from the functional split between elements, and from the type of protocols operating at the radio network layer; and from the end user service.

Even though many aspects in the system architecture of the mobile networks are different, there are also commonalities. The development is in many ways an evolution from the previous system. Similar concepts and common protocols appear, which also leads to some repetition in the text.

A common architectural property of all mobile network generations is that the functionality is split between the radio and core functions as shown in Figure 3.1. Radio network is managing the air interface resources and intends to optimize and maintain a good quality signal to the terminal, and also make efficient use of the resources it has. Core network is free from the radio related functions and is managing the subscriber information, authentication, charging, mobility, and interfacing other networks such as the public Internet (PDN, packet data network) and the public switched telephone network (PSTN).

Transport layer provides a service for the 3GPP radio and core network protocols. Layered protocol architecture isolates layers from each other, and allows communication only via specified service access points. This makes for a clean software interface. Even though the interface between radio and transport is well specified with defined interactions, in practice there are functional interdependencies.

Retransmissions in case of erroneous or lost packets may happen on multiple layers. Scheduling takes place for the air interface and also for the transport interfaces of the BTSs and backhaul elements. Part of the traffic may be encrypted by the mobile system, and then travel further with other traffic into another network element, encapsulated in an IPsec tunnel.

Several topics, QoS, synchronization, security and resilience, have clear dependencies with the radio network. They are studied in further detail in Part II of the book.

3.1.2 Organization

3GPP (initially 3G Global Partnership Project) was formed in 1998, and consists of standardization partner organizations, market representation partners, and individual members. The 3GPP mobile system standardization is carried out within the 3GPP.

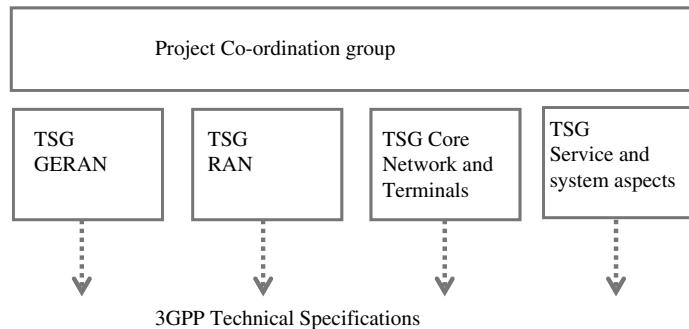


Figure 3.2 3GPP organization [2].

The standardization partner organizations are:

- ARIB (The Association of Radio Industries and Businesses, Japan)
- ATIS (The Alliance for Telecommunications Industry Solutions, USA)
- CCSA (China Communications Standards Association)
- ETSI (The European Telecommunications Standards Institute)
- TTA (Telecommunications Technology Association, Korea)
- TTC (Telecommunication Technology Committee, Japan)

3GPP organization consists of a Project Co-ordination Group, and Technical Specification Groups (TSGs). TSGs exist for the radio access network (TSG GERAN and TSG RAN), core network and terminals (TSG Core Network and Terminals), and for service and system (TSG Service and system aspects). 3GPP organization is depicted in Figure 3.2.

Further information can be found from www.3gpp.org.

3.1.3 *Specifications*

Standards are developed for a standardization release, a 3GPP release. Each specification, such as TS 25.414 ‘UTRAN Iu Interface Data Transport & Transport Signalling’ as an example, has a version in each of the 3GPP releases. The versions differ if new functionality has been added. It may also be that the standard is updated (version number incremented) for the new release without a change in the content.

Release content can be tracked by release specific descriptions. Work items which have been accepted, and whose specification work has been completed in time, make their way to that particular release. After freezing of specifications, change control is applied with a change request procedure. Accepted change requests update the specification content with an incremented version number.

Specifications are divided into stages: Stage 1 specifications give a description from a service user’s viewpoint. In stage 2, the topic is divided into functional elements. Also the access reference points are defined. Stage 3 defines an implementation of the protocol at a physical interface.

3GPP is not creating standards for transport, instead standards from other fora are referred to. In the case of IP transport, the main reference is IETF and its RFCs. With earlier non-IP logical interfaces, standards from ITU-T were often referred to.

3.1.4 Releases

3GPP releases are shown in Figure 3.3 (Stage-3 freeze dates), following approximately a yearly release cycle, with some exceptions. GSM specifications were originally developed by ETSI (European Telecommunications Standardization Institute). This work is not shown here. Since that, GSM further development has been merged into 3GPP. Some major functionality included in the release is shown in Figure 3.3.

Complete content of a 3GPP Release can be found from the release descriptions. IP transport was introduced to 3G (UTRAN) in Rel-5. HSDPA and HSUPA are defined in 3GPP Rel-5 and Rel-6, respectively. LTE standards (first release) were completed in Rel-8. At the time of the writing Rel-11 specification work is ongoing.

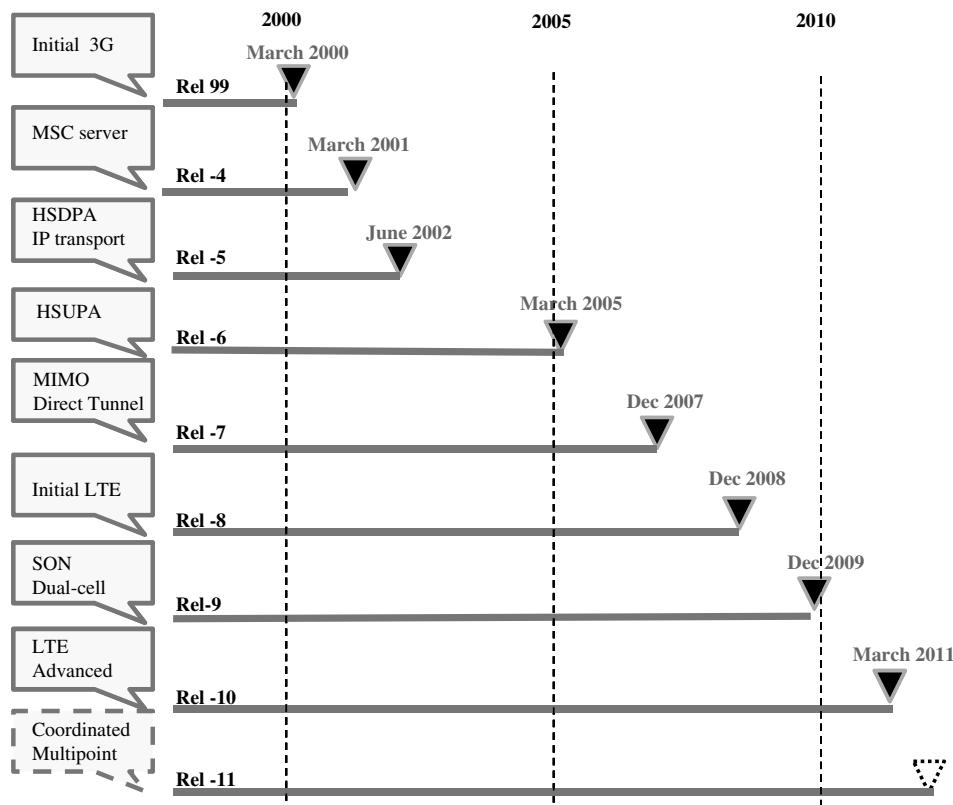


Figure 3.3 3GPP releases [3], [4], [5], [6], [7], [8], [9], [10], [11].

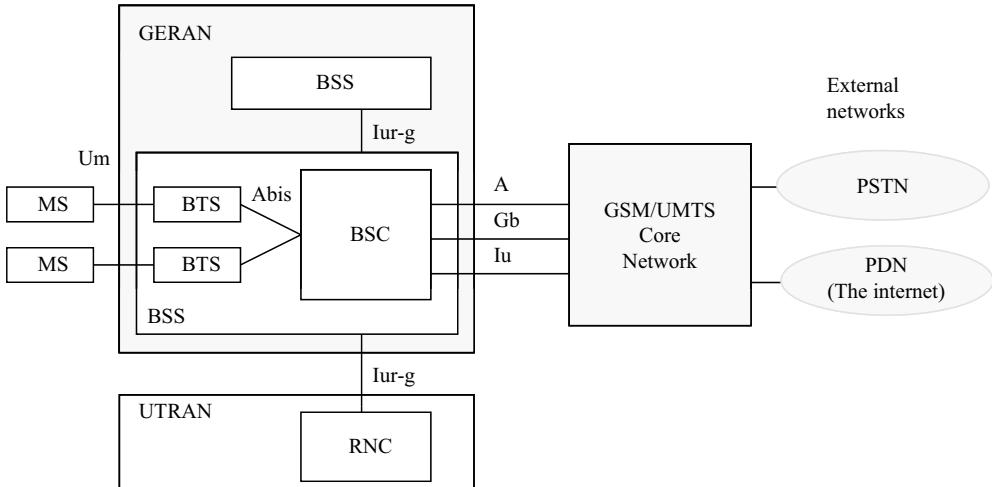


Figure 3.4 GERAN structure [12].

3.2 2G

Since the inception, 2G system has evolved with new functionalities during the years and standardization releases. So for discussion on further detail of the system capabilities, one needs to consider the 3GPP release to which the system complies with. Initially in 2G, the focus was on providing mobile voice service. One technological difference compared to the previous ‘1G’ systems is that in 2G voice and all services are digital, not analogue. For further reading on 2G see [12] and [13].

Enhancements in the 2G network include AMR (Adaptive Multi-Rate) speech, high speed circuit switched data (HSCSD), GPRS (General packet radio service), and EDGE (Enhanced data rates for global evolution). Term GERAN (GSM/EDGE Radio access network) is used to refer to the 2G radio access network with these enhanced capabilities. GERAN also includes an Iu-interface option, which is inherited from the 3G system. Mobile stations may then operate either in A/Gb mode or in Iu mode, depending on how the 2G radio access network interfaces the core network.

The system architecture of the GERAN network is shown in Figure 3.4. The base station subsystem (BSS) internal interface between the BTS and the BSC is Abis.

3.2.1 Circuit Switched Traffic

Figure 3.5 shows the GERAN user plane with both A-mode and Iu mode interfaces. The initial 2G (A-mode) protocols are indicated with a grey colour. The CS user plane is optimized for the transport of voice.

For the A mode, the protocol model for the CS domain user plane is simple: there are only physical and logical channels that are mapped into timeslots. Encryption is supported by the BTS, controlled by the BSC and the core network.

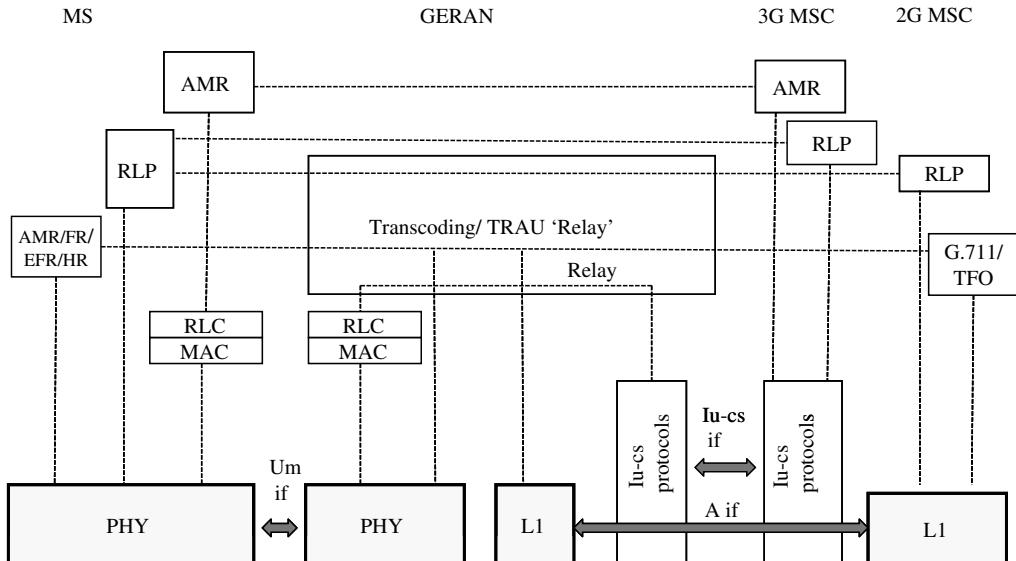


Figure 3.5 2G circuit switched user plane [12].

The initial GSM voice codec is RPE-LTP (Regular pulse excited long term prediction), which results in 13 kbit/s bit rate (full-rate). Before connecting the mobile voice traffic to the traditional PSTN, transcoding is needed via a transcoding and rate adaptation function (TRAU). Voice is carried between the MS (Mobile Station) and the TRAU or MSC. GERAN is carrying the voice frames between the air interface Um and the A interface. In the Iu mode, there are further options.

In practice TRAU is often located close to the core network (MSC), since by this means transport timeslots can be saved. One GSM-coded voice channel occupies a single 16 kbit/s subtimeslot compared to the 64 kbit/s voice rate (A-law or u-law encoded) in the traditional PSTN.

With half-rate, 5.6 kbit/s is required for voice. Enhanced full rate (EFR) is 12.2 kbit/s, increasing speech quality. AMR (Adaptive Multi-Rate) also improves quality for both half-rate and full-rate channels, by adapting its bit-rate to channel conditions. Part of the capacity of the AMR is used for channel coding.

The A interface is in the initial 2G standard a TDM-based interface. In 3GPP Rel-7, A over IP was introduced for the control plane [24]. The user plane protocol stack for A over IP, specified in 3GPP Rel-8, consists of RTP/UDP/IP [25]. For a BSS/MGW pair, two consecutive UDP ports are used, one reserved for RTP (Real-time Transport Protocol), and the other one for RTCP (RTP Control Protocol). The use of RTCP is optional. The same UDP port number is used in both receive and transmit directions. For the layers below IP, basically any L2 protocol is allowed, however 3GPP mandates at least Ethernet.

RTP multiplexing can optionally be supported, if negotiated successfully with RTCP. With RTP multiplexing, several RTP user connections can be carried within the UDP/IP packet. For RTP multiplexing an additional multiplexing header is included as defined in [25], [26]. This saves bandwidth, as many RTP packets may share a single UDP/IP packet. RTP header can also be compressed.

The control plane of 2G circuit switched domain is shown in Figure 3.6.

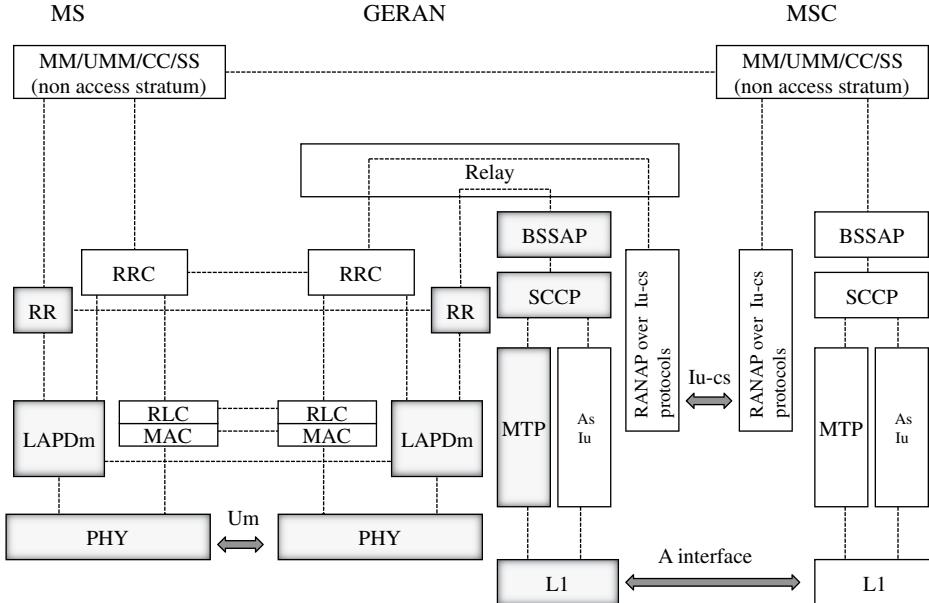


Figure 3.6 2G circuit switched control plane [12].

Figure 3.6 shows the GERAN circuit switched control plane with both A-mode and Iu mode interfaces. The initial 2G (A-mode) protocols are indicated with a grey colour.

Non-access stratum signalling (MM/UMM/CC/SS) is transmitted transparently over the GERAN. Over the A interface they are carried using Direct Transfer application part (DTAP).

BSC is responsible for the radio resource management, RR (Radio resource) in the figure. BSC communicates with the mobile station, but also with the BTS with these RR messages. Part of the messaging from MS to the BSC is transparent to the BTS. BTS then maps these messages to the RR' messages that are carried over the air interface.

In the A mode in the CS domain control plane, LAPD is the control plane protocol over the Abis. LAPDm marks the LAPD protocol over the air interface, while LAPD is used between the BTS and the BSC over the Abis. RR messages are carried over the LAPD protocol.

LAPD is based on the ISDN specifications. GSM specifications refer to EN 300 125 in addition to the CCITT Recommendation Q.921. For a BTS, there may be multiple LAPD links over the Abis, which carry signalling, operation and maintenance, and layer-2 management procedures. LAPD terminal endpoint identifiers (TEIs) are used for addressing.

For the A interface, IP based control plane is supported, with the SIGTRAN protocol stack, BSSAP/SCCP/M3UA/SCTP/IP. BSSAP consists of BSSMAP (BSS Management Application Part) and of DTAP (Direct Transfer Part). BSSMAP is the messaging between the BSS and the core network. SCCP is a Signalling Connection Control Part of the CCITT signalling system No.7, and M3UA (Message Transfer Part 3 User Adaptation Layer) is an adaptation to the SCTP.

3.2.2 Packet Switched Traffic

For the 2G packet switched domain in the user plane, Gb interface supports packet switched data between a Packet Control unit (PCU) and the SGSN. PCU can be considered as an addition to the BSS and implemented e.g. within a BSC.

The Gb interface uses network service virtual links (NS-VL) and network service virtual connections (NS-VC) as abstractions. The underlying transmission is defined as a subnetwork, which initially is standardized to be Frame Relay.

In 3GPP Rel-4, the interface specification introduces IP as an alternative to Frame Relay for the implementation of the subnetwork. The NS-VL is mapped into an IP endpoint, and, in the case of IP transport, the NS-VCs consists of connectivity between source and destination IP addresses and UDP ports [31]. The abstraction intends to hide the realization of the network service from the upper layer (BSS GPRS Protocol).

Figure 3.7 shows the user plane protocol stack carried over the air interface and the GERAN to the SGSN. In the SGSN the user plane traffic is further relayed to the GGSN (Gateway GPRS Support Node) and to the Gi interface.

Figure 3.7 shows GERAN with both A-mode and Iu mode interfaces. The initial 2G/GPRS (Gb-mode) protocols are indicated with a grey colour.

Between the MS and the SGSN, SNDCP (Subnetwork Dependent Convergence Protocol) with Logical Link Control (LLC) protocol is used. LLC is based on LAPD and HDLC (High Level Data Link Control) concepts, supporting acknowledged and unacknowledged mode transfers. Service access point identifiers (SAPIs) identify the LLC layer service access points for the higher layer protocols. LLC includes ciphering, so the user data is encrypted between the MS and the SGSN.

RLC (Radio Link Control) layer provides services like segmentation and reassembly of the LLC layer PDUs and supports both an acknowledged and an unacknowledged mode—these

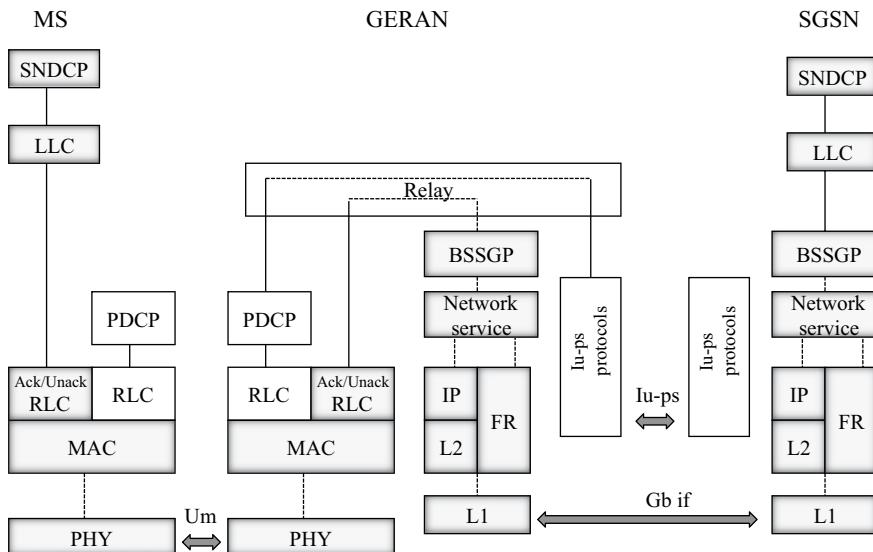


Figure 3.7 2G packet switched user plane [12].

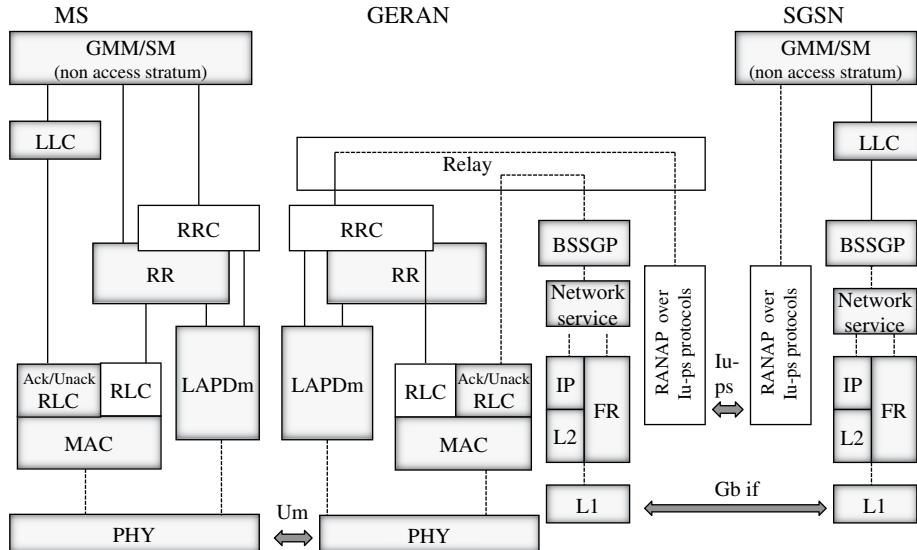


Figure 3.8 2G packet switched domain control plane [12].

are specified in GSM 04.60. For the acknowledged mode delivery, selective retransmission of unsuccessfully delivered RLC data blocks is supported.

MAC (Medium Access Control) layer supports the transmission of logical channels over the physical channels. Physical channels may be dedicated or shared. MAC layer configures the mapping of logical channels into the physical channels.

Between the GERAN and the core network, signalling protocol used is BSSGP, based on the SS7 signalling stack. BSSGP provides buffering and mapping of parameters between the RLC/MAC layers and the BSSGP layer as a relay function. A network service (IP or Frame Relay) transports the BSSGP layer PDUs. Another option is to interface the core network in the Iu mode.

For the packet switched domain control plane, the protocol stack is shown in Figure 3.8.

Figure 3.8 shows the GERAN with both Gb-mode and Iu mode interfaces. The initial 2G/GPRS (Gb-mode) protocols are indicated with a grey colour.

Between the GERAN and the core network, signalling protocol used is BSSGP, based on the SS7 signalling stack. BSSGP uses network service, which can be realized by Frame relay or by IP based stack. Another option is to interface to the core network in the Iu mode.

3.2.3 Abis

Abis interface consists of traffic channels (TCH) for transceivers (TRX) and signalling channels (LAPD channels). Traffic channels are used for voice and data (e.g. GPRS/EDGE). Logical channels include traffic channels (TCHs), intended for encoded speech of full rate (TCH/F) or half rate (TCH/H), or CS data. TCHs are bi-directional. Packet Data Traffic channels (PDTCHs) carry user data. Additionally there are channels for control traffic.

The capacity of one traffic channel depends on the voice codec used (Full-rate, half-rate) and on the modulation and coding scheme (MCS) used for the data traffic. Traffic channel capacities are 8, 16 or 64 kbit/s. Multiple traffic channels may be combined to realize higher data rates.

Signalling channels on the Abis interface include both BSC-BTS signalling, and BSC-MS signalling. Signalling channel capacities are 16, 32 or 64 kbit/s.

Transport on the Abis is TDM between the BTS and the BSC, referring to ITU-T Blue Book definitions of G.703 for physical and electrical characteristics. On the Abis, traffic is mapped into timeslots (64 kbit/s) or sub-timeslots (8, 16, 32 kbit/s). An E1 interface consists of 32 timeslots, a T1/JT1 interface of 24 timeslots, each of 64 kbit/s capacity. One 64 kbit/s timeslot is further divided into four subtimeslots of 16 kbit/s. In the TDM based Abis network, traffic can be optimized by multiplexing traffic, even down to a 8 kbit/s level using TDM multiplexers. With GPRS and EDGE, the Abis definition remains the same. There is no evolution of Abis in the 3GPP for an IP based transport.

Native TDM circuits may be used, or the whole E1/T1 frame may be carried over a packet network using circuit emulation. Vendor specific IP Abis solutions exist, where Abis traffic content is carried over the IP protocol (without emulating the Abis). IP based Abis is not specified in 3GPP. Existing TDM based Abis may be emulated over a packet network e.g. by using a cell site gateway.

Especially with GPRS/EDGE data traffic, TDM based Abis is inefficient, since the timeslots are reserved constantly for each user. To improve the utilization and Abis capacity, timeslots may be allocated dynamically from a pool so that data users share these pooled timeslot resources. In this type of implementation, timeslots are allocated to a user only during a period of time, after which another user gets served. This improves the efficiency of the Abis, as otherwise resources would be reserved also during the idle times of a data session. Solutions for the pooling and flexible resource allocation differ between implementations, as this is not covered in 3GPP.

Inefficiency exists also with voice traffic over the TDM based Abis. Variable rate voice codecs do not transmit at a constant rate. However at Abis a constant rate (a subtimeslot) is consumed. With IP based implementation of Abis this can be addressed.

From a QoS viewpoint the system is simple with a native TDM network. All the traffic channels and traffic types are treated identically. All timeslots pass through and there is no congestion after the traffic has been admitted to the system. Blocking occurs if there are no free timeslots available in the Abis, which is not common, since all traffic channels of the air interface (radio timeslots) are directly mapped to Abis timeslots.

With dynamic or flexible Abis optimization the mapping changes dynamically and the Abis usage becomes more flexible. Still there is no need to differentiate traffic either as all timeslots are transmitted without loss or excessive delay.

3.3 3G

With 3G, there are equally many evolution steps. For mobile backhaul, the most important ones are:

- MSC server based core network architecture in 3GPP Rel-4
- IP transport in 3GPP Rel-5

- HSDPA in 3GPP Rel-5
- HSUPA in 3GPP Rel-6
- One tunnel in 3GPP Rel-7
- 64QAM (Quadrature Amplitude Modulation) (HSDPA), 16QAM (HSUPA) in 3GPP Rel-7
- 64QAM and MIMO (Multiple Input Multiple Output) combination (HSDPA) in 3GPP Rel-8
- Dual-cell and MIMO combination (HSDPA) in 3GPP Rel-9
- Dual-cell (HSUPA) in 3GPP Rel-9
- 4-carrier HSDPA in 3GPP Rel-10
- (Potentially) 8-carrier HSDPA in 3GPP Rel-11 (Rel-11 ongoing)

The above translates to (theoretical maximum) downlink peak rates for a single user so that with Rel-7, HSDPA 64 QAM reaches 21 Mbit/s. Rel-8, HSDPA 64 QAM with 2×2 MIMO achieves 42 Mbit/s. Rel-9, Dual-cell (2×5 MHz), MIMO and 64 QAM produces 84 Mbit/s. With a higher amount of carriers aggregated, the peak rate grows (four carriers 84 Mbit/s, and eight carriers potentially 168 Mbit/s). In practice, availability of terminals may limit the applicability. However, this illustrates that high peak rates are achievable with the enhanced HSPA + networks.

Additionally, a number of improvements to existing functionalities have been included.

A simplified 3G system architecture with a UTRAN focus is shown in Figure 3.9.

Each Node B connects to an RNC (Radio Network Controller) which interfaces the core network (CN). Iu-cs interface carries the circuit-switched traffic, and control plane (RANAP). Iu-ps carries the packet-switched traffic, and control plane (RANAP).

Node-B – RNC configuration is a static one: each NodeB is statically assigned to a specific RNC. This can be changed by a management plane configuration. No dynamic load sharing or multihoming of NodeBs to multiple RNCs exists in the standard.

RNC acts in multiple roles, for UEs and for nodeBs. For a UE, one RNC (serving RNC, SRNC) terminates the Iu link. A drift RNC (DRNC) is any other RNC than the serving RNC.

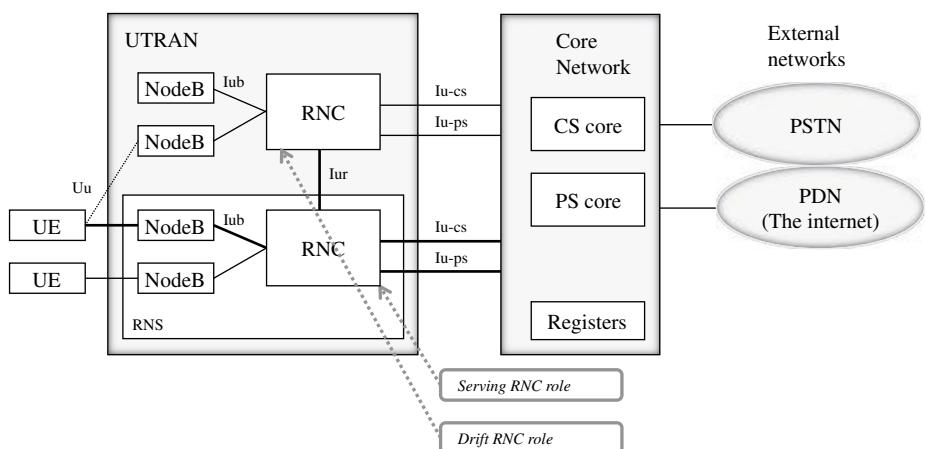


Figure 3.9 3G system and UTRAN [37].

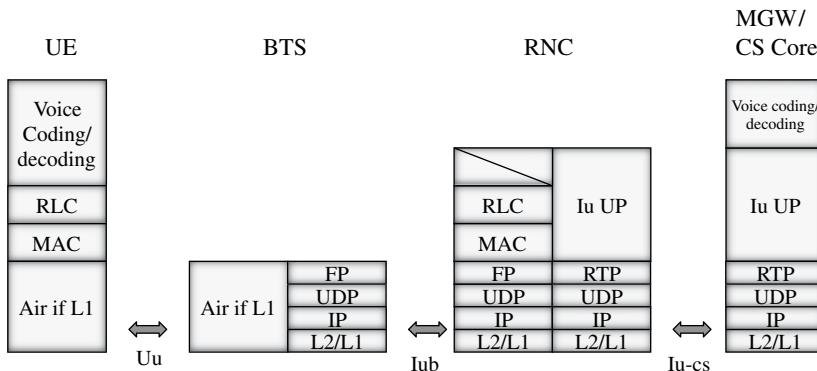


Figure 3.10 3G circuit switched user plane (Voice example) [36], [41].

For each nodeB, there is a single controlling RNC (CRNC). This controlling RNC is responsible for radio resource management of the cells of NodeBs, that are configured to that particular CRNC.

The serving RNC terminates the RANAP link to the core network, and also the RRC signalling from the UE. (Radio) L2 processing takes place in the SRNC, and the L2 protocols are carried via the DRNC to the SRNC, where they are terminated.

The NodeB is responsible for the air interface L1 processing. This includes channel coding and interleaving, spreading, scrambling, etc. With HSPA, high speed mac scheduling (i.e. part of the radio L2 protocols) move to the NodeB.

For the next two sections 3.3.1–3.3.2, only the IP based UTRAN interface stacks are shown. Initially all these interfaces in the UTRAN are based on ATM.

For further reading on 3G, see [37], [38], [39] and [40].

3.3.1 Circuit Switched Traffic

User plane protocol stack for the circuit switched traffic is shown in Figure 3.10.

Circuit-switched traffic, e.g. voice, is over the Iub carried by FP (Frame protocol)/UDP/IP protocol stack, assuming the IP transport option. RLC (Radio Link control)/MAC (Media Access control) layer is implemented into the UE and to the RNC. At the Iu-cs interface, Iu user plane towards the circuit-switched (CS) core, RTP/UDP/IP stack is used. RTCP protocol is optional.

Iu user plane protocol (Iu UP) has two operating modes: transparent mode and a support mode for a predefined SDU size. Iu UP intends to be agnostic to the transport network layer, as well as to the core network (packet or circuit switched).

Control plane protocol stack for the CS traffic is shown in Figure 3.11.

RRC messages (radio network L3 control) over the Iub are carried over FP/UDP/IP. RLC/MAC layer entities are in the UE and in the RNC. RANAP conveys non-access stratum signalling to the CS core (MSC).

In the IP protocol option of Iu, RANAP is mapped over the SIGTRAN protocol stack; SCCP/M3UA/SCTP/IP.

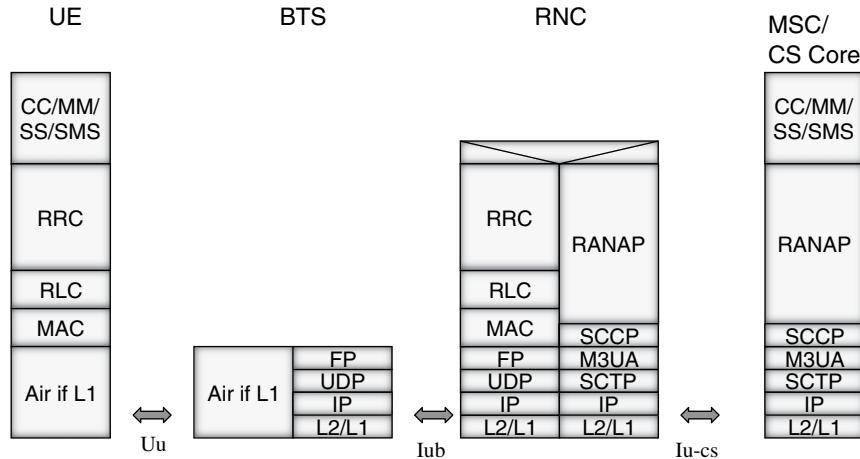


Figure 3.11 3G CS domain control plane [36], [41].

3.3.2 Packet Switched Traffic

Protocols for the user plane packet switched domain are shown in Figure 3.12, using the initial 3GPP Rel-99 DCH (Dedicated Channel) channel. For HSPA, see section 3.3.5. Instead of using DCH for packet switched data, mobile broadband is today implemented more efficiently and with higher data rates with HSPA.

In the example TCP/UDP stack is used for the application in the UE. Peer entity could be for example a server in the Internet. IP transport option is used at the UTRAN interfaces (Iub and Iu).

Within the radio access network, IP packets are in the UE encapsulated within PDCP. PDCP layer is terminated in the RNC, and the user packets are then carried over GPRS tunneling

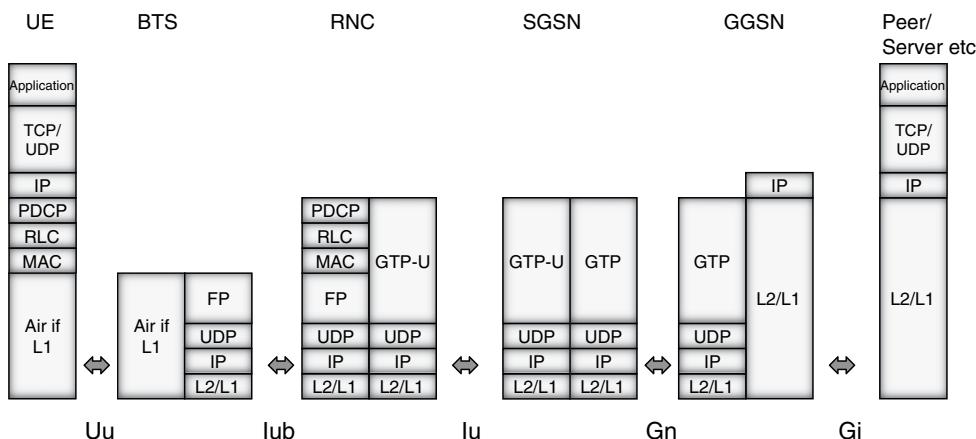


Figure 3.12 3G PS domain protocols [27] (Rel-99 DCH example without one tunnel architecture). For HSPA, see section 3.3.5.

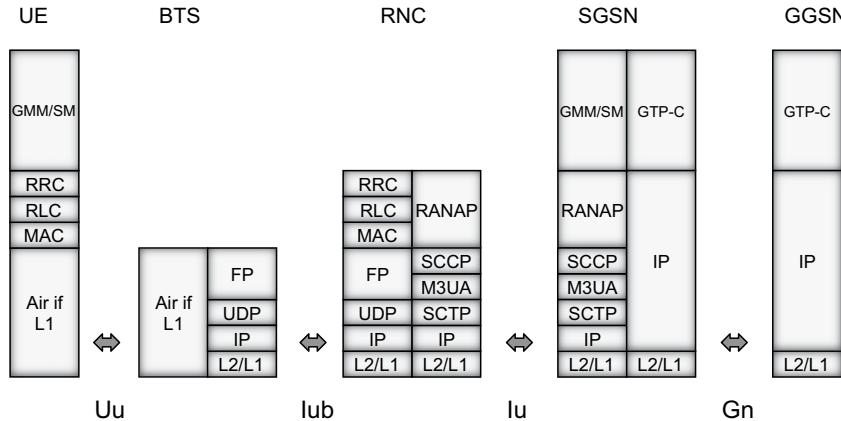


Figure 3.13 3G PS domain control plane [27].

protocol (GTP-U) to the SGSN. Between SGSN and GGSN GTP tunneling is again used. GGSN assigns an IP address to the terminal, and also provides an access point (APN) for the terminal.

Two significant improvements on the PS domain are one tunnel (also often called Direct tunnel) and HSPA. With the one tunnel concept, SGSN can be omitted in the user plane. In this case, the GTP-U tunnel extends between the RNC and the GGSN. In the control plane, SGSN is still needed. This removes one mobile network element (SGSN) completely from the user plane traffic flow.

Another change is in the radio network with the introduction of HSDPA and HSUPA. MAC high speed scheduling entities are added to the NodeB. NodeB is also responsible for fast scheduling of the air interface.

Control plane protocol stack is shown in Figure 3.13.

RRC messages (radio network L3 control) are transported over the Iub with FP/UDP/IP protocol stack. RLC/MAC layer entities are in the UE and in the RNC. RANAP conveys Non-access stratum signalling to the SGSN in PS core. GTP-C is the control plane protocol between the SGSN and the GGSN.

In the IP protocol option of Iu, RANAP is mapped over SIGTRAN protocol stack; SCCP/M3UA/SCTP/IP. This is identical to the Iu-cs.

3.3.3 3G Air Interface Channels

Air interface is divided into three protocol layers: the physical layer (L1), the data link layer (L2), and the network layer (L3). L2 protocols include Medium Access Control (MAC), Radio Link Control (RLC), Packet Data Convergence Protocol (PDCP) and Broadcast/Multicast Control (BMC).

Transport channels depend on the L1 technology, i.e. FDD (Frequency division duplex) or TDD (Time division duplex). The FDD channels are described in the subsequent text.

There are multiple options in how the user traffic is mapped into the channels. DCHs are used e.g. for circuit-switched voice, but DCHs can also carry data. HS-DSCH and E-DCHs support higher data rates than DCHs. HS-DSCH is a shared resource, but can support also

guaranteed bit rates. In the control plane, common channels (such as FACH and RACH) carry control information.

Dedicated channels include Dedicated channel (DCH), uplink or downlink, dedicated for one UE, and an enhanced dedicated channel (E-DCH). E-DCH is often referred to as high speed uplink channel (HSUPA). HSUPA scheduling is in NodeB and supports H-ARQ.

Common transport channels include Random Access Channel (RACH), which is an uplink channel for control traffic. (It may also carry small amounts of user data). Forward Access Channel (FACH), is a common downlink channel for a relatively small amount of data, and for broadcast and multicast data. Broadcast Channel (BCH), is a downlink channel for broadcast of system information. Paging Channel (PCH), is used e.g. for paging and notification, and is a downlink channel. High Speed Downlink Shared Channel (HS-DSCH) is a downlink shared channel, often referred to as HSDPA. HSDPA scheduling is in NodeB and supports H-ARQ.

Service access points (SAPs) provide services between layers. Physical layer provides transport channels. MAC layer provides logical channels (e.g. DCH). RLC layer provides unacknowledged mode, acknowledged mode, and transparent mode services. The services of the RLC layer are called radio bearers. In the control plane, the bearers are signalling radio bearers.

Layer 3 and RLC can be divided into control and user planes. In the control plane, the L3 lowest sublayer, the RRC protocol, (Radio Resource control), terminates in the RNC. The next higher layer, Duplication avoidance, terminates in the core network, but provides services to the higher layers in the access stratum. The higher layer signalling, e.g. mobility management, terminates in the core network, and is part of the non-access stratum. See, for example, Figure 3.13.

In the 3G radio access network system, macrodiversity combining is supported in the RNC (Figure 3.14). In a soft handover, UE maintains parallel user plane connections to the RNC. RNC selects the best stream available, thereby increasing the quality of the connection to the UE. Soft handover adds overhead on the mobile backhaul, as multiple parallel links are required during the soft handover.

In Figure 3.14, Radio Bearers (RB1 and RB2) terminate at the RNC, consisting of transmission over the air interface and over the Iub.

3.3.4 FP, MAC and RLC Protocols

Frame protocol is used for the transport of the radio network protocols over the Iub, and over the Iur. It also carries outer loop power control information to the serving RNC in FP control frames, and supports timing alignment.

The data is passed from the upper layer (MAC) to the FP layer as transport blocks. A transport block set consists of transport blocks for the same transport channel. Transport Time Interval (TTI) is the time interval of transport block sets to be transmitted. TTIs are in multiples of 10 ms. With HSPA, a shorter TTI (2 ms) is supported.

Timing alignment allows adjusting the timing of downlink FP frames relative to the air interface timing. With the procedure, NodeB may instruct RNC to advance or delay sending of the FP frames. A time of arrival window in NodeB is defined, and FP frames are expected to be received within the window. Frames arriving later will miss their air interface slot and have to be discarded. See Figure 3.15. The parameter values are delivered to the NodeB via NBAP.

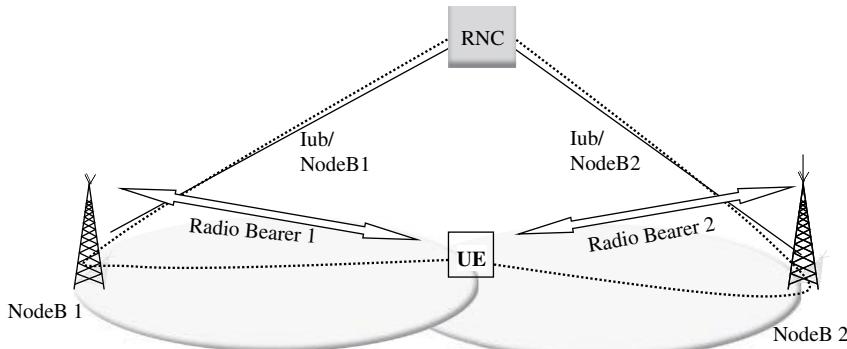


Figure 3.14 Soft handover branches.

Time of Arrival, Window Start (ToAWS), and Time of Arrival, Window Endpoint (ToAWE) define the receiving window.

In the uplink, macrodiversity combining is supported in the RNC. System Frame Number (SFN) and Connection Frame Number (CFN) information is used to synchronize the different branches before combining.

FP packet consists of a header and payload. The FP packet for the downlink DCH, is shown in Figure 3.16. Additionally, padding may be needed to fill the transport blocks at octet boundaries.

The fields are header crc, frame type (FT) for data or control frame, connection frame number (CFN), transport format indicators (TFI), transport blocks, Spare extension and an optional Payload checksum.

One MAC PDU maps into one transport block. Connection frame number informs of the first radio frame, where the transport block should be sent (downlink) or was received (uplink).

Both outer loop power control and scheduling of the transport blocks (scheduling by the MAC layer in the RNC), mean that the transport on the Iub is delay critical. FP control frame functionality is also used to indicate failures to higher layers (RRC). A failure may, via RRC actions, for example, cause termination of the Radio Bearer. A link break on the mobile backhaul is one example of such a failure case. Depending on timers and other implementation specific topics, this happens in 1 sec to few seconds, to even over 10 seconds.

MAC layer is the scheduling entity in the RNC. A transport format defines the format that the physical layer offers for MAC layer, for the delivery of a transport block set during a TTI. A transport format set consists of transport formats associated to a transport channel. A transport

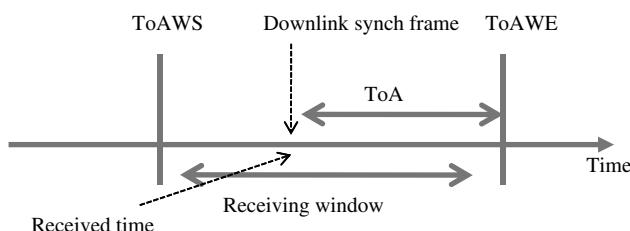


Figure 3.15 Time of arrival.

Header CRC	FT
CFN	
TFI of the first DCH	
...	
...	
TFI for the last DCH	
First TB of first DCH	
...	
...	
Last TB of first DCH	
...	
...	
First TB of last DCH	
...	
...	
Last TB of last DCH	
Spare extension	
Payload checksum (Optional)	
Payload checksum (Optional) (continued)	

Figure 3.16 FP packet. [55] (Header fields with grey color)

format indicator tells the physical layer, which transport format to use. Scheduled bit rate is varied by modifying the transport formats, or the attributes of it. The attributes include transport block size, transport block set size, TTI, error protection, and CRC.

Additional MAC functionality is ciphering. With transparent RLC mode, MAC layer also supports ciphering.

3.3.5 HSDPA (HS-DSCH) and HSUPA (E-DCH)

HSDPA and HSUPA change the functional split within the UTRAN. Radio interface scheduling is moved to the NodeB, and, due to link adaptation and faster scheduling, there is no need for fast power control from the RNC. Higher order modulation is also supported, increasing the bit rate.

A new mac-hs (MAC high speed, see Figure 3.17) scheduling entity is included in the NodeB. Scheduling takes place every 2 ms, as opposed to 10 ms in Rel-99 DCH (or, multiples

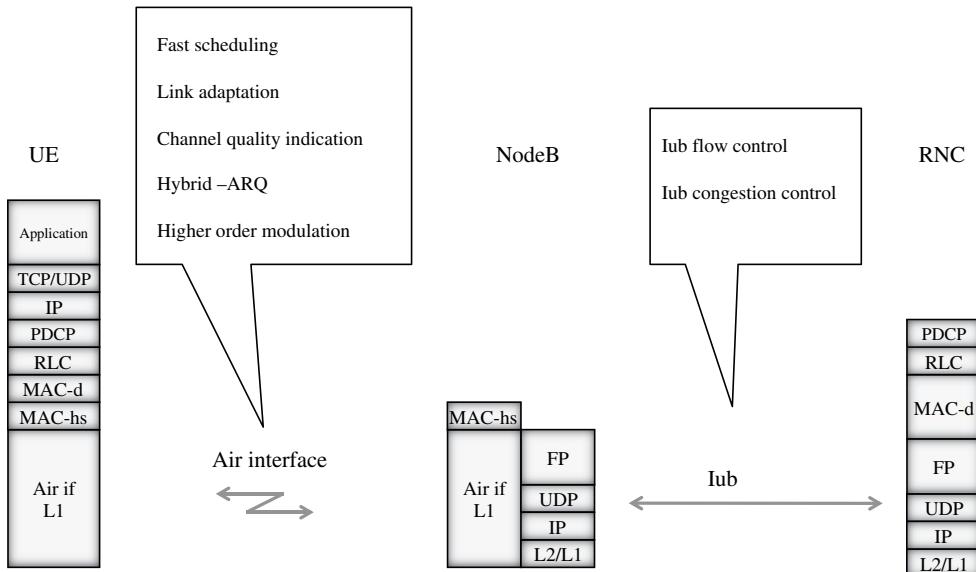


Figure 3.17 HSDPA introduced.

of 10 ms) where scheduling is done in the RNC. Channel quality indication (CQI) is received from the terminal to the mac-hs scheduling entity in the NodeB. The scheduler of the NodeB is now able to rapidly adjust to the varying air interface channel conditions, using link adaptation.

HSDPA supports a shared channel in the downlink. With HSDPA, a single user may get instantaneous high capacity for her data, after which another user is served. This approach fits packet switched services much better than the use of dedicated channels. Reading a web page may take 30 seconds, after which another page is requested. During the reading time, resources are given to transmit data to other users.

Scheduling algorithms include round robin, best C/I (Carrier/Interference), and weighted fair scheduling. Round robin is fair, since all users get a share of the bandwidth. However, simple round robin does not take into account channel conditions, and air interface capacity is not necessarily well utilized. Scheduling according to the best Carrier/Interference uses air interface capacity efficiently, but fairness among users may suffer. Weighted fair scheduling aims at combining efficiency and fairness.

HSDPA supports both non-guaranteed bit rate and guaranteed bit rate bearers. Guaranteed bit rate service allows streaming and conversational service. Background traffic use non-guaranteed bit rate. With non-guaranteed bit-rate, a nominal bit-rate is still defined.

MAC-hs in the NodeB includes the H-ARQ (Hybrid Automatic Repeat Request) function. NodeB is able to retransmit packets that are lost on the air interface. This leads to faster retransmissions and subsequently better performance. For the Iub backhaul the requirements originating from the RNC are now somewhat relaxed compared to Rel-99 DCHs. For the end user experience on HSPA based services, low latency at the Iub however remains essential.

While packets lost in the air interface do not need to be retransmitted from the RNC, packets that are lost at the Iub are still retransmitted by the RLC layer in the RNC (assuming an acknowledged mode RLC).

To the FP layer on the Iub, HSDPA introduces HS-DSCH capacity request and capacity allocation messages (HSDPA Iub flow control). NodeB allocates the amounts that can be sent from the RNC over the Iub. The target is that the MAC entities in the NodeB and in the RNC select packets that can be sent to the MAC-hs function. Capacity allocation defines the amount of data allowed in terms such as number of PDUs, interval during which the granted capacity is to be transmitted, and a maximum PDU size.

With 16-QAM modulation, the RLC data rate is over 13 Mbit/s per user (Category 10 terminal, 15 codes used). This is further enhanced with MIMO (Multiple input, multiple output), 64-QAM modulation, and a possible use of multiple 5 MHz carriers. (Recall that the rate with four carriers achieves 168 Mbit/s.) Even the data rates with the 16-QAM resulting in data rates over 10 Mbit/s are more than a decade larger than what is achievable with the DCH. (Assuming 384 kbit/s DCH for the comparison).

Considering a NodeB air interface upgraded to HSDPA, depending on the configuration and other factors, it means that the site can source well over 10 Mbit/s traffic to the Iub backhaul. HSDPA, as well as 3G in general, allows dimensioning the Iub transport separately from the air interface capacity. The supported Iub capacity may be less than what is the available peak data rate in the air interface. Clearly then congestion may occur, and Iub traffic types need to be prioritized so that control channels and real-time services (voice) are not impaired during congestion. HSPA service may be poor, however, due to the heavy congestion on the Iub. For HSDPA Iub dimensioning, see [40] with further references.

If the Iub has not been upgraded accordingly, it easily becomes a severe capacity bottleneck. In good radio conditions, the capacity allocations given based on the HSPA air interface exceed the available Iub capacity. In this situation, the throughput is poor, due to congestion on the Iub. Air interface capacity has been expanded with HSPA, but this will not lead to higher data rates for the users, unless the mobile backhaul is similarly expanded.

While congestion cannot be removed without ultimately adding Iub capacity, a congestion control function has been defined in 3GPP to address the topic. This can significantly improve the performance with a narrowband Iub. With HSDPA congestion control, delay and loss can be detected over the Iub between the NodeB and the RNC. HSDPA congestion control functions over the Iub, and addresses the backhaul bottleneck. This is discussed further in the QoS chapter.

During the introduction of the HSDPA, many of the existing NodeBs were backhauled with ATM Iub over E1/T1/JT1s, compliant with 3GPP Rel-99. Due to the need for higher HSDPA data rates, and cost pressures, adding E1/T1/JT1s is often not feasible economically. An alternative is to offload HSDPA traffic into a parallel IP/Ethernet path, meaning that HSDPA is mapped into an IP path while other traffic remains on ATM. This also relieves the congestion discussed previously. The approach is illustrated in Figure 3.18.

In 3GPP viewpoint this is a single logical Iub interface, where different traffic types comply with different transport protocol options: HSDPA is implemented with 3GPP Rel-5 IP transport, and Rel-99 DCH and control traffic (such as NBAP) with the Rel-99 ATM protocol stack.

HSDPA increased the data rates in downlink. High speed uplink packet access (HSUPA) in 3GPP Rel-6 introduced an E-DCH (enhanced DCH) to bring data speed improvement for the uplink transmission from the UE to the NodeB. E-DCH is still a dedicated channel as the name

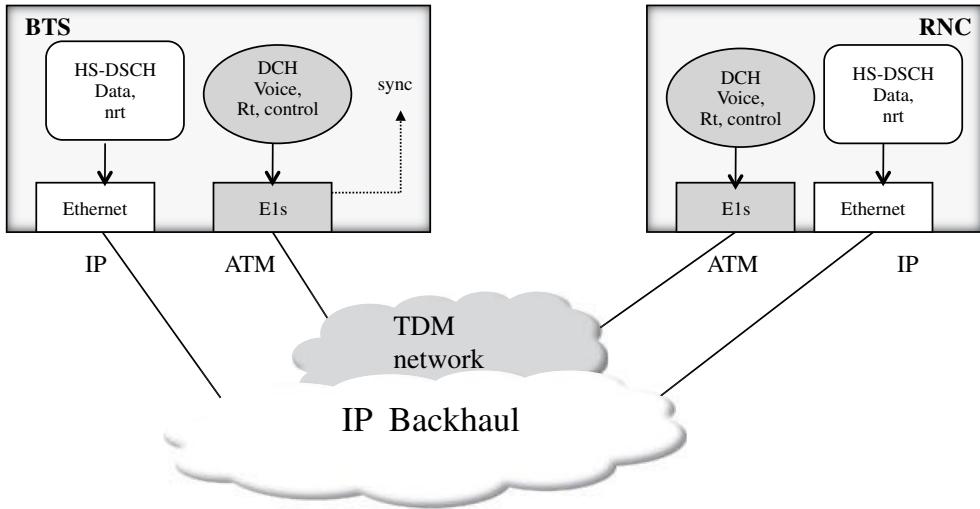


Figure 3.18 Use of IP/Ethernet to backhaul high speed data.

implies. E-DCH however borrowed the support of H-ARQ and fast MAC scheduling (10 ms, and 2 ms) from the HSDPA. As a dedicated channel, it has fast power control, and also supports soft handovers.

With HSUPA, the mac-e scheduler in the NodeB aims to achieve a high utilization, without causing an excessive noise rise in the cell.

Over the Iub, Frame protocol layer capacity requests and capacity allocations are supported, as well as a HSUPA congestion control functionality.

3.3.6 Iub

3G RAN (Radio access network) architecture is shown in Figure 3.19.

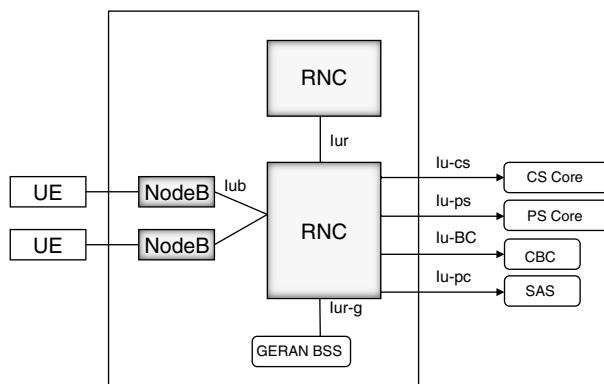


Figure 3.19 Interfaces of and within UTRAN [37].

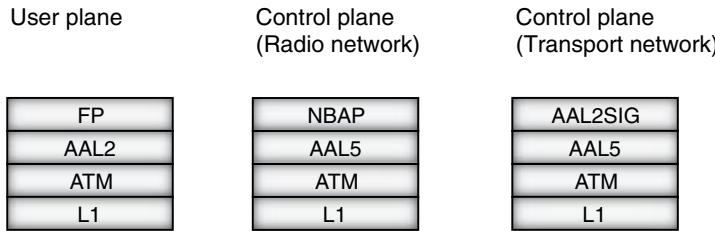


Figure 3.20 ATM based Iub interface [54], [58].

Each NodeB connects to the RNC via an Iub interface. From a transport point of view, all traffic from the NodeBs is directed to the RNC. Physical transmission topology of the Iub interface, and for the NodeB access, is not shown in the figure. Typically, the transmission topology is different from that of the radio network logical topology.

RNC interfaces the core network with the Iu interface (Iu-cs, Iu-ps, Iu-BC and Iu-pc). Iu-cs is the interface to the circuit-switched core, Iu-ps interfaces the packet-switched core, Iu-BC is for interfacing Cell Broadcast Centre (CBC), and Iu-pc interfaces Standalone Serving Mobile Location Centre (SAS). The RNC may be connected to BSS supporting GERAN Iu mode via the Iur-g interface.

From the mobile backhaul perspective, Iub, Iur, Iu-ps and Iu-cs are the interfaces that are discussed further, as these are typically always implemented in the UTRAN.

3G Iub interface initially used ATM as defined in 3GPP Rel-99. In 3GPP Rel-5, IP transport was introduced. ATM based protocol stacks for user and control plane are shown in Figure 3.20.

With the 3GPP initial Rel-99 standard, user plane traffic is carried over AAL2/ATM. Each user bearer maps to an AAL2 CID (Channel identifier), and these AAL2 bearers are set-up using AAL2 signaling. The AAL2 signaling VCC (Virtual Circuit Connection) is carried over AAL5/ATM. ATM allows separating traffic into VCCs of different characteristics, such as Constant Bit Rate (CBR), real-time Variable Bit Rate (rt-VBR) and Unspecified Bit Rate (UBR).

With ATM narrowband time-division-multiplexed E1, T1, JT1 interfaces are often used. Multiples of these physical interfaces can be combined with ATM IMA (Inverse Multiplexing over ATM). Also Sonet/SDH interfaces can be used with ATM.

ATM Iub can be emulated over the IP backhaul. ATM Virtual Circuits are, for example, mapped into MPLS pseudowires, or alternatively, the whole interface (E1/T1/JT1) traffic is mapped into a MPLS pseudowire.

Native IP based protocol stacks for user and control plane are shown in Figure 3.21 for the Iub interface.

Frame protocol is the user plane radio network layer protocol carried on top of the transport layer (UDP/IP transport in the figure). Each user plane bearer is mapped into an UDP port and an IP address. With IP, the user plane bearers over the Iub (UDP/IP connectivity) are established via NBAP signaling. There is no need for a transport specific signaling protocol like AAL2 signalling in the case of ATM. In the bearer set-up phase, the receiving node informs the IP address into which the bearer shall be terminated.

SCTP/IP is used for NBAP. NBAP is the control plane protocol for the radio network signaling between NodeB and the RNC. NBAP procedures are needed e.g. to bring up and

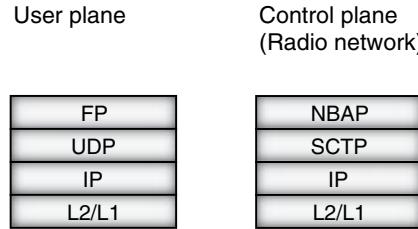


Figure 3.21 IP Iub [54], [58].

manage the cells. The same SCTP association is used for uplink and downlink directions, and bearers are mapped into a single stream in the uplink, and to a single stream in downlink direction. RFC3309 defined checksum shall be used.

Operation of the NBAP is critical for the NodeB. If IP connectivity is lost, SCTP times out (value depending on the parameters, e.g. after several seconds to tens of seconds) and consequently radio network layer recovery actions are started. The recovery may be a reconfiguration or a restart of the NodeB, which causes a service outage on that NodeB site.

For the IP layer, 3GPP Rel-5 defines IPv6 as mandatory, while IPv4 is optional, for all UTRAN terrestrial interfaces. However, a note mentions that this shall not preclude a single implementation of IPv4. IP addresses of the mobile network elements do not need to be publicly routable, so there is typically no pressing need for IPv6 from an addressing viewpoint. Private IPv4 addresses can be used.

For the data link layer, 3GPP UTRAN specifications require that with the IP transport option, PPP with HDLC framing shall be supported. However a note adds that a single implementation of any other Layer-2 protocol is not precluded. As a conclusion, for UTRAN both IPv4 and IPv6 is allowed and the underlying layers (Layer-2, Layer-1) are not defined.

With the introduction of HSDPA and HSUPA, the Iub transport protocol stack is not modified. Use of HSDPA and HSUPA is in standardization not coupled to the introduction of the IP transport. In practice, the data rates often require an expansion of the backhaul network – which then leads to the introduction of IP transport. To the FP layer, new HS-DSCH Frame protocol and E-DCH frame protocol formats are introduced with HSPA.

3.3.7 *Iur*

Iur interface connects the RNCs to the neighbour RNCs. Protocol stacks for the IP based *Iur* is shown in Figure 3.22. IP protocol stack was included into the standard in 3GPP Rel-5. Prior to that, *Iur* is ATM based (3GPP Rel-99 definition).

3.3.8 *Iu-cs*

Iu-cs interface connects the UTRAN to the circuit-switched core network. Protocol stacks for the IP based *Iu-cs* are shown in Figure 3.23. IP protocol stack was included into the standard in 3GPP Rel-5. Prior to that, the *Iu-cs* is ATM based (3GPP Rel-99 definition).

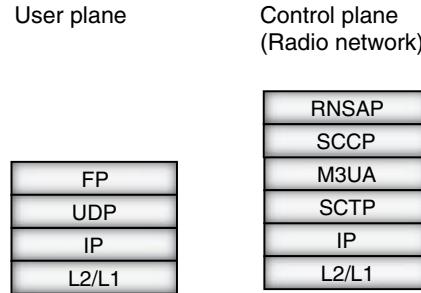


Figure 3.22 IP based Iur [50], [54].

In the user plane, RTP protocol is used, over the UDP/IP. As a difference to A over IP protocols, RTP multiplexing is not defined for the Iu-cs. In the control plane, RANAP is mapped into the SIGTRAN protocol stack, SCCP/M3UA/SCTP/IP.

User plane bearer IDs, including IP addresses, are exchanged via the RANAP. In the bearer set-up phase the receiving node informs of the IP address into which the bearer shall be terminated.

3GPP defines IPv6 as mandatory, while IPv4 is optional, for all UTRAN terrestrial interfaces. However, a note mentions that this shall not preclude a single implementation of IPv4.

For the data link layer, 3GPP specifications require that with the IP transport option, PPP with HDLC framing shall be supported. However a note adds that a single implementation of any other Layer-2 protocol is not precluded. As a conclusion, for UTRAN both IPv4 and IPv6 is allowed and the underlying layers (Layer-2, Layer-1) are not defined.

3.3.9 Iu-ps

Iu-ps interface connects the UTRAN to the packet-switched core network. Protocol stacks for the IP based Iu-ps is shown in Figure 3.24. IP protocol stack was included into the standard in 3GPP Rel-5. Prior to that, the Iu-ps is ATM based (3GPP Rel-99 definition).

3GPP Rel-5 IP specifications define IPv6 as mandatory, while IPv4 is optional, for all UTRAN terrestrial interfaces. However, a note mentions that this shall not preclude a single implementation of IPv4.

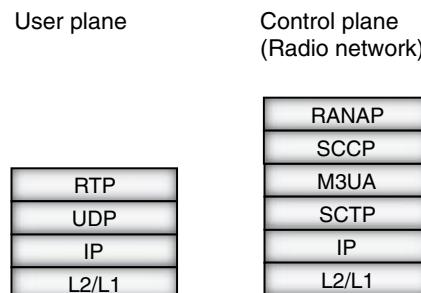


Figure 3.23 IP Iu-ps [41].

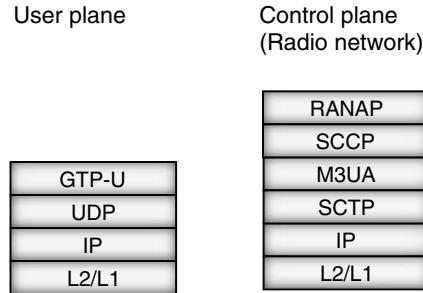


Figure 3.24 IP Iu-ps [41].

User plane bearer IDs, including IP addresses, are exchanged via the RANAP. In the bearer set-up phase the receiving node informs into which IP address the bearer shall be terminated.

For the data link layer, 3GPP UTRAN specifications require that with the IP transport option, PPP with HDLC framing shall be supported. However, a note adds that a single implementation of any other Layer-2 protocol is not precluded (as defined in 3GPP TS25.414). As a conclusion, for UTRAN both IPv4 and IPv6 are allowed and the underlying layers (Layer-2, Layer-1) are not defined.

3.3.10 GTP-U Protocol

GTP-U (GPRS Tunneling Protocol) tunnels user IP from/to the RNC, over the Iu-ps interface. GTP protocol has two variants, GTP-U (user plane), and GTP-C (control plane). GTP-C protocol is not needed in the radio access network. Additionally, there are different versions of the protocol. For GTP-U, v1 is mandated in Rel-8 and onwards and the GTP-Uv1 specification is TS29.281. Before Rel-8, the normative reference of GTP-U is 3GPP TS29.060. GTP-v1 protocol does not need to listen to GTPv0, a well-known port 3386, and GTPv0 messages can be silently ignored.

A GTP-U tunnel is identified using a Tunnel Endpoint Identifier (TEID), IP address, and UDP port number. The UDP destination port number for GTP-U is 2152. The UDP source port is locally allocated by the sending node.

TEID value is exchanged by the radio network control plane protocol, RANAP. One GTP-U protocol entity exists for an IP Address. TEID allows multiplexing users, different packet protocols and different QoS levels. Different GTP-U Endpoints use different TEID values.

At the Iu-ps interface, 3GPP TS 29.281 defines that IPv4 shall be supported, and IPv6 should be supported. The Iu interface specification, 3GPPTS 25.414 defines that with the IP transport option, the RNC/core network node shall support IPv6, while the support of IPv4 is optional. A note adds that a single implementation of IPv4 is not precluded. There is seemingly contradiction, however, in practice it is a question of interoperation between RAN and core network, and on selecting a common protocol.

Figure 3.25 shows the GTP-U header.

The header is at minimum 8 bytes, as the fields that are always present are the version field, protocol type (PT), a Spare Bit (marked*, sent as '0' and not evaluated by the receiver), extension header flag (E), sequence number flag (S), N-PDU number flag (PN), Message type,

Octets	Bits							
	8	7	6	5	4	3	2	1
1	Version		PT	*	E	S	PN	
2	Message type							
3	Length							
4								
5								
6	Tunnel Endpoint identifier							
7								
8								
9	Sequence number							
10								
11	N-PDU number							
12	Next extension header type							

Figure 3.25 GTP-U header [70].

Length and Tunnel Endpoint Identifier. Optional header fields are sequence number, N-PDU number, and next extension header type.

Version field defines the version of the protocol, with GTP-U v1, the number is ‘1’. Protocol type differentiates between GTP (value ‘1’) and GTP’ (value ‘0’) protocols.

Flags E, S, and N, indicate the presence of the corresponding headers, E for Extension header, S for Sequence number, and PN for N-PDU Number field.

GTP-U Message Types are shown in Figure 3.26.

Length field means the length of the payload in octets. The optional header fields are considered payload.

Message Type	Message
1	Echo Request
2	Echo Response
26	Error indication
31	Supported Extension headers notification
254	End marker
255	G-PDU

Figure 3.26 GTP-U message types [70].

Tunnel endpoint identifier, TEID, identifies the endpoint in the receiving entity. The TEID value is assigned locally by the receiving side.

For the optional fields, Sequence number is for preserving transmission order. N-PDU number field is used in coordinating acknowledged mode communication and needed in inter SGSN routing area update and with certain inter-system handover cases. Next extension header type defines the type of the extension header that follows.

GTP-U messaging supports both signalling messages and user plane G-PDU messages. Signalling messages are intended for user plane path/tunnel management. User plane messages, G-PDUs, carry the original packets, T-PDUs.

For the signalling messages, Echo request and Echo response can be used to detect that the peer is alive. The frequency of the messages is implementation specific. Supported extension header notification message indicates which extension headers are supported. Error indication is defined e.g. to inform of a received GTP-U PDU, which does not have a valid context (RAB in the case of Iu-ps). End marker indicates the end of a payload stream of tunnel. G-PDUs arriving on the tunnel after the end marker may be silently discarded.

User plane message (G-PDU) structure is GTP-U header, followed by a T-PDU. T-PDU is the inner IP packet in the GTP-U packet. The inner IP packet is an IP packet sent to the terminal in the downlink direction from the external network (identified by the APN), or, in the uplink direction, an IP packet sent from the terminal to the external network. One or more tunnels may be used.

3GPP TS TS23.060 defines the maximum inner IP packet size that can be transmitted without fragmentation in the terminal or GGSN, to be 1502 octets (for PDP of type PPP), or 1500 octets (for other cases). For the MTU calculation of the Iu-ps links (outer IP packet), IP, UDP, and GTP-U headers have to be taken into account in order to avoid fragmentation. Fragmentation is avoided, if links can be configured to support the required outer packet MTU.

3.4 LTE

3.4.1 Architecture

LTE system architecture is shown in Figure 3.27.

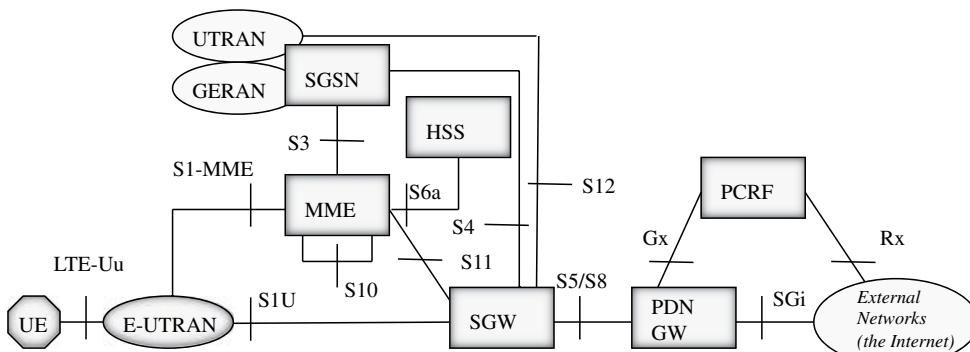


Figure 3.27 LTE architecture [72].

In the LTE system, the radio access network is called E-UTRAN (Evolved UTRAN). LTE is a packet switched network with no native support for circuit switched services. Instead, services of a circuit switched nature, e.g. voice, can be supported through the use of guaranteed bit rate bearers. The core network is clearly divided into a control plane entity MME (Mobility Management Entity) and a user plane entity consisting of a SGW (Serving Gateway) and a PDN GW (Packet Data Network Gateway).

E-UTRAN interfaces the core network elements via S1-U in the user plane, and via S1-MME in the control plane. MME is responsible for managing the bearers, authenticating the UEs, and managing mobility. The role of MME is comparable to that of the SGSN control plane.

In the user plane, SGW interfaces the radio network (eNodeB). It is the anchor point to the inter-eNodeB mobility and for handovers towards 2G and 3G systems. This is comparable to the SGSN user plane functionality. PDN GW interfaces the LTE system to the external networks, and it allocates IP addresses to the UEs. This is comparable to the GGSN role of 2G and 3G networks.

MME has an interface for the Home Subscriber Server (HSS), which stores subscriber information. HSS also supports the authentication of the subscriber, instead of a separate authentication centre.

The E-UTRAN architecture is shown in Figure 3.28.

E-UTRAN interfaces the core network via the S1 interface. S1 interface consists of S1-MME which is a control plane interface between an eNodeB and MME, and of S1-U as a user plane interface, between an eNodeB and the SGW. In the figure the user and control plane S1 interface are not shown separately.

For resilience against core network node failures, eNodeBs can interface many core network elements. This is shown in Figure 3.28 by multiple parallel S1 lines (the rightmost eNodeB).

Within an E-UTRAN, eNodeBs interface other eNodeBs using the X2 interface. X2 interface is used for handovers, so eNodeBs, with their coverage areas overlapping, can benefit from the X2 interface. If X2 interface is not configured, the handover uses S1 interface

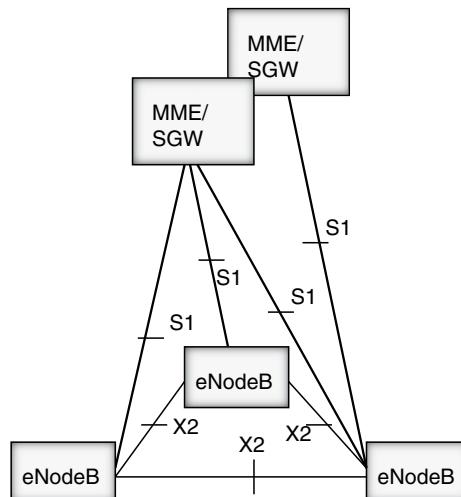


Figure 3.28 E-UTRAN architecture [73].

Table 3.1 Functions in LTE [72].

Element	Function
eNodeB	Inter cell RRM
eNodeB	Radio bearer control
eNodeB	Connection mobility control
eNodeB	Radio admission control
eNodeB	eNodeB Measurement configuration and provisioning
eNodeB	Dynamic resource allocation (scheduling)
eNodeB	Protocols: RRC, PDCP, RLC, MAC, PHY
MME	NAS security
MME	Idle state mobility handling
MME	EPS bearer control
S-GW	Mobility anchoring
P-GW	UE IP Address allocation
P-GW	Packet filtering

instead, so X2 is an optimization. If possible, it is preferable to perform an X2 handover compared to the S1 handover. The signaling procedure is simpler and allows for a loss-less handover.

X2 is a logical interface. In a practical mobile backhaul implementation, X2 traffic would often be routed via an aggregation node higher in the access/aggregation tier of the network, as implementing a direct physical layer connectivity between adjacent eNodeBs is costly. X2 consists of user and control traffic during an X2 handover. Traffic volume of X2 is not significant, however, it is an important interface that needs to be considered in the backhaul design.

Comparing the E-UTRAN architecture to the 3G UTRAN architecture, there are key differences from the mobile backhaul viewpoint. First, the UTRAN is collapsed into a single node (eNodeB), and consequently there is no need for a controller-BTS interface, which would impose real-time requirements to the backhaul. Second, X2 introduces direct eNodeB-eNodeB traffic, which does not exist in 2G or 3G. Third, LTE is natively an all-IP network, as there is no other standardized transport alternative, like ATM in 3G.

A summary of key functions of elements is shown in Table 3.1.

In LTE, all the radio network functions are managed by the eNodeB only, as it is the only node in the E-UTRAN radio network. This is a main difference to 2G and 3G, where radio protocols are distributed between the BTS and the controller. Now these protocols (PDCP, RLC, MAC, RRC, etc) are all in the eNodeB, as well as the interface to the core network.

For further reading on LTE, see [69], [70] and [71].

3.4.2 Packet Switched Traffic

User plane protocol stack is shown in Figure 3.29.

IP traffic from the UE is carried over the Uu air interface using PDCP/RLC/MAC layers. Over the S1u interface, user IP packets are tunneled using the GTP-U protocol. Packet data network gateway interfaces the external network (e.g. the Internet) using the SGi interface.

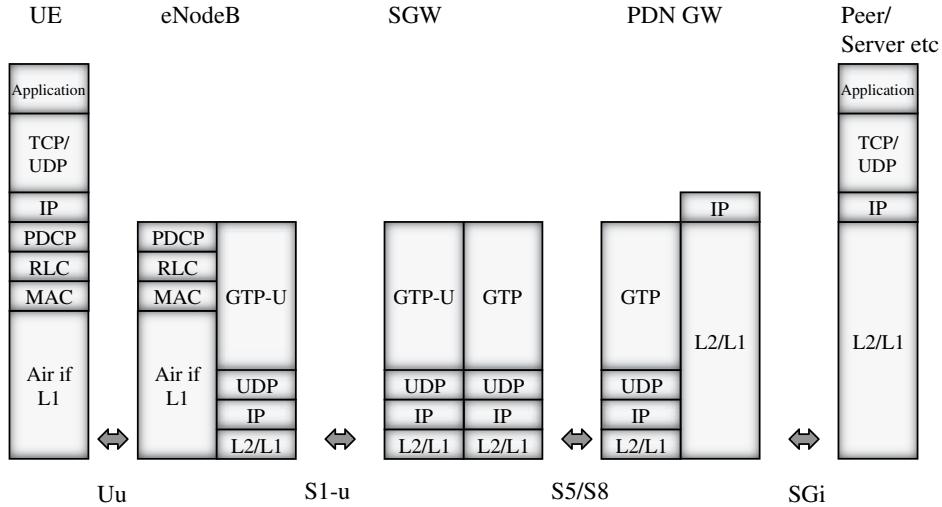


Figure 3.29 LTE user plane [73].

As a difference to 2G or 3G, now the air interface channels all terminate in the eNodeB. These channels do not need to be carried over the backhaul. This simplifies the backhaul considerably, since the air interface performance is not as directly coupled to the backhaul as is the case with Abis or Iub.

Note that S5/S8 interface has two options, 3GPP protocol stack option (GTP based), as shown in the figure above, and IETF Mobile Ipv6 (Proxy Mobile IP) as another option.

Control plane protocol stack is shown in Figure 3.30, for both Uu (air interface) and for the S1-MME interface.

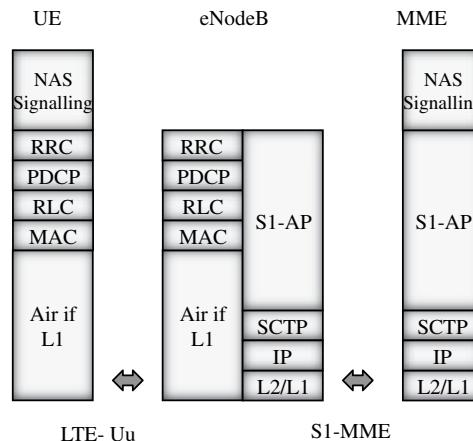


Figure 3.30 LTE control plane [73].

RRC (Radio Resource Control) is now located at the eNodeB, as there is no controller element. RRC functions include radio resource management functions; such as setting up and deleting radio bearers, air interface admission control, and managing handovers.

PDCP layer includes ciphering functionality in addition to the IP header compression. Both user plane IP traffic, and the RRC signaling messages are encrypted.

UE communicates with the MME with the non-access stratum (NAS) signaling, which is transparent to the eNodeB and to the E-UTRAN. This signaling is relayed by the eNodeB to the MME.

3.4.3 Air Interface

The air interface of LTE is using OFDMA in downlink (Orthogonal Frequency Division Multiple Access) and single-carrier FDMA (SCFDM) in uplink. In the uplink, power consumption of the terminal is enhanced when using SCFDM compared to OFDM. This is a key design criterium for a mobile system.

In the frequency domain, OFDMA is divided into subcarriers, with 15kHz spacing. By using a different amount of subcarriers, spectrum can be used flexibly, up to a 20 MHz maximum spectra. This alone increases the data speeds significantly over previous systems (3G WCDMA carrier is 5 MHz). Note however the possibility for dual-cell and carrier aggregation with HSPA). Additionally, as opposed to W-CDMA, scheduling can be done in the frequency domain in addition to scheduling in the time domain.

Adaptive modulation and coding (AMC) is used. This means that modulation and coding schemes are varied and selected, to optimize spectrum usage and throughput in the given channel conditions. When the channel quality is good, high bit rates are achieved by a high order modulation, such as 64-QAM. Another modulation, such as QPSK, provides better coverage, although the data rate is lower. Coding scheme means adding redundant bits to the information, for a coding gain. Varying the amount of redundant information allows fine tuning the communication to the channel conditions.

H-ARQ (Hybrid Automatic Repeat Request) procedure at Layer-1 is used to retransmit those frames that are not received correctly over the air interface.

To illustrate the downlink peak rates of LTE, 3GPP TS36.306 defines e.g. the following UE capabilities (see also [71]):

- Category 3 UE : 100 Mbit/s
- Category 4 UE : 150 Mbit/s
- Category 6 and 7 UEs (3GPP Rel-10) : 300 Mbit/s
- Category 8 UE (Further capability) : 3 Gbit/s.

The 3GPP Rel-10 UE categories with 300 Mbit/s peak downlink data rates provide the next step (with solutions such as carrier aggregation) after the first phase of LTE. A theoretical peak data rate capability of 3GBit/s exists with a Category 8 UE. This would require a large bandwidth of 100 MHz and the support of eight antennas in the UE. Thus it is not expected to enter the LTE market in the near future.

3.4.4 S1

S1 interfaces are shown in Figure 3.31.

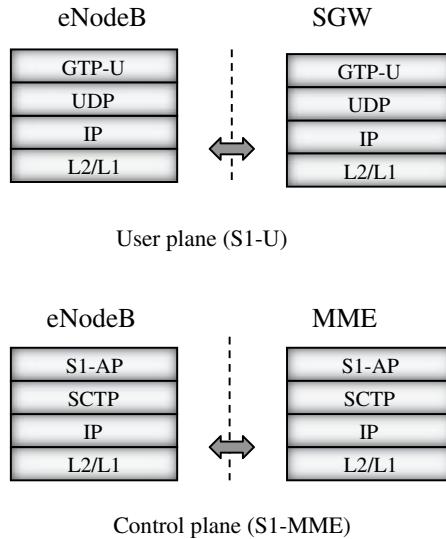


Figure 3.31 S1 interface for user and control plane [75], [77], [79].

An eNode B communicates directly with the core network in the user plane over the S1-u interface. The protocol stack for the transport consists of GTP-U layer carried over UDP/IP. Layers underlying IP are not standardized, and can be for example Ethernet or Point-to-point protocol.

In the control plane (S1-MME interface) uses S1-AP (S1 application protocol) between the eNodeB and the MME. SCTP/IP is the underlying protocol layer for S1-AP.

As the PDCP layer is terminated in the eNodeB, user plane traffic is passed on to the core network unprotected. IPsec protocol is typically deployed for the S1-u interface to provide for confidentiality, authentication, and encryption.

GTP-U tunneling protocol carries the user IP packets to the SGW. A GTP tunnel is identified by the GTP tunneling endpoint identifiers (GTPU-TE IDs). GTP-U uses UDP/IP. IP typically utilizes Ethernet as the physical interface, and as the L2 protocol, but this is not covered in the 3GPP specifications, so basically any L2/L1 that carries IP can be used.

For LTE, 3GPP TS36.412 (signalling) and 3GPP TS36.414 (data transport) specify the use of IPv6 and/or IPv4. GTPv1-User plane specification, 3GPP TS29.281 has IPv4 mandatory ('shall' be supported) while IPv6 'should' be supported. Also with LTE, private addresses can be used in the mobile network elements, so there is typically no driver for IPv6 from an address shortage point of view. At the IP layer, fragmentation is required for the eNodeB and for the EPC (Evolved Packet Core) by 3GPP TS36.414.

3.4.5 X2

X2 is an optional interface. If it does not exist, there is no parallel communication between eNodeBs, and handovers via X2 are not possible, in which case handovers are supported only via the S1 interface. Benefits of the X2 interface are in having a reduced signaling for inter-eNodeB handovers, the capability for loss-less X2 handovers, shorter interruption time in data transfer, and a reduced latency in the handover execution.

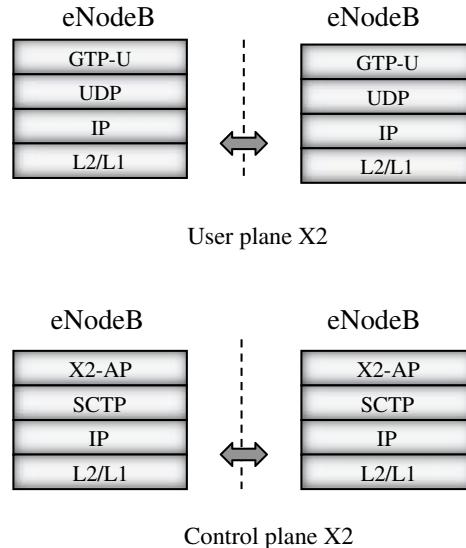


Figure 3.32 User and control plane stack for X2 [80], [82], [84].

X2 does not need to be a physical direct link between two eNodeBs, but traffic can be routed via a common L2 or IP point higher in the network. In the L2 connectivity case, the neighbour eNodeBs need to share the same VLAN. With IP routing is used.

X2 interface is used during the inter-eNodeB handover. While the handover is in progress, packets arriving in the downlink direction can be directed to the target eNodeB using the X2 interface. The benefit of this is in avoiding loss of traffic during the handover process. In the uplink, the UE buffers the information until communication via the new target eNodeB can start.

X2 User and control plane protocol stacks are shown in Figure 3.32. At the control plane, the protocol stack of the X2 interface between eNodeBs, is identical to that of the S1-MME, with, however, X2-AP as the application protocol between two eNodeBs. At the user plane, GTP-U tunneling of user packets is used.

3.4.6 Bearers

With LTE, a default EPS bearer is established when the UE is attached to the network/enters the ECM connected state. This bearer can be used immediately for transmitting user traffic. With the EPS bearer an IP address is allocated to the UE.

The default bearer removes the need to frequently set up and remove bearers based on user activity, as is the case with dedicated channels of 2G and 3G. In particular, low volume data, keep-alives and status updates to web pages can now be efficiently transmitted, without extra signalling. A default bearer can also be released in case of inactivity, but the duration of the timer can be longer than in 2G or 3G.

EPS Bearer consists of a radio bearer, S1 bearer, and of S5/S8 bearer (see Figure 3.33). The radio bearer connects the UE to the eNodeB, S1 bearer between the eNodeB and the SGW, and

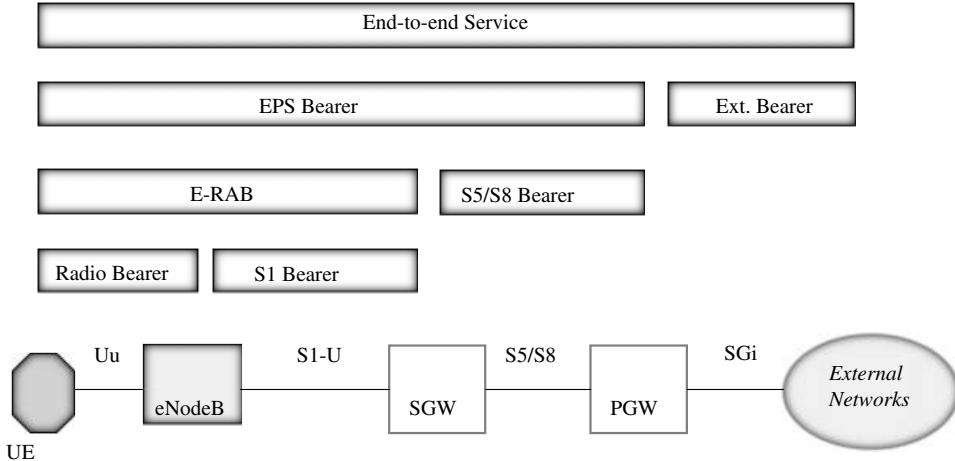


Figure 3.33 LTE bearers [72].

S5/S8 bearer from the SGW to the PGW. The end-to-end service, UE connectivity to the Internet server (or equivalent), takes place via the external bearer over the SGi interface.

As an analogy to the 3G system, the PDP context of 3G is equivalent to the EPS bearer. 3G Radio access bearer would map to the radio bearer and the S1 bearer.

EPS bearer is identified by an ID given by the MME. The S1 bearer is set up by the MME, and is a GTP tunnel, identified by the GTP tunnel endpoint identifiers (TEIDs). The radio bearer is set up by the eNodeB. Due to inactivity, S1 bearer and related resources may be released (S1 release request).

For the S5/S8 bearer there are two options: GTP option, identified by the TEID, or IETF based Proxy Mobile IPv6 tunnel. S5/S8 bearer is established by the SGW and the PGW.

The need to set up the bearer may come either from the UE or from the PDN. If data arrives for the UE from the PDN, EPS bearer establishment is initiated by the PGW. This happens via the PCRF (Policy and Charging Rules Function). According to the QoS requirements, the EPS bearer may be a dedicated bearer. As already described, a default bearer is set up by the MME, when the UE attaches to the network.

The concept of EPS bearers is related to the Quality of Service in LTE. Each bearer is associated with a Quality of Service class. This class is identified by the Quality of Service identifier (QCI). QoS is discussed further in a dedicated chapter in Part II of this book.

3.4.7 Mobility Management

Simplified, MME manages the mobility when the UE is in the ECM (EPS Connection Management) idle state. In the ECM connected state, mobility is managed by handovers using eNodeB.

MME keeps track of the location of the UE in the network. When the UE is in the EMM (EPS Mobility Management) deregistered state, there is no information of the location of the UE. For UE to send or receive data (data arriving to PDN GW), a state change to EMM registered is required.

Change to the EMM registered state is accomplished via the attach procedure or via a tracking area update. If the trigger for the state change is the data arrival at the PDN GW, MME pages the UE. UE needs an RRC connection over the air interface to eNodeB. RRC connection is established using a contention based random access procedure.

UE sends an attach request which is forwarded by the eNodeB to the MME. Subscription information is provided by the HSS to the MME, and the MME sets up a user plane connection (default EPS bearer). An IP address is allocated to the LTE terminal by the PDN GW. Default EPS bearer includes the radio bearer, the S1 bearer, and the S5/S8 bearer. For the S1 bearer GTP-U tunnels are set up. With LTE attach, the UE state is changed to EMM registered, and ECM connected. At the same time, the initial EPS default bearer is established, and an IP address is given to the UE.

The MME keeps track of the UE location at a tracking area level. Tracking area consists of multiple cells. UE may be registered to one or multiple tracking areas. If UE changes location to a tracking area where it is not registered, it needs to update the MME (a tracking area update). This way the MME stays up-to-date of the current UE location. By paging MME may instruct the UE to move to the ECM connected state when required.

When the UE is in the EMM registered state, MME knows the location of the UE at a tracking area level. When the UE is in the ECM connected state, cell level location of the UE is available.

Connectivity to the PDN GW and to the external network is active for the life time of the PDN connectivity. Due to inactivity, the ECM state may be changed to the ECM idle state. In this state, the UE still holds an IP address. This enables moving fast back to an active state.

When the UE is in the ECM connected state, mobility is managed by handovers. Intra-eNodeB handover takes place within the same eNodeB (different cells). Inter-eNodeB handover takes place either via the X2 interface or via the S1 interface.

A mobility anchor point means that the external PDN network continues to see the same ‘location’ of the UE, even though the connectivity internally in the mobile system from the UE to this anchor point changes. Mobility anchor point depends on the type of handover. With intra-eNodeB handover (handover between cells of a single eNodeB), the anchor point is the eNodeB. With inter-eNodeB handover, SGW serves as the anchor point. With inter-RAT handovers, SGW is also the anchor point, and with non-3GPP system handover, it is the PDN GW.

As eNodeBs are responsible for the radio resource management, measurement reports from the UE are sent to the eNodeB. The decision to move to a new target cell, is done by the source eNodeB. Source eNodeB requests a handover to the target eNodeB. Target eNodeB acknowledges the handover (if resources can be granted) and allocates a new temporary identifier (C-RNTI). UE then synchronizes with this new target cell. PDCP sequence numbers are used between the eNodeBs to detect already transferred packets – enabling a lossless handover.

Inter-eNodeB handover takes place via the X2 logical interface, when existing. If X2 interface does not exist, the handover is performed over the S1 interface.

When the UE has an RRC connection to the eNodeB, it is in RRC connected state. When additionally a S1 bearer exists in the user plane, the UE is in ECM connected state. In this state, MME and UE are able to exchange signaling messages. These are referred to as non-access stratum signaling messages, as they are transparent to the E-UTRAN.

3.4.8 Interworking with 2G and 3G

Mobility is supported within the different 3GPP radio access technologies (RATs). Coverage of LTE is initially limited, so it is beneficial if the UE can be served by 2G or 3G radio access network as well. This naturally requires a terminal supporting multiple radio access network types.

With a handover between the LTE and a 3G network, SGW serves as the mobility anchor point. New S3 and S4 interfaces are required. In the control plane, SGSN interfaces the MME using the S3 interface, and the SGW via the S4 interface. Both S3 and S4 are new logical interfaces for the SGSN. Alternatively, a direct tunnel can be set up between the RNC and the SGW via the S12 interface. Interworking with 2G radio access follows the same principles.

The control plane S3 interface is based on GTP-C, a control plane version of the GPRS tunneling protocol. GTP-C is carried on top of UDP/IP. S4 interface is using GTP (and GTP-C for control plane) as well, over the UDP/IP.

LTE system can also interface non-3GPP networks, such as WiMax. In this case the mobility anchor point is the PDN GW.

3.4.9 Voice Support

As there is no circuit-switched core in the LTE system, voice service (and messaging service, such as text messages) has to be supported by an alternate way. Multiple possibilities exist. First, voice service may be delivered as voice over IP (VoIP) with the IMS (IP multimedia subsystem). Second, LTE may be used as an IP network providing connectivity for VoIP clients (such as Skype™) of the terminal and for the peer somewhere in the PDN (the Internet). This solution is sometimes referred to as ‘over the top’, since the mobile network is now a transparent IP bearer used by the application layer clients. Third, circuit switched fallback (CSFB) is defined, which redirects the terminal to use 2G or 3G system for the circuit-switched voice call.

Additionally, single-radio voice call continuity (SRVCC) is defined to enable an IMS-based VoIP call to continue as a circuit switched voice call in a 2G or a 3G network. This is useful in case the coverage area of LTE is limited. Voice calls started in the LTE system do not need to be dropped at the edge of the LTE coverage, if 2G or 3G system still provide service.

In the Voice over LTE/IMS-based solution, the LTE core network needs to include the IMS system and interfaces, and terminals need to have a VoIP/IMS compatible client. SRVCC may be used to complement the solution, in a phase when LTE coverage is still thin.

With circuit switched fallback, a circuit-switched voice call is set-up via the 2G or 3G system. In the mobile originated call, the eNodeB instruct the UE to move to a 2G or a 3G network. Once UE is in the 2G or 3G network a circuit-switched voice call is set-up via this system. Similarly in the case of a mobile terminated call, paging occurs via the LTE system, and then the circuit-switched voice call is initiated in the 2G or 3G network. Circuit-switched fall-back may be considered as an interim solution, as ultimately voice is assumed to be supported without the circuit-switched core.

GSM Association (GSMA) announced the Voice over LTE (VoLTE) initiative in 2010, targeting for a standard way to support voice and messaging services over LTE. The GSMA proposal is based on the IMS specifications for support of voice over LTE.

3.4.10 Self Configuration and Self-Optimization

The LTE system includes functionality for self-configuration and self-optimization. With self-configuration, new nodes obtain the necessary basic configuration for system operation. With self-optimization process, network can be optimized using measurements from UE and the eNodeB.

Self-configuration assumes a pre-operational state; eNodeB is powered on and has backbone connectivity, but the RF transmitter is off. With the self-configuration process, basic set-up is obtained, as well as the initial radio configuration. Self-configuration involves setting up the S1-MME and the X2 interfaces.

For the S1-MME, it is assumed that the eNodeB is aware of the remote end-point address(es) of the MME(s), for creating an SCTP association. How this is realized, is outside the scope of the LTE specifications. In Chapter 4 some generic IP networking protocols for obtaining IP addresses and related configuration, are discussed.

With the end-point-address available, eNodeB tries to establish an SCTP connection to the MME. When the SCTP association is successfully established, eNodeB and MME exchange application layer configuration information. This includes e.g. Tracking area and PLMN ID information. After this initialization is completed, S1-MME interface is operational.

Automatic neighbour relation (ANR) function aims at automatically finding out (radio network) neighbours. Neighbour cell relations consists of source-cell – target-cell – relations. Each relation has three attributes: No remove-attribute means that the relation shall not be removed from the Neighbour relation table (NRT). No HO (handover) – attribute means that the relation is not to be used for handovers. No X2 – means that the X2 interface shall not be used.

In addition to the ANR function, O&M may modify the neighbour relations and the attributes in the NRT. How the ANR function finds out the radio network neighbours, is discussed in further detail in 3GPP TS36.300.

For the X2, the process is similar to the S1-MME. eNodeB establishes an SCTP association towards the remote end point, with the known remote IP endpoint. For obtaining the X2 remote end point address, the eNodeB may additionally use information obtained via the ANR. If the eNodeB has learned the eNodeB ID of a candidate X2 peer by the ANR, it can then request from MME the IP address of this candidate X2 peer, identified by the eNodeB ID. Configuration transfer message is sent to the MME, which relays the request to the eNodeB holding the eNodeB ID. In the reply, suitable transport layer (IP) addresses are given to the MME. MME relays this information to the requesting eNodeB, and SCTP association establishment can proceed.

3.5 Summary

Mobile backhaul is an integral part of the mobile network. Logically, the mobile network layer is isolated from the transport network layer. This accounts for a ‘clean’ interface design between the radio network and the transport/backhaul network layers. In practice, interdependencies exist: many functions and characteristics in the backhaul do have an impact on the mobile system performance, and to the end user experience of the service. Also, characteristics of the mobile network influence the backhaul design.

The capabilities of mobile systems have evolved. 2G shifted voice largely to the mobile network. 3G/HSPA, HSPA evolution, and LTE are shifting broadband connectivity as well to the mobile network. With this type of evolution, expansion of the mobile backhaul is often needed. Flat rate mobile broadband tariffs, spectral efficiency improvements, and high data rates of the new radio technologies, have led to an explosion of network traffic in terms of Mbytes per user per month. In 3GPP, this is addressed by defining IP based protocol stacks for the logical interfaces.

3GPP is in general agnostic to layers below the IP layer. For the IP layer, 3GPP includes requirements for both IPv4 and IPv6, and as well in some cases for IPsec. The way the logical interface definitions are written varies in interface standards. As a simplified summary, the backhaul may from a 3GPP viewpoint be implemented with any L2/L1 technology. Also for IPv4/IPv6 it is a network implementation issue. As IPv4 private addressing can be used for the mobile network elements (apart from a number of core network interfaces), there typically is no shortage of IPv4 addresses. Other benefits of IPv6 remain.

In the 2G system, GERAN interfaces towards the core network (A and Gb) include IP based protocol options. On the other hand, the GERAN internal BSC-BTS interface (Abis) does not have an IP based standardized protocol option. The only standard compliant definition is the initial TDM-based Abis. Vendor specific native IP implementations exist. Another alternative for an evolution to a packet based Abis is to deploy a circuit emulation service, which transports the TDM-based Abis over a packet network.

Initially, 3G UTRAN interfaces were defined as ATM based. 3GPP Rel-5 introduced an IP transport alternative for all of the UTRAN interfaces (Iub, Iur, Iu-cs, Iu-ps). As a difference to 2G, also the UTRAN internal RNC-NodeB interface has a standard compliant definition for IP transport. LTE is an all-IP network from the start and only IP based logical interfaces are defined.

LTE has simplified the system architecture by removing network elements and circuit switched functions, and by optimizing the system for an always-on IP connectivity. Main benefits are high speeds and low latency, making LTE well suited for mobile broadband. For the delivery of services like voice, different alternatives have been specified, including a fallback to an existing 2G/3G network for the voice call.

References

3.1

- [1] 3GPP TS 23.101: ‘Universal Mobile Telecommunications System (UMTS): General UMTS Architecture’, v10.0.0
- [2] http://www.3gpp.org/ftp/Inbox/2008_web_files/3GPP_Scopeando310807.pdf, retrieved Oct 2011.
- [3] 3GPP Overview of 3GPP Release 99, v0.1.1
- [4] 3GPP Overview of 3GPP Release 4, v1.1.2
- [5] 3GPP Overview of 3GPP Release 5, v0.1.1
- [6] 3GPP Overview of 3GPP Release 6, v0.1.1
- [7] 3GPP Overview of 3GPP Release 7, v0.9.15
- [8] 3GPP Overview of 3GPP Release 8, v0.2.4
- [9] 3GPP Overview of 3GPP Release 9, v0.2.3
- [10] 3GPP Overview of 3GPP Release 10, v0.1.2
- [11] 3GPP Overview of 3GPP Release 11, v0.0.1

3.2 2G/GERAN

- [12] 3GPP TS43.051 GSM/EDGE Radio Access Network (GERAN); Overall Description. v10.0.0
- [13] Eberspächer, Vögel, Bettstetter: GSM Switching, Services and Protocols, Second Edition, Wiley 2001.
- [14] 3GPP TS 08.01 General Aspects on the BSS-MSC Interface, v8.0.0
- [15] 3GPP TS 08.02 BSS-MSC interface; Interface principles, v8.0.0
- [16] 3GPP TS 08.04 BSS-MSC interface; Layer 1 specification, v8.0.0
- [17] 3GPP TS 08.06 Signalling transport mechanism (BSS - MSC) interface, v8.0.0
- [18] 3GPP TS 08.08 MSC-BSS interface; Layer 3 specification v9.0.0
- [19] 3GPP TS 08.51 BSC-BTS Interface General Aspects, v8.0.0
- [20] 3GPP TS 08.52 BSC-BTS Interface - Interface Principles, v8.0.0
- [21] 3GPP TS 08.54 BSC-BTS Layer 1; Structure of Physical Circuits, v8.0.0
- [22] 3GPP TS 08.56 BSC-BTS Layer 2; Specification, v8.0.0
- [23] 3GPP TS 08.58 BCS-BTS Interface Layer 3 Specification, v8.6.0
- [24] 3GPP TS 48.006 Signalling transport mechanism Specification for the Base Station System - Mobile Services Switching Centre (BSS - MSC) interface, v10.0.0
- [25] 3GPP TS 48.103 Base Station System - Media GateWay (BSS-MGW) interface; User plane transport mechanism, v10.0.0
- [26] 3GPP TS 29.414 Core network Nb data transport and transport signaling, v10.0.2
- [27] 3GPP TS 23.060 General Packet Radio Service (GPRS); Service description. v3.17.0
- [28] 3GPP TS 03.64 General Packet Radio Service (GPRS); Overall description of the GPRS radio interface; Stage 2, v8.12.0
- [29] 3GPP TS 08.14 General Packet Radio Service (GPRS); (BSS) - (SGSN) interface; Gb interface Layer 1, v8.0.0
- [30] 3GPP TS 08.16 General Packet Radio Service (GPRS); (BSS) - (SGSN) interface; Network service, v.8.0.0
- [31] 3GPP TS 48.016 General Packet Radio Service (GPRS); Base Station System (BSS) - Serving GPRS Support Node (SGSN) interface; Network service, v10.0.0
- [32] 3GPP TS 08.18 General Packet Radio Service (GPRS); (BSS) - (SGSN) BSS GPRS protocol (BSSGP), v8.12.0
- [33] 3GPP TS 04.60 General Packet Radio Service (GPRS); (MS) - (BSS) interface; Radio Link Control/Medium Access Control (RLC/MAC) protocol, v8.27.0
- [34] 3GPP TS 04.64 MS-SGSN; Logical Link Control (LLC) Layer Specification, v.8.7.0
- [35] 3GPP TS 04.65 MS-SGSN; Subnetwork Dependent Convergence Protocol (SNDCP), v8.2.0

3.3 3G

- [36] 3GPP TS 23.002 Network architecture, v10.4.0
- [37] 3GPP TS 25.401 UTRAN overall description (Release 10), v10.2.0
- [38] Kaaranen, Ahtiainen, Laitinen, Naghian, Niemi: UMTS Networks, Architecture, Mobility and Services, Second Edition, Wiley 2005
- [39] Holma, Toskala: WCDMA for UMTS, 2nd Edition. Wiley, 2002.
- [40] Holma, Toskala: WCDMA for UMTS: HSPA evolution and LTE, 5th Edition. Wiley, 2010.
- [41] 3GPP TS 25.410 UTRAN Iu interface: General aspects and principles, v10.2.0
- [42] 3GPP TS 25.411 UTRAN Iu interface layer 1, v10.1.0
- [43] 3GPP TS 25.412 UTRAN Iu interface signalling transport, v10.1.0.
- [44] 3GPP TS 25.413 UTRAN Iu interface Radio Access Network Application Part (RANAP) signalling, v10.3.0
- [45] 3GPP TS 25.414 UTRAN Iu interface data transport and transport signalling, v10.1.0
- [46] 3GPP TS 25.415 UTRAN Iu interface user plane protocols, v10.1.0
- [47] 3GPP TS 25.419 UTRAN Iu-BC interface: Service Area Broadcast Protocol (SABP), v10.2.0.
- [48] 3GPP TS 25.420 UTRAN Iur interface general aspects and principles, v10.1.0.
- [49] 3GPP TS 25.421 UTRAN Iur interface layer 1, v10.1.0
- [50] 3GPP TS 25.422 UTRAN Iur interface signalling transport, v10.1.0.
- [51] 3GPP TS 25.423 UTRAN Iur interface Radio Network Subsystem Application Part (RNSAP) signalling, v10.4.0.
- [52] 3GPP TS 25.424 UTRAN Iur interface data transport&transport signalling for Common Transport Channel data streams, v10.1.0
- [53] 3GPP TS 25.425 UTRAN Iur interface user plane protocols for Common Transport Channel data streams, v10.1.0
- [54] 3GPP TS 25.426 UTRAN Iur and Iub interface data transport & transport signalling for DCH data streams, v10.1.0.

- [55] 3GPP TS 25.427 UTRAN Iub/Iur interface user plane protocol for DCH data streams, v10.1.0
- [56] 3GPP TS 25.430 UTRAN Iub Interface: general aspects and principles, v10.1.0.
- [57] 3GPP TS 25.431 UTRAN Iub interface Layer 1, v10.1.0
- [58] 3GPP TS 25.432 UTRAN Iub interface: signalling transport, v10.1.0.
- [59] 3GPP TS 25.433 UTRAN Iub interface Node B Application Part (NBAP) signalling, v10.4.0
- [60] 3GPP TS 25.434 UTRAN Iub interface data transport and transport signalling for Common Transport Channel data streams, v10.1.0
- [61] 3GPP TS 25.435 UTRAN Iub interface user plane protocols for Common Transport Channel data streams, v10.3.0.
- [62] 3GPP TS 25.442 UTRAN implementation-specific O&M transport, v10.1.0
- [63] 3GPP TS 29.281 General Packet Radio System (GPRS) Tunnelling Protocol User Plane (GTPv1-U), v10.3.0
- [64] 3GPP TS 25.301 Radio Interface Protocol Architecture, v10.0.0
- [65] 3GPP TS 25.321 MAC Protocol Specification, v10.4.0
- [66] 3GPP TS 25.322 RLC Protocol Specification, v10.1.0
- [67] 3GPP TS 25.323 PDCP Protocol Specification, v10.1.0
- [68] 3GPP TS 25.324 BMC Protocol Specification, v10.0.0
- [69] 3GPP TS 25.331 RRC Protocol Specification, v10.5.0
- [70] 3GPP TS 29.281 General Packet Radio System (GPRS) Tunnelling Protocol User Plane (GTPv1-U), v10.3.0

3.4 LTE

- [71] 3GPP TS 23.401 General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access, v10.5.0
- [72] 3GPP TS 36.300 Evolved Universal Terrestrial Radio Access (E-UTRA), Overall description, v10.5.0
- [73] 3GPP TS 36.401 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Architecture description, v10.3.0
- [74] Holma, Toskala: LTE for UMTS, Evolution to LTE-Advanced. 2nd Edition. Wiley, 2011
- [75] 3GPP TS 36.410 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); S1 general aspects and principles, v10.2.0
- [76] 3GPP TS 36.411 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); S1 layer 1, v10.1.0
- [77] 3GPP TS 36.412 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); S1 signalling transport, v10.1.0
- [78] 3GPP TS 36.413 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); S1 Application Protocol (S1AP), v10.3.0
- [79] 3GPP TS 36.414 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); S1 data transport, v10.1.0
- [80] 3GPP TS 36.420 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 general aspects and principles, v10.2.0
- [81] 3GPP TS 36.421 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 layer 1, v10.0.0
- [82] 3GPP TS 36.422 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 signalling transport, v10.1.0
- [83] 3GPP TS 36.423 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 Application Protocol (X2AP), v10.3.0
- [84] 3GPP TS 36.424 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 data transport, v10.1.0
- [85] 3GPP TS 29.275 Proxy Mobile IPv6 (PMIPv6) based Mobility and Tunnelling protocols; Stage 3, v10.3.0
- [86] 3GPP TS 23.107 Quality of Service (QoS) concept and architecture, v10.1.0
- [87] 3GPP TS 23.203 Policy and charging control architecture, v10.4.0
- [88] 3GPP TS 29.061 Interworking between the Public Land Mobile Network (PLMN) supporting packet based services and Packet Data Networks (PDN), v10.4.0
- [89] 3GPP TS 23.272 Circuit Switched (CS) fallback in Evolved Packet System (EPS), v10.5.0
- [90] GSMA IR.92 IMS Profile for voice and SMS, v4.0
- [91] 3GPP TR 36.902 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Self-configuring and selfoptimizing network (SON) use cases and solutions, v9.3.1

OTHER

Further IP transport references (IETF RFCs etc) are listed at the end of Chapter 4.

4

Packet Networks

Esa Metsälä

It was stated in the very first pages that 3GPP view of the mobile backhaul is often a line. For a packet network, it could also be represented as a cloud. This does not provide much more insight into the topic. What is a packet based mobile backhaul then?

This question is addressed in this chapter: by looking at the services mobile backhaul offers (Section 4.1), standardization (Section 4.2), and the functionalities and protocols at different layers: Physical interfaces (Section 4.3), L2 with PPP (Section 4.4) and L2 with Ethernet (Section 4.5), IP (Section 4.6) and finally MPLS and IP applications (Section 4.7). In Chapter 5 the focus is on the physical layer/media (microwave radio, fibre, copper) and on MEF services. With these building blocks a mobile backhaul service is implemented.

4.1 Mobile Backhaul Application

4.1.1 Backhaul Service

Radio network (together with the core network) offers a service that is accessible for users: a mobile voice call, or internet connectivity for a smartphone, as examples. For a functioning mobile network, a transport service is needed.

Transport service to the radio network is modeled in Figure 4.1 as:

- *an integrated transport*, which is implemented within a mobile network node, and
- *a backhaul service*, providing a service between the peer mobile network elements via the external physical interfaces of the mobile network elements and via the intermediate backhaul network.

Both integrated transport, and the backhaul service are required for serving the mobile network application.

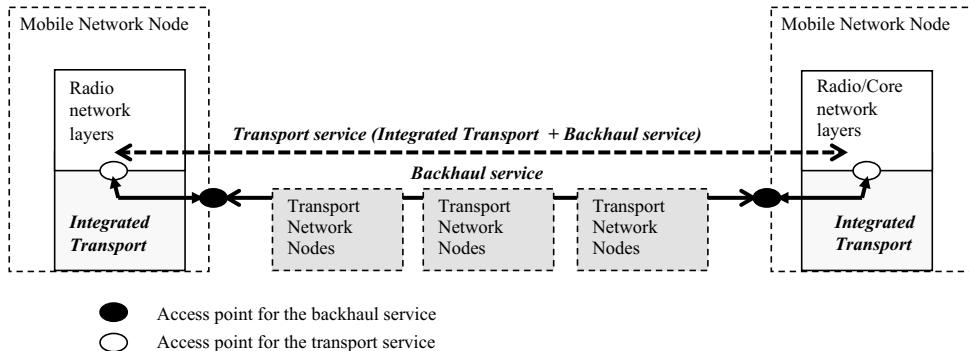


Figure 4.1 Transport service.

In Figure 4.1, a transport service consisting of the integrated transport, and the backhaul service, is offered for the radio network layer. The access to the transport service is only available in the mobile elements and not in the intermediate transport network nodes. This is e.g. UDP/IP connectivity for radio network layer Frame Protocol packets with the specified delay, loss, etc. (3G example). Frame protocol over UDP/IP is then terminated at the peer mobile network element, not within the transport network nodes.

At the external interface of a mobile network node (e.g. a NodeB), a backhaul service provides connectivity to the peer mobile network node (e.g. a RNC) at the IP layer. The access to the backhaul service is provided by the physical interface of the mobile network node. Intermediate transport nodes may operate at the IP layer or at L2 or L1, and the backhaul service may consist of multiple legs.

The backhaul service may be a physical layer service, L2 service, or IP layer service. What is required between the peer entities, is an IP layer connectivity. The backhaul transport service is provided by intermediate transport network nodes, either as self-deployed, or as a service of a third party service provider.

The main topic is the backhaul service. Anyway, a bigger part of characteristics of the transport service depend on the backhaul service than on the integrated transport. The integrated transport at a minimum is a termination function and a mapping of the radio network layer PDUs into UDP/IP or similar transport bearers.

The model allows the part of functionality that is implemented into the mobile network node to be discussed. This functionality is also essential. As an example, QoS mapping from the radio network layer attributes to the Differentiated Services code points takes place within the mobile network nodes – it would in fact not be possible in the intermediate transport nodes, since the radio network layer information is not available in the transport nodes.

Some of the aspects for the backhaul service are influenced by the radio network layer requirements. This depends on the radio access technology used (2G, 3G, LTE): every radio access technology has its own characteristics. Another part of the requirements originate from the end user service. Both are important.

In addition to user traffic, network control and management traffic is carried in the mobile backhaul. This includes radio network layer signaling, transport layer control protocols, O&M and potentially synchronization. These protocols are critical, as without a service for these functions, user plane services cannot be maintained either.

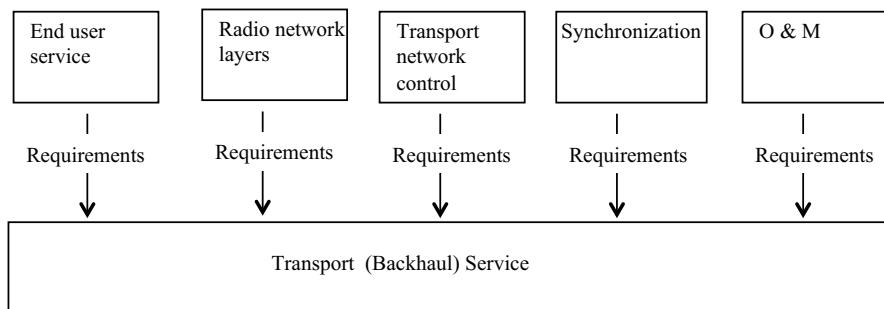


Figure 4.2 Transport service.

Figure 4.2 shows requirements for a transport service – originating from the end user service, radio network layers, transport network control, synchronization and O&M.

As discussed in Chapter 3, currently three generations of 3GPP mobile systems exist: 2G, 3G, and LTE. Among these systems there are various enhancements introduced along the way. Mobile packet backhaul supports all these systems, offering different types of services, such as:

- Pseudowire emulation service for 2G BTSs (which is natively TDM).
- Pseudowire emulation service for pre- Rel5 3G NodeBs (which is natively ATM).
- Service for native IP 3G NodeBs (Rel-5 onwards).
- Service for LTE IP eNodeBs (all-IP from the start).
- Control plane connectivity.
- Transport control protocols.
- O&M.
- Synchronization.

The basic service is simply that of information transfer: transfer of higher layer protocol data units between the service access points, timely and without impairments. The service as a whole can be abstracted with the help of a Service Level Agreement and Specifications, which help to describe the externally visible characteristics of the service (delay, packet loss, availability, security, etc.).

The service can be leased from a third party service provider. It can equally be supplied in-house by a transport network department. It is often a combination of the two. In both cases it is useful to model the backhaul as a service. It also allows separating the service from the underlying networking technology.

4.1.2 Access, Aggregation and Core

The single line model of transport between the mobile network elements can be replaced by dividing the backhaul into areas of access, aggregation and core. In Figure 4.3, access and aggregation tiers are shown since they are the most relevant for the mobile backhaul.

Networking technologies are often different in these three backhaul segments. In this text the focus is on access and aggregation. For access, it is critical to provide the ‘first mile’ physical connectivity to the BTS: support the traffic mix with the required characteristics. Cost is an issue, due to the amount of access lines needed.

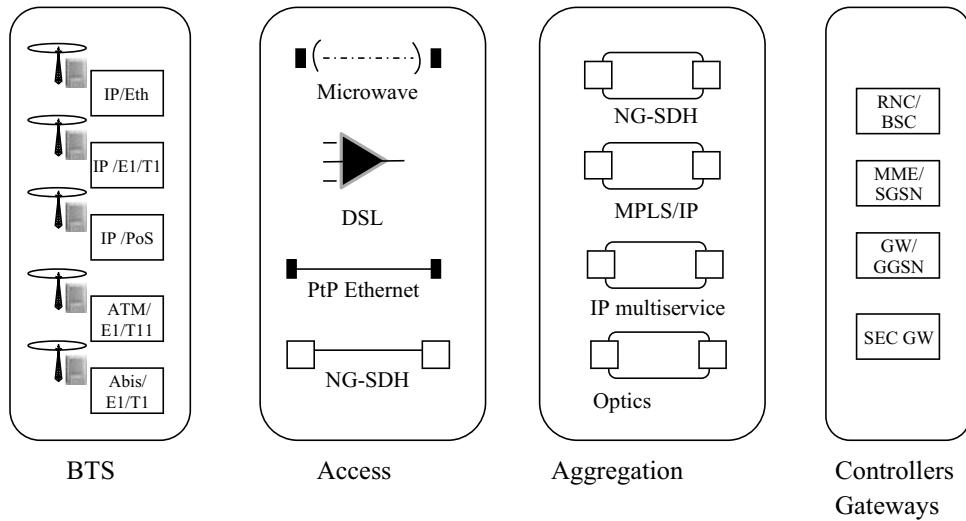


Figure 4.3 Access and aggregation tiers.

At the aggregation edge, one needs more functions: carrier grade resilience for high availability, QoS mappings, potentially further access control, and multiplexing and mapping of the access lines into the aggregation network.

4.1.3 3GPP Guidance for the Backhaul

3GPP has little guidance on how to build the backhaul: what technology to use and how to use it. The logical interfaces mandate the use of IP (IPv4 or IPv6). At the logical interface, the layer below IP can be Ethernet, PPP, ATM or something else.

In Figure 4.4, mobile elements may connect with any backhaul tier, access, aggregation or backbone. Mobile elements use a service of the backhaul. 3GPP specified protocols are carried transparently over the mobile backhaul. There are no protocol interactions between the radio network layer and the transport network layer in the intermediate backhaul network.

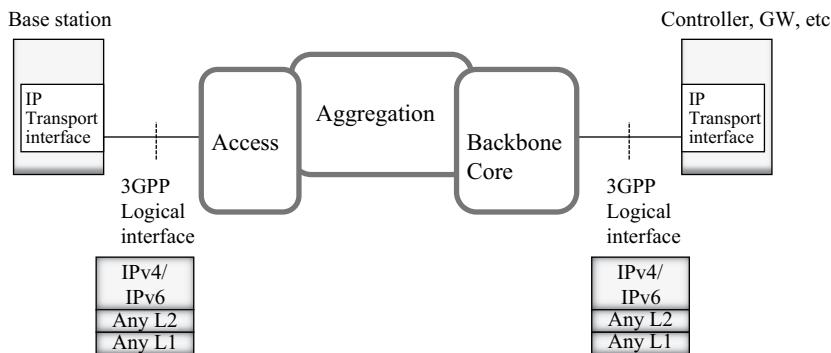


Figure 4.4 3GPP logical interfaces.

Above it was stated that mobile elements use a service of the backhaul. Would a Service Level Agreement (SLA) or other such type of specification with technical requirements for the backhaul be available directly from the 3GPP?

This would be very useful for the design of the backhaul, however, such specification does not exist. Many topics are implementation dependent: radio layer protocol options and timers used, algorithms, and so on. Due to the functionality being split to the mobile network elements, e.g. a delay budget needs to be allocated for every element and for every protocol layer. Also, the service offered by the mobile network to users introduces service-specific requirements to the backhaul. Mapping all these topics into a single generic document would be a difficult task.

4.1.4 Networking and Backhaul

Networking in general uses a model of local area and wide area networks. How do the general networking principles apply to the mobile backhaul?

In the case of an enterprise, a simplified model looks like that of Figure 4.5.

The model is relevant for the mobile backhaul as well: Technologies used in the packet mobile backhaul originate from the IT (Information Technology) and the enterprise world. Protocols are originally designed for this type of application.

In Figure 4.5, hosts connect to a Local Area Network (LAN), which could be a single office building, a floor in that building, or a campus area (with a limited physical distance). For communication between sites, Wide Area Network (WAN) links are needed. IP routers act as gateways to networks on other sites. Each WAN link is treated as a network of its own.

The same networking technologies are used in the mobile backhaul application. A similarity is that the amount of BTSs is rather high even though not fully comparable to the number of hosts. For cost optimization, it is crucial to have low cost ports at the hosts, and a low cost technology for any LAN.

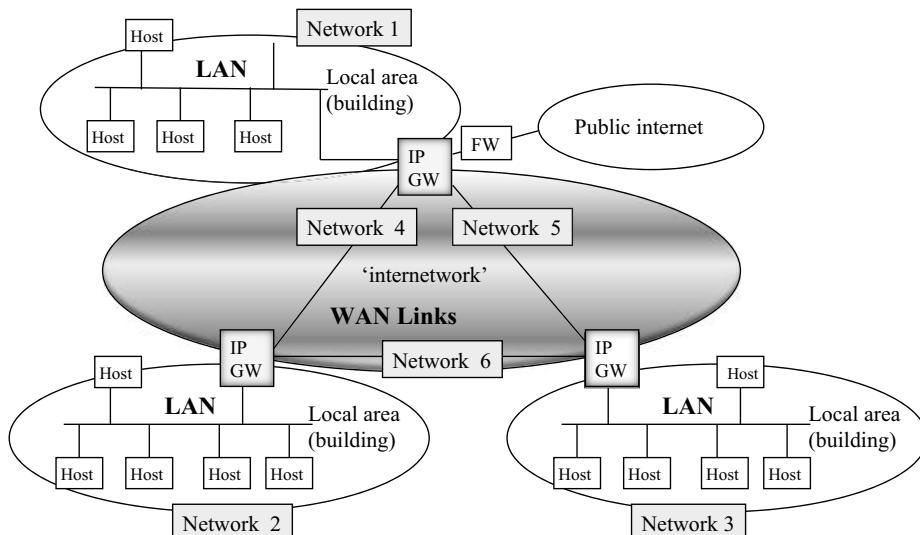


Figure 4.5 A simplified enterprise network. (For enterprise network architectures, see e.g. [1]).

Traffic flow follows a different topology in the mobile backhaul. In the enterprise LAN, not all traffic needs to exit the LAN. Hosts may contact other hosts or servers within the LAN, so direct connectivity via the LAN is a benefit. In the mobile backhaul, all traffic flows from the BTSs are directed towards remote sites. So the emphasis is on arranging site-to-site links in the mobile backhaul. Both applications, the enterprise network and the mobile backhaul, also require wide area connectivity that has high availability, is secure, and supports QoS for the services.

Often the mobile backhaul is not managed commonly with the radio network. Responsibility for the operation of the transport network is separate from the responsibility for the operation of the radio network. This has commonality with enterprise networks: individual sites and local area networks are managed by the enterprise itself. A service provider is often responsible for the wide-area network (WAN) connections. Naturally it is possible to deploy the whole network by the enterprise, or by the mobile operator (a self-deployed backhaul).

In the leased service model, BTSs and other mobile network elements are in the role of customer equipment (CE). At some point in the network, the mobile system interfaces the provider network, through a provider edge (PE) node. As an example, microwave radios may be used as a first mile access technology managed by the mobile operator. Subsequent legs of the transport service may be provided by a service provider.

4.2 Standardization

For the networking and transport technology, there is no single standardization body. Depending on the field, the related standardization forum varies. 3GPP mostly refers to existing standards concerning transport, developed by other organizations.

In the networking area, a lot of innovations have been driven by an implementation into a specific product. Later on, the feature or a variation of it, might have entered a standardization body, such as IEEE or IETF. It may also be that there is no direct counterpart for the functionality in standards, which leaves the feature proprietary. Despite this, many features of this kind have been widely deployed and have proven useful.

Often work is ongoing simultaneously in more than one standardization organization. This has the side effect that there will likely be a variation in the content of the standards, even if harmonization would be beneficial for the industry as a whole. At least terminology often differs and this alone causes confusion.

A brief introduction to standardization organizations is given in the subsequent section. Not all relevant standardization bodies are covered, in order to keep the focus on the most important activities within the field of a packet mobile backhaul. Also pre-packet era standardization is not covered.

4.2.1 IEEE

IEEE (Institute of Electrical and Electronics Engineers) is a professional association, whose aim is in advancing technological innovation. In addition to standards, IEEE publishes technical literature in electrical engineering, computer science and electronics. IEEE has a regional and a technical structure.

IEEE 802 LAN/MAN Standards Committee is the standardization body for the Ethernet LAN related standards, including categories like 802.1 Bridging & Management, 802.2

Logical Link Control and 802.3 Ethernet. Another example area relevant for the backhaul is synchronization. IEEE1588 precision timing protocol is one alternative for BTS synchronization in the packet network.

4.2.2 IETF

The Internet Engineering Task Force (IETF) is an open community of network designers, researchers, vendors and network operators. The objective of IETF is to make the Internet work better, with high quality technical documents. Technical work is done in working groups. The Internet Assigned Numbers Authority (IANA) coordinates unique parameter values for the Internet protocols.

IETF specifications that intend to become internet standards, are marked as standards-track documents. Related publications are RFCs (Request for Comments) and draft RFCs. Not all RFCs become internet standards, or Best Current Practices (BCPs). The RFCs that evolve to standards are additionally marked with ‘STDxxx’ while keeping the original RFC number as well. The RFCs that evolve into best current practices, are additionally marked with ‘BCPxxx’.

Best current practices document the best way to perform a certain operation or a process. Non-standard track specifications may be published as informational or experimental RFCs. Retired specifications may be published as historic RFCs.

As opposed to published specifications (RFCs), internet drafts do not have an official specification status. An Internet draft may be made available for informal comments in the internet-drafts directory. An unchanged internet draft is removed from the directory after 6 months, unless it is recommended to be published as an RFC.

For the standard track specifications, the maturity levels are Proposed Standard, Draft Standard, and Internet Standard. (Note: With RFC 6410, Oct 2011, maturity levels are reduced to two: ‘Proposed Standard’ and ‘Internet Standard’.) For the non-standard track specifications, maturity levels of Experimental, Informational, and Historic are used.

As RFCs are less formal and precise as standards from e.g. ITU-T, interoperability needs to be ensured by separate interoperability testing (IOT) of RFC compliant implementations from different vendors.

4.2.3 ISO

International Organization for Standardization (ISO) is a non-governmental organization, and a network of national standards institutes, publishing International Standards in diverse areas.

One of the most quoted ISO standard is the OSI (Open Systems Interconnection) model, defining OSI protocol layering. Another one which is widely used in service provider networks is the IS-IS routing protocol, the source of which is in ISO/IEC JTC 001 ‘Information technology’ committee, where subcommittee 6 is named ‘JTC 1/SC 6 - Telecommunications and information exchange between systems’.

4.2.4 ITU-T

ITU, International Telecommunications Union is a specialized agency of the United Nations. ITU-T, Telecommunications Standardization sector, covers a wide area of information and communications technology standards.

Some of the relevant ITU-T study groups for the mobile backhaul are:

- ITU-T Study Group 2 – Operational aspects of service provision and telecommunications management.
- ITU-T Study Group 12 – Performance, QoS and QoE.
- ITU-T Study Group 13 – Future networks including mobile and NGN.
- ITU-T Study Group 15 – Optical transport networks and access network infrastructures.
- ITU-T Study Group 16 – Multimedia coding, systems and applications.
- ITU-T Study Group 17 – Security.

ITU-T has modified its approval process in 2001, in order to respond to the need for faster standards development, with the Alternative Approval Process (AAP). However the traditional approval process (TAP) can also be used. Standardization Domain 04 (numbering/addressing) and Domain 11 (tariff/charging/accounting) are assumed to follow TAP, and other recommendations follow AAP, however this can be changed with a decision in the corresponding Study Group.

Examples of packet mobile backhaul related recent standards are e.g. in Operation, Administration and Maintenance (OAM), Optical networks, Synchronization, and Security.

4.2.5 MEF

Metro Ethernet Forum (MEF) has its purpose in developing technical specifications and implementation agreements to promote interoperability and deployment of Carrier Ethernet worldwide.

Some key documents related to the Mobile Backhaul are:

- MEF 2 Requirements and Framework for Ethernet Service Protection.
- MEF 3 Circuit Emulation Service Definitions, Framework and Requirements in Metro Ethernet Networks.
- MEF 4 Metro Ethernet Network Architecture Framework Part 1: Generic Framework.
- MEF 6.1 Metro Ethernet Services Definitions Phase 2.
- MEF 10.2 Ethernet Services Attributes Phase 2.
- MEF 11 User Network Interface (UNI) Requirements and Framework.
- MEF 12.1 Carrier Ethernet Network Architecture Framework Part 2: Ethernet Services Layer - Basic Elements.
- MEF 17 Service OAM Framework and Requirements.
- MEF 22 Mobile Backhaul Implementation Agreement (2/09).
- MEF 23 Class of Service Phase 1 Implementation Agreement.

For practical carrier Ethernet networks the most relevant areas for the mobile backhaul are the service definitions and MEF22. Ethernet service attributes document is also referred to in the IEEE 802.1Q standard.

4.2.6 IP/MPLS Forum

IP/MPLS Forum has its aim in driving broadband wireline solutions, converged packet networks and next generation IP network specifications. The specifications address interoperability, architecture, and management.

IP/MPLS Forum has its history as ATM Forum, Frame Relay Forum, and MPLS Forum. ATM Forum joined the MPLS and Frame relay alliance, which became the IP/MPLS Forum. Broadband forum's background is in DSL Forum. Broadband forum then united with the IP/MPLS Forum in 2009.

Technical Working Groups of the IP/MPLS Forum are:

- BroadbandHome.
- End to End Architecture.
- Fiber Access Network.
- IP/MPLS & Core.
- Metallic Transmission, and
- Operations & Network Management.

Forum has published IP/MPLS related specification concerning mobile backhaul, 'MPLS in Mobile Backhaul Networks Framework and Requirements'. A lot of work done on the broadband area is relevant for the mobile backhaul even though not directly referred to in 3GPP.

4.3 Physical Interfaces

4.3.1 High Data Rates

Theoretical air interface single user peak rates for selected radio technologies in downlink (DL) and uplink (UL) is shown in Figure 4.6 together with the Ethernet port capacities. This illustrates one of the key benefits of an Ethernet port, high capacity.

The above shows the capability of Ethernet, e.g. as a single port in an eNodeB to scale capacity-wise, and match the peak rates of the air interface. The other benefit of Ethernet is low cost. Due to the widespread high volume usage of Ethernet ports, the cost per bit in routers and switches is typically lower with Ethernet compared to other ports, such as Packet over Sonet, or IP over PPP over E1/T1s. Due to these reasons, Ethernet is often the first alternative for the physical interface port, when connecting IP based BTSs to the packet mobile backhaul.

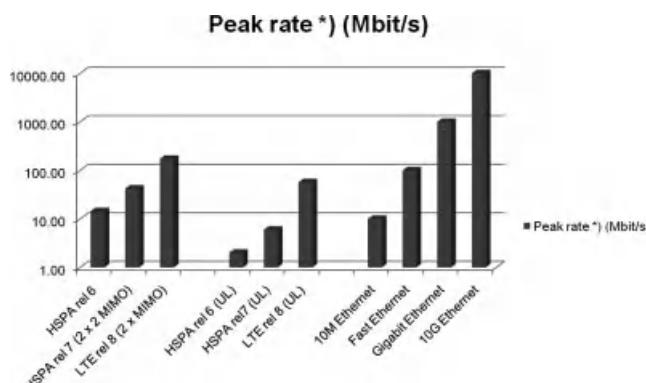


Figure 4.6 Peak rates of 3GPP air interfaces vs Ethernet. *)Peak rate of air interface is the theoretical maximum peak rate per single user. Note: HSPA + evolution not included.

4.3.2 Ethernet Ports

IEEE standards include data rates for 10 Mbit/s (10 BaseT), 100 Mbit/s (FastEthernet), 1 Gbit/s (Gigabit Ethernet) and 10 Gbit/s [12]. 40 Gbit/s and 100 Gbit/s standards have also been completed. With Ethernet, there is less need to install or upgrade physical ports, as an initial rate of e.g. 100 Mbit/s or 1 Gbit/s is adequate and accounts also for some growth. A single electrical Ethernet port may support speeds of e.g. 100/1000 Mbit/s.

For the physical connection, Ethernet supports both twisted-pair cable (unshielded twisted pair, UTP, and shielded twisted pair, STP) and optical fibre interfaces. With high speeds, there are more requirements for twisted-pair cabling.

With twisted-pair cables, one pair is used per direction in order to support full-duplex transmission. Similarly, two fibres are used for full-duplex transmission on fibre media. The twisted-pair cables may be provided e.g. with a RJ-45 connector, with 8-pins, allowing four twisted-pair cables. A minimum of two of these cable pairs are then used. More pairs may be needed, depending on the transmission rate.

When one twisted-pair cable is used for the transmit, and another pair for the receive direction, a cross-over cable is required to connect two devices, unless automatic detection feature is used. Automatic MDI/MDIX detection allows switching between receive and transmit directions.

Twisted-pair cable media is not suited for long-distance transmission, instead it supports site internal connectivity, e.g. between a BTS and a cell-site router, or a controller and a controller site device. The exact distance depends on the speed and on the cable type.

For the fibre interfaces, SFP (Small Form Factor Pluggable) transceiver modules are common, with LC type of fibre connector. Fibre may be multi-mode or single-mode. Laser devices support either short range (hundreds of meters) or long range (several kilometers, or tens of kilometers).

Ethernet is a common interface, which is available not only in networking equipment such as switches and routers, but also in different types of transmission and ‘telecom’ equipment. Ethernet port to the BTS can be offered over fibre, passive optical network, copper, SDH/Sonet including NG-SDH, or microwave radio, as shown in Figure 4.7.

Yet another viewpoint is the higher layer protocol carried over Ethernet. The Ethernet physical port and the Ethernet L2 frame is basically agnostic to the higher layer protocol (protocol type field depends on the upper layer protocol), so, for example, IPv4 or IPv6 can be carried.

4.3.3 E1/T1/JT1

A lot of telecommunications infrastructure exists that support time-division-multiplexed (TDM) interfaces, such as E1s, T1s and JT1s, and these interfaces can be used for native IP traffic as well. Compared to Ethernet, the transmission rates of E1s and T1s are limited, however. E1 interface is 2.048 Mbit/s and T1/JT1 is 1.544 Mbit/s, consisting of 32 and 24 timeslots, respectively.

Native Abis transmission is specified to carry the radio timeslots in the time-division-multiplexed PCM (pulse coded modulation) interfaces. Similarly, a lot of ATM based 3G NodeBs are connected via E1/T1 lines. So a lot of sites exist with already installed TDM lines.

Concerning capacity, a single E1 supports around 120 voice channels for 2G (16 kbit/s each, not considering signaling and O&M channels). This may be enough for a 2G site of voice-dominated

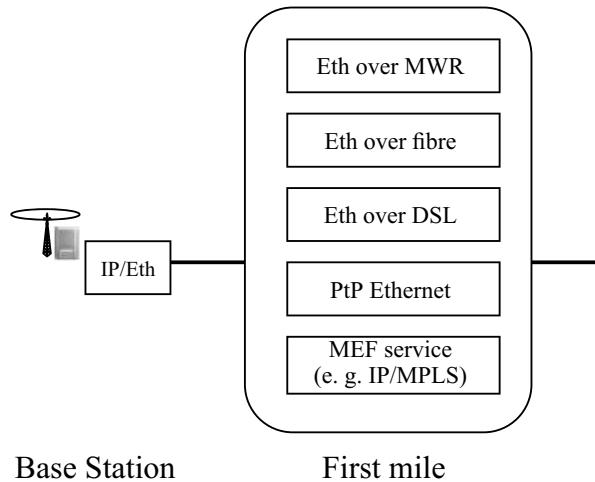


Figure 4.7 Ethernet on the first mile BTS access.

traffic mix. However it clearly limits the peak rate capability with HSPA and LTE to less than 2 Mbit/s. While E1s and T1s/JT1s can support the capacity of 2G for voice-only sites, they do not easily scale to the high data rates of HSPA and LTE and to a mobile broadband application.

Low interface capacity also means additional delay due to serialization. For example, when a 1500 octet packet is transmitted over an E1 (assuming 30 usable timeslots) this requires 50 E1 frames, or a delay of $50 \times 125 \mu\text{s}$ (6.25 ms). With smaller packets, the extra delay is still considerable. It also means that small voice packets need to wait until a larger data packet has been transmitted. (This, however, can be addressed by the ML-PPP protocol and Multi-Class Extension.)

By adding E1/T1 lines (using ML-PPP), the capacity grows, however in practice it is difficult to go much beyond 8..16 E1/T1s. Simply fitting the connectors and cables to the BTS becomes an issue. This is especially true for small footprint BTSSs.

Functionally narrowband TDM interfaces for the native IP traffic are feasible with the characteristics of low to modest capacity and notable delay. Economically it makes sense if the TDM infrastructure is available. With leased lines the costs often become prohibitive. This is market dependent.

The narrowband TDM interface of the BTS typically connect to a separate transmission device. This could be a Network Termination (NT) device for a leased line or a microwave radio indoor unit. The physical E1/T1 interface itself is not meant for long-haul transmission, although long-haul versions of the electrical interface exist also.

The physical interface of the E1 is either a 75 ohm coaxial cable (asymmetric) or 120 ohm twisted-pair (symmetrical), and for T1/JT1, the twisted-pair (symmetric) cable. The E1/T1 interface definitions can be found in ITU-T G.703/G.704 and ANSI T1.403/T.408.

A benefit of a E1/T1/JT1 physical interface is that it also carries synchronization. Synchronization is obtained via the line interface. The line interface clock must then be traceable to a primary reference. The whole TDM network is often obtaining its timing from a single source.

The E1/T1 framing support signaling alarms, loss of signal/loss of frame alignment, and a remote alarm/defect indication. Loss of signal (LOS) is detected, when the incoming signal has no transitions (meaning in practice, no signal) for a period of time. As a consequence, Alarm Indication Signal (AIS) can be sent. AIS is a signal with all ‘1’s. As a response to the received AIS, Remote Defect Indication (RDI, sometimes called Remote Alarm Indication, RAI) can be sent back.

Point-to-point protocol, and its’ extensions, are described further in a separate section.

4.3.4 SDH/Sonet

Synchronous Digital Hierarchy (SDH)/Synchronous Optical Networking (Sonet) provides interface rates from 51.84 Mbit/s to 9 953.280 Mbit/s as shown in Table 4.1.

Instead of mapping IP via PPP to the Sonet/SDH (Packet over Sonet, PoS), Next generation SDH (NG-SDH) has an added support for the Ethernet. In this case, the use of Sonet/SDH as an underlying layer is not visible to the customer and a standard Ethernet port is used. Often, this is a preferred option to PoS, since then there is no need to separately arrange for the PoS interfaces. NG-SDH will be explored in Chapter 5.

For Packet over Sonet applications within the mobile backhaul, the first two rates best match the BTS site capacities. Either the whole net interface capacity may be used for a single IP stream, or multiple E1/T1s may be multiplexed into the STS-3/STM-1 rate of 155.25 Mbit/s. In this case IP is first mapped into E1/T1s using ML-PPP.

A BTS may have SDH/STM-1/VC-4 (Virtual container) or Sonet/OC3/STS-3c SPE (Synchronous payload envelope) interface directly. The VC-4/STS-3c SPE bit rate is 150.336 Mbit/s, with a net rate of 149.760 Mbit/s, and IP can be carried in this using PPP protocol.

Alternatively, the capacity of STM-1/STS-3 may be used for multiplexing 63 E1s (using e.g. VC-12 Virtual containers for E1s), or 84 T1s (using VT1.5 virtual tributaries for T1s). In this case, IP is first mapped to E1/T1s and then further into the SDH/Sonet structures. With Multilink-PPP, a number of E1/T1s can be bundled. In this case, the BTS can interface the SDH/Sonet multiplexer using E1/T1 interfaces.

The electrical STM-0 and STM-1 interfaces use one coaxial cable (75 ohm) per direction.

The optical interfaces are categorized for intra-office short-haul (< 2 km), inter-office short-haul (15 km) and inter-office long-haul (40 km – 80 km). The distances are target distances of ITU-T G.957.

Depending on the application, the interface bit-rate and the distance, transmitters are either LED (Light-emitting diode) transmitters or single mode or multi mode lasers. Similarly, the fibre may be multi mode or single mode, with multimode capable only for short distances. The

Table 4.1 SDH/Sonet interface rates [18].

Interface	Physical layer rate
STS-1/STM-0	51.84 Mbit/s
STS-1/STM-1	155.52 Mbit/s
STS-1/STM-4	622.08 Mbit/s
STS-1/STM-16	2 488.320 Mbit/s
STS-1/STM-64	9 953.280 Mbit/s

Octet	
1	Flag
2	Address
3	Control
4	Protocol
5	Protocol
6	Information (incl. Padding)
7	...
8	Information (incl. Padding)
9	Frame control sequence
10	Frame control sequence
11	Frame control sequence
12	Frame control sequence

Note 1) Protocol field is 1 or 2 octets (2 octets shown)

Note 2) Frame control sequence fields is 2 or 4 octets (4 octets shown)

Figure 4.8 PPP frame format [24].

single mode has a nominal wavelength of 1310 or 1550 nm. Fibre variants are defined in ITU-T G.652/G.653/G.654/G.655.

Sonet/SDH has overhead (OH) structures, that are used for OAM and alarm indication at different levels. A transmission path consists of regenerator sections and multiplex sections, and the OAM can function per these segments. The functions include Loss of Signal (LOS) and Loss of Frame (LOF) detection, AIS and RDI, among other failure indications. For performance monitoring, bit error monitoring is supported at the regenerator and multiplex section levels, and at the path level.

4.4 PPP and ML-PPP

4.4.1 PPP over E1/T1/JT1

PPP (Point-to-Point Protocol) defines an encapsulation method assuming a full-duplex communication. The frame format is HDLC-like (High level data link control). From 3GPP standards viewpoint, PPP is one option for Layer-2, below the IP as a network layer protocol. PPP in HDLC-like framing is defined in RFC1662 [25].

The L2 frame format is shown in Figure 4.8.

A flag marks the start of the frame with a bit sequence ‘01111110’. Address field is typically a broadcast field of all ‘1’s, meaning all stations. Control fields consists of ‘00000011’, other values may be used if so agreed. Protocol indicates the protocol carried in the information field. Information consists of data to be transmitted, and padding may be included. Together the information and padding may include up to 1500 octets, which is the default value for the Maximum Receive Unit (MRU). Other values may be negotiated.

Frame Check Sequence (FCS), is calculated over the entire frame. 2 byte FCS is the default, 4 byte FCS may be negotiated. In case of mobile backhaul, IP is encapsulated into PPP by marking the protocol field accordingly, and then including data and optional padding.

In IP addressing, the PPP link can be treated as an unnumbered serial interface (no IP address assigned) or then as a network. In the latter case, the network consumes one address in addition to the interface addresses.

In the control plane, Link Control Protocol (LCP) may be utilized. Link control protocol functions include establishing, configuring and testing the connection, with the target of supporting ease of configuration with an auto-negotiation of parameters.

Authentication is also possible with the PPP protocol. For this purpose, Password Authentication Protocol (PAP) or Challenge Handshake Authentication Protocol (CHAP) can be used. In the dial-up Internet connectivity, the Internet service provider may authenticate the user with this method.

4.4.2 ML-PPP

In order to increase the data rate capability with the IP over PPP over E1/T1/JT1s, it is possible to combine a number of these lines to create a higher capacity service. This is accomplished through the use of multi-link point-to-point (ML-PPP) protocol. Aggregating for example eight E1 lines into a group results in a IP/ML-PPP over 8xE1 interface, with the physical layer rate of 8×2.048 Mbit/s or 16.384 Mbit/s. ML-PPP may be negotiated using the LCP.

Figure 4.9 shows an application of IP based NodeBs connected via E1/T1 based TDM network to the RNC.

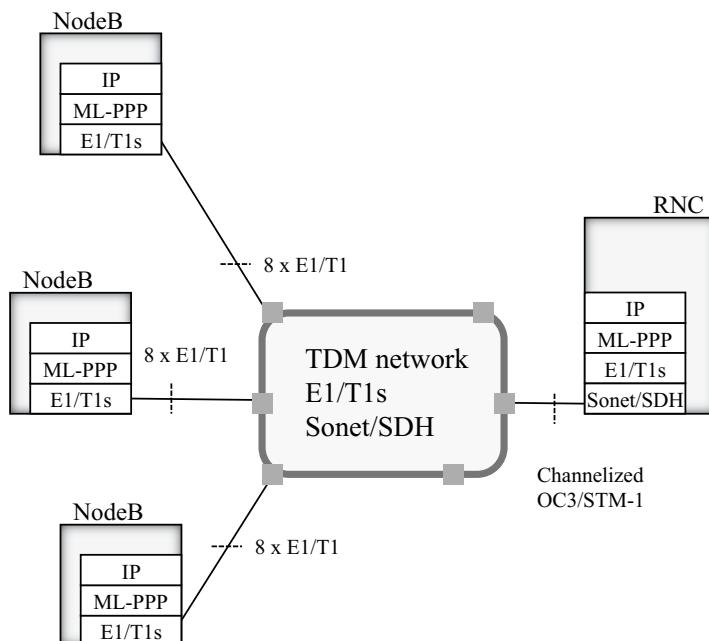


Figure 4.9 Using ML-PPP over E1/T1s for IP Iub.

Flag	Flag
Address	Address
Control	Control
Protocol	Protocol
Protocol	Protocol
Begin/End	Begin/End
Sequence	Sequence
Sequence	Sequence
Sequence	Sequence
Information (incl. Padding)	Information (incl. Padding)
...	...
Information (incl. Padding)	Information (incl. Padding)
Frame control sequence	Frame control sequence
Frame control sequence	Frame control sequence
Frame control sequence	Frame control sequence
Frame control sequence	Frame control sequence

Note 1) Frame control sequence fields is 2 or 4 octets
(4 octets shown)

Note 1) Frame control sequence fields is 2 or 4 octets
(4 octets shown)

Figure 4.10 ML-PPP header [28].

With ML-PPP, multiple E1/T1 lines are aggregated into a single higher capacity interface. In the example eight lines were used for each NodeB. In the TDM network E1/T1s are collected with a Sonet/SDH multiplexer, so that the traffic can interface the RNC via a few channelized STM-1/OC3 interfaces, instead of a large number of E1/T1 ports.

Another possibility is to have a ML-PPP gateway terminating the PPP protocol, and supporting an Ethernet interface towards the RNC. In this case the ML-PPP gateway provides the conversion from the E1/T1 infrastructure into packet networking.

In the mobile backhaul application in the BTS access area, ML-PPP interfaces provide capacity up to a few tens of Mbit/s, depending on the amount of TDM lines available. Instead of using E1/T1/JT1 lines all the way from the BTSs to the controllers, they can be terminated in an aggregation device. This device interfaces the higher aggregation layer, e.g. with an Ethernet interface.

With ML-PPP, large packets may be fragmented into multiple segments, which then are transmitted over the physical links constituting the bundle. Segment sizes can be optimized to match the speed of the links, in case links have unequal speeds.

For ML-PPP, the fragments are encapsulated as shown in the header of ML-PPP in Figure 4.10. The left hand side shows a long sequence field format (4 bytes). Short sequence format (2 bytes), depicted on the right hand side, may be negotiated using the LCP. A protocol ID of 003DH is reserved for the ML-PPP. The ML-PPP header consists of Begin/End bits and the sequence number field. Begin/End bits tell whether the fragment begins or ends the packet carried as data. Sequence field is incremented for transmitted fragments.

In order to reduce blocking for voice due to large data packets, optimization is possible with a Multi-Class Extension. Multi-Class Extension essentially supports running multiple instances of the ML-PPP protocol. Each class operates on its own instance, with separate

sequence numbering. The classes are identified with two unused bits in the short sequence header, and with four of the six unused bits of the long sequence header. This arrangement allows for four classes in the case of short sequence headers, and 16 classes in the case of the long sequence header. With the classes, voice can be separated from data, as an example.

IP Control Protocol (IPCP) for the PPP allows configuration of the IP layer protocol parameters. One of such options is IP header compression. Header compression method can be RFC1144 for TCP/IP, which is included in the original IP CP RFC. More recently, Robust Header Compression (ROHC) defines profiles for compressing IP, UDP, RTP, and ESP protocols.

Header compression saves bandwidth on the low-speed links, as headers can be substantially compressed. The main drawback is the added complexity and cost. If header compression is implemented for 3G as an example, each NodeB needs a counterpart which terminates the header compression function. Header compression is computing intensive and increases cost of the elements.

Another PPP supported means of increasing efficiency on low speed links is PPP multiplexing [30]. This allows multiplexing several PPP encapsulated packets within a single PPP frame. PPP multiplexing adds a delimiter to separate the PPP packets at the de-multiplexing end. A PPP Mux control protocol is used to negotiate this option. PPP multiplexing occurs before the ML-PPP encapsulation, so that PPP multiplexing is then used for the whole ML-PPP bundle.

4.4.3 PPP over Sonet/SDH

Similarly to the IP over PPP (or ML-PPP) over E1/T1s, IP can be mapped into PPP and further to the Sonet/SDH containers, as Sonet/SDH essentially is a point-to-point full-duplex link. Often the solution is called Packet over Sonet.

There are no major differences in Sonet and SDH concerning PPP operation. HDLC-like framing is used as discussed for PPP over E1/T1/JT1. IP over PPP is mapped into Sonet STS-3c-SPE or SDH VC-4, in the case of the STS-3/STM-1 interface. The payload is scrambled before inserting into the Sonet/SDH container, due to issues discovered with the original RFC1619 specification.

Note that for carrying IP in Sonet/SDH network, NG-SDH allows the use of an Ethernet port, and the Ethernet is then mapped to the Sonet/SDH containers, possibly with virtual concatenation.

4.5 Ethernet and Carrier Ethernet

Ethernet's success in the enterprise and LAN area has enabled low port costs and also low cost for basic Ethernet hardware for Ethernet bridging. Ethernet also has the general benefit in the enterprise LANs that due to the flat MAC address space, hosts can be moved within the LAN without a need to reconfigure IP addresses.

Also, Ethernet is known for its autoconfiguration capabilities within the LAN. Due to MAC address learning, there is no need to install routes to destinations or run a routing protocol to learn those routes. Instead, bridges forward unknown frames to all stations in the LAN and gradually automatically learn of the port behind which the MAC address resides.

Originally Ethernet is a LAN technology, approved as IEEE 802.3 in 1983, supporting 10 Mbit/s, and in general standardized in IEEE 802 standards. What is a bit confusing, is that Ethernet is at the same time a L1 (Physical layer) and a L2 (Link layer) technology.

Since the inception, Ethernet (especially at L1) has evolved and the current Ethernet is quite different from the initial Ethernet LAN standards. Today, Ethernet is typically not a shared media on the physical layer, but a point-to-point link built with a pair of copper or fibre cables. At the center site, a switch connects all stations together.

Basic MAC bridging and frame forwarding concepts are, however, mostly kept intact. Ethernet also is known for its backwards-compatibility: New versions of standards allow operation with older bridges.

4.5.1 Carrier Ethernet

The five attributes of carrier Ethernet in the Metro Ethernet Forum (MEF) are Standardized Services, Scalability, Reliability, Quality of Service and Service Management [41]. MEF view is service oriented; instead of standardizing technologies and protocols, MEF focuses on characterizing the services at Ethernet layer, their behaviour and their attributes. The actual standardization of functionalities is done in IEEE (native Ethernet and its evolution), IETF (MPLS and IP based implementation for Carrier Ethernet) and ITU-T (NG-SDH, Ethernet protection, and many OAM related functions).

One shortcoming of native Ethernet is in troubleshooting and OAM. Ethernet did not have a way of monitoring whether the connectivity at the Ethernet layer exists. There were no simple checks like ‘ping’ at the IP layer. Ethernet OAM addresses this need.

Similarly, the original Ethernet frame lacks QoS support. Since that, IEEE 802.1Q standard introduced priority bits for class of service marking. For scalability, in terms of service instances (VLANs), 802.1Q was limited to a 12-bit field. IEEE 802.1ad (Provider bridging) adds another 12-bit field, so that customer VLANs can be kept separate from provider VLANs.

With provider bridging, a single MAC address field is still used for both customer and provider addresses. To enhance the scalability related to MAC addresses, Provider backbone bridging (PBB) of IEEE802.1ah adds separate MAC address fields for the provider (and yet further fields).

With native Ethernet bridging, redundant switch topologies rely on spanning tree. The combination of L2 bridging and spanning tree in native Ethernet is a limitation for its use in carrier grade applications.

In IETF, carrier Ethernet is addressed by the use of MPLS and IP: MPLS forwarding in the user plane with an IP control plane. Additionally for E-LAN (multipoint) service, the topology in the MPLS core has to be full-mesh, with split-horizon forwarding. With these restrictions, IETF based carrier Ethernet does not need spanning tree in the MPLS core.

As mentioned, MEF services may be delivered basically on any technology. For the mobile backhaul, the mobile operator view is that of a service user. In this case the UNI (User to Network Interface) interface and the SLA (Service Level Agreement) matter more than the underlying technology. For the service provider the implementation of the service is essential. The mobile operator does not in general need to see beyond the UNI. Knowing how the service might have been implemented gives however some insight to the service and to its characteristics. For troubleshooting it may also be essential.

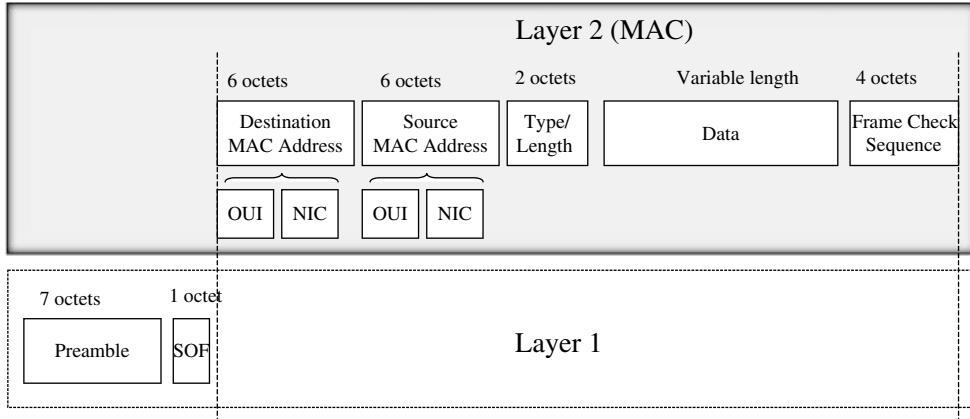


Figure 4.11 Ethernet frame [12].

4.5.2 Ethernet and Ethernet Bridging

Ethernet frame format is shown in Figure 4.11.

At the Ethernet L2 (Ethernet MAC layer), the header consists of destination and source MAC addresses, Ethernet type (Ethernet II frame, see¹) and a checksum (Frame check sequence, FCS). At Layer 1, a 7-octet preamble and 1-octet Start of frame delimiter are included.

The MAC address consists of an organizationally unique identifier (OUI) and a unique network interface card (NIC) field. Both of these fields are three octets long and together the L2 MAC address sums up to six octets. MAC addressing supports unicast, multicast and broadcast addresses. Broadcast address consists of all ones (FFFF FFFF FFFFH).

MAC addresses have no hierarchy. Certain MAC addresses are reserved and have a special purpose. MAC addresses have to be unique within the L2 domain.

Ethernet type field is included in the Ethernet II frame. This indicates the protocol that is carried by the Ethernet frame. As an example, 8000H is used for IPv4, 86DDH for IPv6, 0806H for ARP, and 8100H for a VLAN-tagged frame. The header includes a 4-octet checksum, which is calculated over the header fields and data (destination address, source address, length/type, data and padding). Erroneous frames are discarded.

An Ethernet bridge works on the MAC layer, and is transparent to the network level protocol. In the case of mobile backhaul, the network level protocol is IP. Underneath the L2, different Ethernet physical layers may be used. The Ethernet bridge can forward traffic between different Ethernet physical segments, as shown in Figure 4.12.

Ethernet bridge is the term used in the IEEE standard. Often Ethernet switch is used interchangeably. A multilayer L2/L3 switch is capable for both L2 bridging, as well as for IP forwarding. Bridges and switches both limit collision domain to a port.

¹ According to IEEE 802.3, Ethernet Type/Length field is interpreted as Length in case the value is less than or equal to 1500. If the value is greater than or equal to 1536 (0600H) then the field is interpreted as Type. In practice, Type interpretation is commonly used.

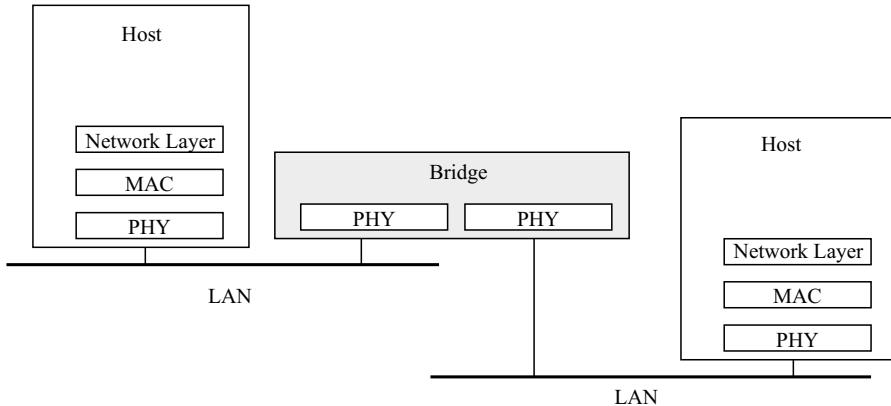


Figure 4.12 Ethernet bridging [43].

The main operations of Ethernet bridging are:

- forwarding of the L2 frames, based on filtering rules (e.g. MAC address table), and
- creating the filtering rules (by MAC address learning and/or other methods).

A bridge that supports MAC address learning is a learning bridge.

If there is no filtering rule, a bridge forwards received frames out on all ports (except on the port where it came from). This is called flooding. A filtering rule exists due to a management plane action (manual configuration), or it is created by MAC address learning. Filtering occurs also if a port is blocked (e.g. due to spanning tree protocol). In addition, hosts may also indicate the destination addresses they wish to receive with multicast registration protocols.

Bridges learn MAC addresses by observing the source MAC addresses of the incoming frames, and associating these with the ports where the frames came from. A MAC address table consists of MAC addresses and of the ports associated with those addresses.

Due to the learning capability and due to the flooding function, there is no necessity for configuring the MAC address table in advance. If the MAC address table is empty, the frame is flooded. So a L2 frame finds its destination, provided that this destination is within the flooding domain. When the MAC address table has an entry for the destination MAC address, the frame will only be forwarded to the port corresponding to the MAC address.

A drawback related to unknown unicast flooding is that a frame is forwarded also to links where there are no recipients for that frame. Capacity on the links is consumed. Flooding unknown unicast frames, or broadcast frames, may lead to a broadcast storm, if the frames find a path back to the originating switch. In this case the whole broadcast domain of the network suffers from low or no throughput. Spanning tree is designed to address this risk, by blocking ports so that loops do not exist.

Troubleshooting the topology and configuration during a broadcast storm is difficult.

Due to this, configurations and failure situations that may lead to L2 loops have to be analyzed. Keeping broadcast domains small in size helps. Spanning tree is discussed in Part II in Resilience chapter.

MAC address table in the Ethernet switch can hold a certain maximum amount of entries. The amount of MAC addresses may thus limit the size of the network. MAC address table entries also age out. If the size of the MAC address table is small, unused MAC addresses

should be removed in order to make room for the new addresses. On the other hand, removing entries rapidly leads to re-learning of the ‘old’ address. These topics are dependent on the implementation of the switch.

In some cases it is useful to remove an entry from the MAC address table prior to the expiration of the timer. When the switch knows that the device on the other end of the link is not operational, the related entry can be removed. As an example, after protection switching in the network, destination MAC address may reside behind a different port. Removal of the initial entry allows faster restoration of the L2 service.

4.5.3 Ethernet Link Aggregation

Link aggregation, defined in IEEE 802.1AX-2008 [44], supports combining multiple Ethernet links into a group, which is seen by the Ethernet MAC client as a single link. The benefits are increased capacity and also resilience, as failure of a single link is tolerated. Link aggregation also allows load sharing which is otherwise not supported on the Ethernet. (With the exception of VLAN based load sharing when using Multiple Spanning Tree Protocol).

Link aggregation adds a sublayer to the Ethernet L2, between the MAC layer and the MAC client. MAC client communicates with an aggregator function which hides the individual ports. The ports are bound to the aggregator, which is responsible for distributing and collecting frames. Link aggregation control protocol (LACP) supports controlling and configuring the operation, such as binding ports to the aggregator.

Link aggregation group (LAG) is considered operational, if a port belonging to that group is up. LAG becomes unoperational, when all ports are down. A port may be detected to be down by multiple methods. It may be physically down (‘no signal’). A LACP failure may be detected by not receiving LACPDUs or by other means.

Links need to operate at the same rate and they need to be full-duplex. A distribution algorithm allocates traffic to the individual links. The algorithm can be based solely on source and destination MAC addresses, but this may not provide adequate information for load distribution. The algorithm may also use higher layer information; typically IP addresses and port numbers.

The MAC address of the Link aggregation group may be the same as the MAC address of one of the ports.

4.5.4 VLANs

Virtual LAN (VLAN) support was added into Ethernet with the IEEE802.1Q standard. A 2 byte (16 bit) tag is added to the Ethernet frame, to identify the VLAN (VLAN ID), and also to include a Class of Service marking (Priority code point bits, PCP). Adding a VLAN tag increases the Ethernet frame header by 2 bytes, which needs to be taken into account when calculating Maximum Transfer Unit (MTU) at the Ethernet layer.

The new fields due to the 802.1Q standard are shown in Figure 4.13. The EtherType allocated for 802.1Q is 8100H.

Priority code point (PCP) allows eight values for priorities. Canonical format indicator (CFI) is used for compatibility with Token Ring bridges. VLAN identifier (VLAN ID) is 12 bits and so allows for 4096 values altogether.



Figure 4.13 IEEE 802.1Q header fields [45].

In the standard VLAN ID 0 is reserved for priority-tagged frames: VLAN ID is 0, but priority is encoded into the priority bits. VLAN ID 1 is the default for port VLAN ID (PVID). Value of FFFH (4096 decimal) is reserved.

Without VLANs each Ethernet port is a broadcast domain. With VLANs, each VLAN is a broadcast domain of its own, so unknown Ethernet frames are not flooded into other VLANs.

Virtual LANs are used to separate traffic logically – e.g. based on traffic type (voice, data) or based on functionalities, such as departments of an enterprise (manufacturing, accounting, etc). If traffic is destined for a VLAN other than where it originated a router is needed.

4.5.5 Class of Service

Originally Ethernet did not include support for differentiating traffic and marking the frame according to the quality of service needs of the traffic carried in the Ethernet frame. The VLAN 802.1Q frame includes three bits (Priority Code Points, PCPs) for indicating Class of Service. Priority code points allow the Ethernet switch to be QoS aware by taking into account the marking in scheduling.

The two highest PCP values 6 and 7 are commonly reserved for network control traffic. The highest user traffic priority is then 5, which can be used e.g. for voice. Use of priority bits is further discussed in the QoS chapter.

4.5.6 VLAN Example

An example application is shown in Figure 4.14 for IP Iub nodeBs (3G). Microwave radio link transmission is built for the NodeB access. An ethernet switch (named as MWR L2 switch in Figure 4.14) is used as the first (pre-) aggregation device, combining traffic from a few NodeBs.

In Figure 4.14 Virtual LANs (VLANs) logically separate O&M from the user/control traffic into a different VLAN, as one possible application. There are dedicated VLANs per NodeB, since there is no need to connect directly two NodeBs. This increases isolation and separates broadcast domains. A downside is the extra configuration effort.

An alternative is to share a single VLAN for a number of NodeBs. This reduces the configuration effort, however as a result the broadcast domain extends over a number of sites. Having several stations (NodeBs) in the same broadcast domain causes a risk of losing connectivity to all of the sites in an error case where due to some reason a L2 broadcast storm results.

With the exception of a direct X2 implementation of the LTE X2 logical interface there is no need for a BTS to BTS connectivity. X2 is often implemented not as a direct

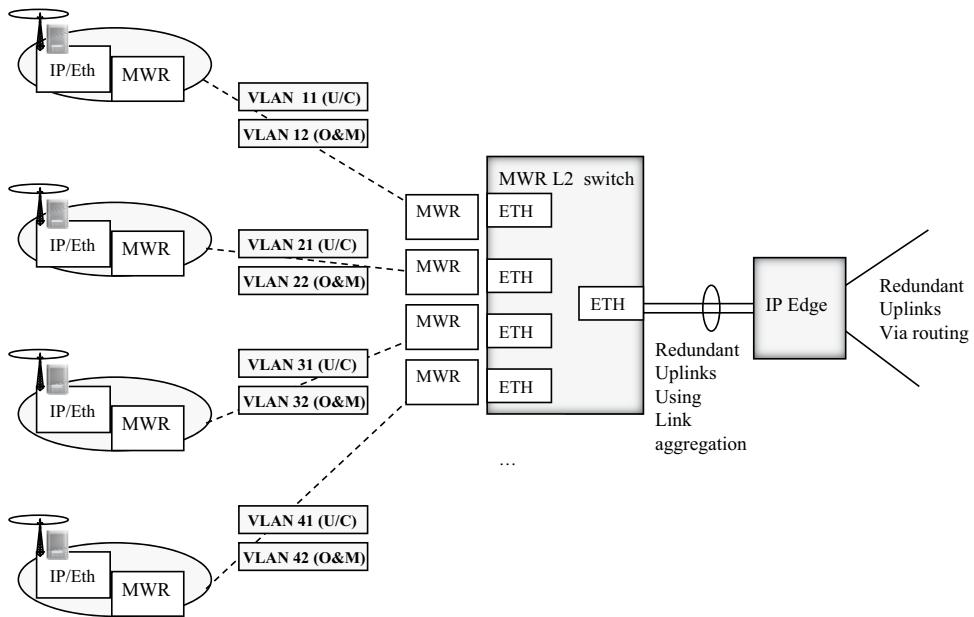


Figure 4.14 Ethernet aggregation.

eNodeB-eNodeB link, but rather via a point higher in the aggregation network. If X2 connectivity is wished to be arranged directly at the Ethernet layer, then LTE eNodeBs need to share the same VLAN so that eNodeBs can communicate directly via a L2 switch.

A microwave radio access L2 switch collects traffic from the NodeBs. The switch has a redundant uplink connection to the IP edge device. Ethernet link aggregation is used for redundancy for the switch uplinks. From the IP edge device onwards, two or more paths exist via IP and possibly MPLS. Alternatively, IP capability can be brought already to the pre-aggregation layer.

4.5.7 Ethernet OAM

How are faults on Ethernet links detected? How is it assured that connectivity on a VLAN over multiple switches exists?

Operation, administration and maintenance (OAM) with Ethernet is addressed in both IEEE and ITU-T standardization. Link-layer OAM targets OAM over a single hop: Monitoring that the link to the next Ethernet device is working. Service level OAM works on VLAN level end-to-end – over multiple network segments and even over different operator domains. Service level performance monitoring complements the OAM solution.

Figure 4.15 shows Link level OAM, defined in IEEE802.3 ah. The Link OAM aims for

- Remote Failure Indication.
- Remote Loopback.
- Link Monitoring.

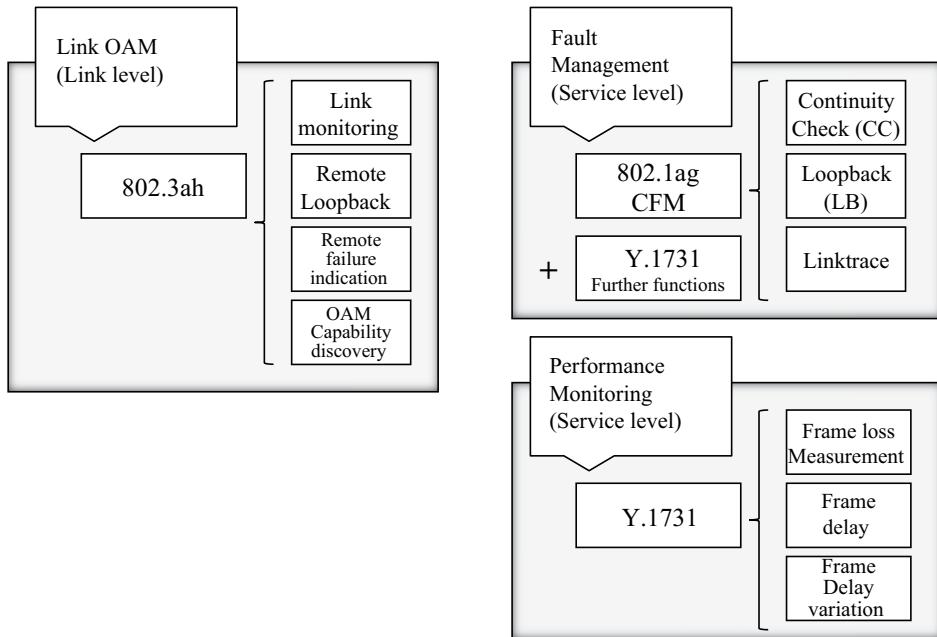


Figure 4.15 Key protocols for OAM

Critical link events are a Link fault, a Dying gasp (an unrecoverable failure), and a Critical event.

Link Layer OAM adds a new sublayer that is located between the MAC layer and the MAC client. If Ethernet link aggregation is in use, Link layer OAM is below the Ethernet link aggregator function.

Link OAM is sent in OAM Protocol data units (OAM PDUs) which are carried as Slow protocol mode frames (Annex 43B). OAM PDUs are carried only over a single link hop. They are not forwarded by the MAC layer.

OAM PDUs use a slow protocols_multicast address (01-80-C2-00-00-02) as the destination MAC address. The source MAC address is the individual MAC address of the bridge port that initiates the frame. At least one OAMPDU is sent per second.

For the service management, both IEEE and ITU-T have contributed to the standards. IEEE 802.1ag and ITU-T Y.1731 are largely compatible with different terminology, but also some variation in the content.

Service OAM targets end-to-end scenarios. With Maintenance Domains (MDs), service OAM capabilities in this aspect are comparable to Sonet/SDH transport networks. Up to eight levels of Maintenance Domains are allowed. MDs support a nested structure. Maintenance domain is controlled by a single network operator. Maintenance points (MPs) are either Maintenance Association End Points (MEPs) or Maintenance Domain Intermediate Points (MIPs).

Connectivity Fault Management aims at isolating connectivity faults to a single bridge or LAN. CFM entities are addressed by MAC addresses. MPs recognize a group address (CCM and LTM PDUs), and in addition an individual MAC address.

Functions supported are:

- Path discovery. Path discovery uses Linktrace protocol, which can track the path to the specific destination MAC address.
- Fault detection using Connectivity Check Messages (CCMs). CCMs are periodically transmitted as multicast. Continuity check is performed once per second, but other values may be configured. Transmit intervals as short as 3.3 ms are included in the specification.
- Fault verification, using the Loopback protocol. Loopback is essentially an ‘Ethernet Ping’.
- Fault notification, using the MEP.

Additionally fault recovery is mentioned in the IEEE specification using spanning tree protocol.

ITU-T Y.1731 defines additional functionality. Alarm Indication Signal (ETH-AIS) and Remote Defect Indication (ETH-RDI) are specified. Other Y.1731 functions include Ethernet Automatic Protection Switching (ETH-APS), Ethernet Test Signal (ETH-Test), Ethernet Locked Signal (ETH-LCK), Maintenance Communication Channel (ETH-MCC), Experimental OAM (ETH-EXP), and Vendor specific OAM (ETH-VSP).

ETH-APS is specified separately in ITU-T G.8031. ETH-Test may be used for in-service or out-of-service diagnostics. ETH-LCK indicates that a MEP is administratively locked, and helps the receiver to differentiate between administrative action and a failure condition. ETH-MCC is a communication channel between MEPs. ETH-EXP and ETH-VSP are not defined in the Y.1731.

For performance monitoring, a frame loss measurement (ETH-LM) and a frame delay measurement (ETH-DM) are included in ITU-T Y.1731.

4.5.8 *Provider Bridging*

Provider Bridging (PB) and Provider Backbone Bridging (PBB) enhance native Ethernet with further scalability and with a separation of service provider’s Ethernet network from the customer’s VLANs. With PB and PBB this is implemented with an enhanced Ethernet.

Customer network may run spanning tree, OAM, Link Aggregation Control Protocol (LACP) or other control protocols. These control protocols are either carried transparently through the service provider network, are blocked, or, the customer protocol instance has a peer entity in the service provider network.

Provider Bridging, IEEE802.1ad increases Ethernet scalability in terms of customers, and customer VLANs. With original 802.1Q, up to 4096 VLANs were defined. Provider Bridging adds another VLAN tag for service provider’s use. Due to this, 802.1ad is also called ‘QinQ’. Now the customer VLANs are transparent to the service provider’s network. See Figure 4.16.

The new QinQ frame structure with the S-tag basically repeats the 802.1Q structure (as in the C-Tag). A new Ethertype 88A8H marks the QinQ frame. PCP is the priority code point and VLAN ID is as with 802.1Q. DEI means a drop eligible indicator. CTAq (Customer Tag) has the original structure as with 802.1Q.

In fixed broadband, subscribers may e.g. be mapped into a C-VLAN/S-VLAN pair. Another option is to map all subscribers into one S-VLAN. In the mobile backhaul, the provider bridge

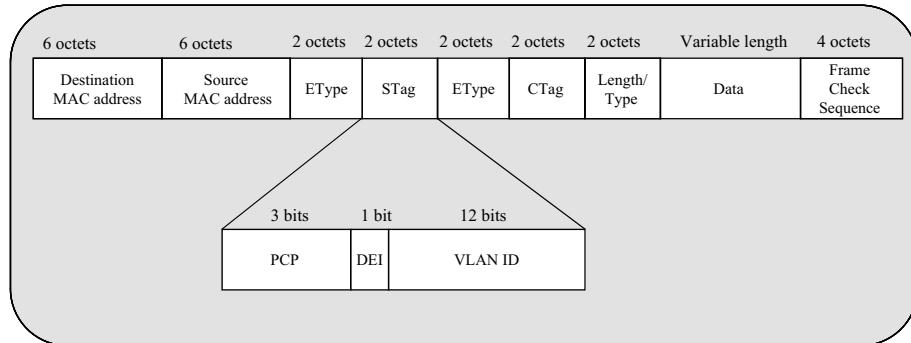


Figure 4.16 QinQ frame [49].

could e.g. map all BTS signaling VLANs into one S-VLAN, all BTS user plane VLANs into another S-VLAN and so on.

4.5.9 Provider Backbone Bridging

Provider backbone bridges (PBB) is defined in IEEE 802.1ah [46], and it further enhances Ethernet capabilities and scalability with a new Ethernet header including provider MAC source and destination addresses that are separate from the customer MAC source and destination addresses. Due to this, IEEE 802.1ah is also called ‘mac in mac’. Additionally a new 24-bit customer VLAN tag is included. For PBB Ethertype of 88E7H is allocated.

The main benefit of PBB is that customer MAC addresses learning can be kept at the edge devices. Core bridges do not need to know of any customer MAC addresses due to the separate MAC address field of the provider.

4.5.10 MPLS Based Carrier Ethernet

Since MPLS/IP is commonly used in service provider networks it is often the technology of choice when implementing MEF services. Point-to-point (E-Line) service is defined in IETF as VPWS (virtual private wire service). Multipoint service (E-LAN) is standardized as VPLS (Virtual private LAN service). Rooted multipoint (E-Tree) can be seen as a variant of the E-LAN, where connectivity between the leaves is restricted. While VPWS and VPLS have reached RFC status, rooted multipoint work is ongoing.

4.6 IP and Transport Layer Protocols

For the terrestrial transport, both 3G and LTE specifications define an IP based protocol stack for the logical interfaces; Iub, Iur, Iu-cs, Iu-ps for the 3G, and S1 and X2 for the LTE. For 2G, A and Gb have been standardized also for IP transport as discussed in Chapter 3. 2G Abis interface does not have a standardized IP alternative.

Base stations and other radio network elements (controllers) are basically hosts with an IP stack.² RNC, even if it has all logical interfaces based on IP, is not a router that routes IP packets between its ports. Instead, it terminates the radio network layer and the transport protocol stacks. For another logical interface a new IP packet is created (after the radio network processing).

UDP is used as the upper layer protocol in the user plane, SCTP in the control plane. The intermediate transport network may use Ethernet in the access, IP or MPLS for the aggregation, or a combination of those. At IP, routes may be learned by a routing protocol, or static routes may be configured.

Details of how mobile elements obtain their IP addresses and what is the addressing structure, is not defined in 3GPP. Networking and transport standards are not in the scope of 3GPP.

Since the mobile network elements must be able to interface standard compliant packet networks, it can be assumed that the mobile network elements need to comply with the relevant IETF, IEEE and other standards bodies' definitions, even when these are not always referred to by 3GPP.

4.6.1 IP

The Internet Protocol, IP, is defined in RFC 791, Internet Protocol (IP), for IPv4 [55]. IPv6 is defined in RFC 2460 Internet Protocol, Version 6 (IPv6) Specification [56]. Both protocols, IPv4 and IPv6, are relevant for the mobile backhaul and included in 3GPP standards. IPv4 has a focus in this book. Throughout the text IP refers to IPv4, unless specifically mentioned.

In addition to the IP layer present in the mobile backhaul, end user applications use IP. The user IP layer is transparent to the mobile backhaul as it is encapsulated within the radio network layer protocols. User IP layer can be IPv6 even though the mobile backhaul is using IPv4. Effectively these two layers (end user IP and mobile backhaul layer IP) are isolated.

For the IP protocol and the TCP/IP protocol suite, further information can be found e.g. in [61], [62], [63]. A short summary of main functions related to the IP protocol follows.

IP is connectionless and forwards packets hop-by-hop. There is no need to inform in advance or set up any connection between the sending and receiving nodes. IP packets may be lost or duplicated on the way. Receiving node does not acknowledge or reorder received packets. Higher layers need to deal with this.

IPv4 supports unicast, multicast and broadcast forwarding. With mobile backhaul, the most common traffic types are directed from a single node to another node and are point-to-point in nature, so unicast forwarding is used. For MBMS (Multimedia Broadcast Multicast Service) multicast would be of potential benefit.

IP was first introduced into the mobile backhaul access transport in 3GPP Rel-5 for UTRAN as an alternative to ATM. For Rel-5, the 3GPP TR25.933 in 2003 [64] documented motivation for the introduction of IP into UTRAN as:

- support of a mix of traffic types, and support for low-speed links;
- popularity of World Wide Web;

² In the core network, GGSN may be considered as a router, with an added mobile network specific functionality. Basically any other mobile network element may as well be built on a router platform and may have routing capability.

- price pressure on networking equipment;
- most applications will be based on IP;
- harmonization with operation and maintenance networks, that will be based on IP;
- packet switching allows efficient use of transport resources;
- independence of L2 technology;
- autoconfiguration and dynamic routing capabilities.

Use of narrowband E1/T1 links is in many cases economically not feasible for a mobile broadband application. This is due to the need of high user data rates, packet dominated traffic mixes, and cost pressure. The introduction of IP is in these cases often coupled to the availability of an Ethernet port to the BTS site.

With Ethernet, IP over 100M/1G/10G Ethernet ports deliver high capacity with a single low cost physical port. In the aggregation tier e.g. carrier Ethernet (IP/MPLS) is used for the aggregation of the access traffic from the BTSs. IP(MPLS) network can also be shared with other applications such as residential fixed broadband.

In Chapter 3 the protocol stacks specified by 3GPP for the different radio access technologies and their logical interfaces were reviewed. While there is variation in how the specifications are formulated, both IPv4 and IPv6 have been included into the 3GPP specifications, while in general the layers below IP (L2 and the physical layer) are not defined. For 2G and 3G also non-IP interfaces are defined. For these interfaces, there is no requirement by 3GPP for conversion to IP transport. Part of the base stations may stay with TDM (or ATM) -based logical interfaces while others already become IP based.

Similarly it is possible that a 3G NodeB uses ATM Iub until the RNC, but that the RNC interfaces the core network (Iu-cs, Iu-ps) with IP transport. Mobile network element (base station and controller) interface and protocol availability, such as the amount and type of transport interfaces and protocols, is clearly implementation specific, which of course may limit configuration options. With LTE the situation is simpler, as LTE is an ‘all-IP’ network from the start, 3GPP Rel-8. All logical interfaces are only defined as IP based.

Additionally, IPsec is in several cases mandated by 3GPP. A discussion on IPsec and backhaul protection follows in Chapter 9.

Ethernet is often used as an access technology. A question is the location of the IP edge: how close is the first IP device to the BTS. Another question is that should the same IP network also carry 2G Abis and possibly 3G ATM/Iub. These topics are not covered by 3GPP and are dependent of the network implementation.

An IPv4 header is shown in Figure 4.17.

The header fields are:

- Version describes the version of the protocol, IP v4 or IPv6. Value 4 is used for IPv4, and 6 for IPv6.
- Internet Header Length is the length of the header. The header length varies depending on the amount of optional header fields. As a minimum the header consists of 20 bytes.
- DS/ECN. The field consists of a 6-bit field for Differentiated Services (DS) and a 2-bit field for Explicit Congestion Notification.
- Total length indicates the total length of the packet, including the header and data fields.

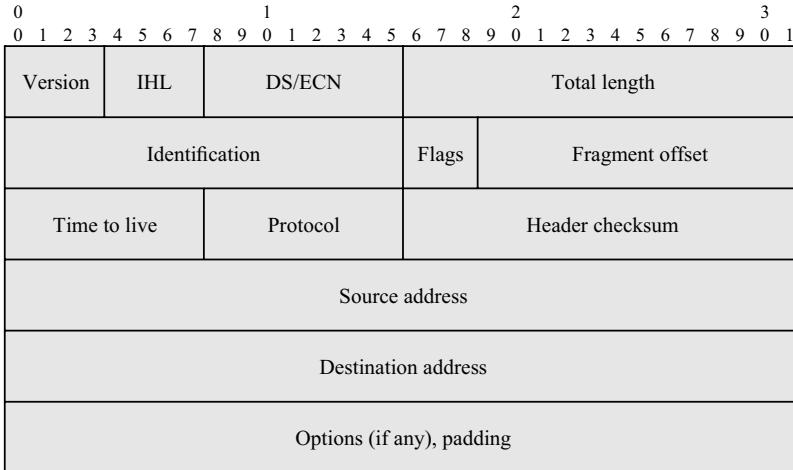


Figure 4.17 IPv4 header [55].

With IPv4, intermediate routers on the path may need to fragment an IP packet, if the network maximum transfer unit (MTU) is smaller than the length of the IP packet.³ The Ethernet MTU size supported may vary. Ethernet jumbo frames allow up to 9000 octets, which is much larger than the commonly referred 1500 octets for Ethernet. Typically bridges and switches can support frames longer than 1500 octets by configuration.

If the MTU of the network does not support the size of the packet to be sent, the router needs to fragment the packet, i.e. divide it into parts so that the length of each part is less than the MTU and send those fragments separately. Fragments travel through the network until they reach the recipient, which can then assemble the fragments back into a single IP packet.

Fragmentation is an issue for the network performance, as it often needs to be done in software, both for the fragmenting and defragmenting functions. This slows down IP forwarding and consumes resources from the devices.

Fragmentation is required for the eNodeB and for the EPC (Evolved Packet Core) by 3GPP TS36.414. Similarly, RNC supports fragmentation of the GTP packets at the Iu interface [81]. TS29.281 (GTP-U v1) instructs to avoid fragmentation by trying to match the path MTU with the inner IP packet size plus tunnel headers (outer UDP, outer IP and GTP headers).

- Identification field can be used together with the flags and fragment offset fields for identifying fragments of the original packet.
- Flags. This field consists of three bits: Bit 0: Reserved, must be zero, Bit 1: Don't Fragment (DF), Bit 2: More Fragments (MF).
- Fragment offset tells the offset from the beginning of the original packet. These fields are used in defragmentation.

If the Don't Fragment (DF) flag is set, the datagram (packet) must not be fragmented. If the network MTU does not support the packet size, the packet is dropped.

³ With IPv6, fragmentation (if needed) is performed by hosts and not by the intermediate routers.

24-bit block	10.0.0.0 - 10.255.255.255
20-bit block	172.16.0.0 - 172.31.255.255
16-bit block	192.168.0.0 - 192.168.255.255

Figure 4.18 RFC 1918 private address ranges.

If a packet is fragmented (DF flag not set), the More Fragments (MF) flag is set. The last fragment travels with the MF set false (value 0).

- Time to live – field provides a means to prevent packets from circulating in the IP network forever. This could happen e.g. due to a misconfiguration such as inserting a static route that causes a loop.

Time-to-live is initially set to some value (e.g. 64) and then decremented by one by each router on the path. When the field reaches 0, the packet is discarded and an ICMP control message is returned to the sender. Time-to-live field also allows the span of routers that a packet may cross to be limited. For example, it can be set to 1 in which case the packet will only reach the next hop router.

Note that GTP-U tunnel endpoints do not need to change the inner packet TTL value.

- Protocol – field indicates the protocol that is carried on top of IP. Examples are UDP, TCP, SCTP, and so on. Protocol field is useful in separating different applications within the same IP host address.
- Header checksum is calculated over the IP header fields only. The data part is not included in the calculation. As each router processes the IP header, and modifies the content of it (e.g. by decrementing the TTL field), a new header checksum is calculated before the packet is forwarded. Note that in IPv6, no checksum exists. This simplifies the router's task as checksum calculation does not need to be performed.

Each IP address contains a network portion and a host portion. Network address is used to transmit packets between networks.

- Source address specifies the sender of the IP packet.
- Destination address defines the receiving node of the packet.
- Options field include further functionality that may be requested.

4.6.2 IP Addresses and Address Assignment

With IP transport, mobile network elements, BTSSs, controllers and gateways use IP addresses as source and destination addresses for the mobile backhaul traffic. None of these addresses

need to be publicly routable IP addresses (with the exception of some of the core interfaces), so private addresses can be used within the radio network. Due to this, there is also typically no shortage of IP addresses with IPv4.

For clarity, end user IP traffic is carried transparently between the terminal and the core network, as discussed in Chapter 3, over GTP-U and other protocols. GGSN/PDN GW allocates IP addresses to terminals and acts as the interface between the mobile network and the Internet. The way IP addresses are assigned to the mobile elements is separate from the assignment of user IP addresses. For the mobile network elements, IP addresses are either statically configured or assigned dynamically via DHCP (Dynamic Host Configuration Protocol).

For IPv4, private IP addresses are defined in RFC 1918.

Networks reserved for private usage include addresses from 10.0.0.0 to 10.255.255.255, from 172.16.0.0 to 172.31.255.255, and from 192.168.0.0 to 192.168.255.255 as shown in Figure 4.18. These networks may be used for addressing mobile network elements, and the network/host portion may be divided flexibly. Refer to an example in Figure 4.19. There is no mobile network originating need to comply with address class boundaries (class A, B, or C and so on). Variable length subnetting can be used. As addressing is not defined in 3GPP it is a question of the network implementation. RFC1918 private address ranges allow for more than 16 million addresses.

IP addresses given to the mobile elements should be unique and the same address should not be used for another element, even if the elements would be in isolated areas. Routing utilizes the network prefix part of the address field in routing decisions.

The division of the IP address field into a network prefix and a host field depends on the amount of hosts in that network. The IP addressing plan should cover the anticipated growth in

		Network prefix				Host bits	
Subnetting 10.0.0.0/27 into /29 subnets		←				→	
Subnet mask for /29		11111111	11111111	11111111	11111111	000	
Subnetwork #1	Network address	10.0.0.0	00001010	00000000	00000000	00000000	000
	First host address	10.0.0.1	00001010	00000000	00000000	00000000	001
	Last host address	10.0.0.6	00001010	00000000	00000000	00000000	110
	Broadcast address	10.0.0.7	00001010	00000000	00000000	00000000	111
Subnetwork #2	Network address	10.0.0.8	00001010	00000000	00000000	00000000	000
	First host address	10.0.0.9	00001010	00000000	00000000	00000000	001
	Last host address	10.0.0.14	00001010	00000000	00000000	00000000	110
	Broadcast address	10.0.0.15	00001010	00000000	00000000	00000000	111
Subnetwork #3	Network address	10.0.0.16	00001010	00000000	00000000	00010000	000
	First host address	10.0.0.17	00001010	00000000	00000000	00010000	001
	Last host address	10.0.0.22	00001010	00000000	00000000	00010000	110
	Broadcast address	10.0.0.23	00001010	00000000	00000000	00010000	111
Subnetwork #4	Network address	10.0.0.24	00001010	00000000	00000000	00011000	000
	First host address	10.0.0.25	00001010	00000000	00000000	00011000	001
	Last host address	10.0.0.30	00001010	00000000	00000000	00011000	110
	Broadcast address	10.0.0.31	00001010	00000000	00000000	00011000	111

Figure 4.19 Network prefix and host address example.

the number of hosts in order to avoid frequent changes. It should allow summarizing routes in the network. This optimizes routing performance and avoids a large memory consumption in the routers by reducing the number of entries routers need to maintain and exchange.

Base stations may consist of one or multiple IP subnetworks (subnets). For LTE, TS36.414 states this explicitly. Subnetwork and addressing planning is a question of network planning and implementation.

IP addresses need to be assigned to the mobile network elements. IP layer connectivity needs to exist in the user, control, synchronization and management plane between the peer entities (e.g. a 3G NodeB and a 3G RNC). User plane bearer IDs, including IP addresses, are then exchanged via the mobile network control plane (signalling via NBAP, RANAP; S1-AP.) For example, LTE S1 bearer is identified with an IP address, port info, and GTP-U Tunnel Endpoint IDs. In the bearer set-up phase the receiving node informs of the IP address into which the bearer shall be terminated.

In general there may be one or multiple IP addresses per BTS. If a single IP address is used for multiple applications (e.g. user plane and control plane), port and higher layer protocol information can be used to direct the traffic to the correct termination point within the BTS.

Traffic of a BTS may be logically separated into different VLANs: control plane, user plane, management plane VLAN, and so on. Each VLAN forms its own IP subnet. If the BTS access is L2 Ethernet based, each of these VLANs may connect multiple BTSs – grouping logically e.g. control plane traffic into a single VLAN. Alternatively, VLANs to individual BTSs can be kept separate from VLANs of other BTSs. This forms a point-to-point structure, and the amount of VLANs grows.

Use of L2 and VLANs in the first place is a network design topic, not defined by 3GPP.

Static configuration is the simplest option for obtaining an IP address. The IP address, including the network prefix and the length of the prefix (subnet mask), and the host portion is configured manually.

An additional topic related to the IP addressing is the use of virtual or loopback addresses. When IP addresses are tied to the physical ports and the port fails, the IP address is unreachable, causing service downtime. A loopback address may still be reachable via another port.

In residential (fixed) internet access and within corporate networks, automatic assignment of IP addresses is commonly used. Dynamic Host Configuration Protocol, DHCP, is used to obtain an IP address, subnet mask and the IP address of the default gateway. Basically similar methods are usable for the mobile backhaul, although for efficient address assignment the element addresses may not be randomly chosen. This topic is not covered by 3GPP so there may be different implementations for automatic assignment of addresses and other configuration information.

With DHCP, a host sends first a DHCP broadcast message (DHCP discovery) which reaches all devices on the same L2 broadcast domain. The first router, if not a DHCP server, should be configured to relay this message to a DHCP server. The DHCP server maintains a pool of addresses and assigns an address to the host from this pool.

The addresses given by the DHCP server include a lease time. Lease time is a period in time which the address is owned and usable by the end node. A benefit of the lease time is that addresses can be freed for other nodes, after the expiry of the timer.

The total number of BTSs in a network may be in the thousands range or more. Network management system can be used for the IP address configuration. When the management

plane connection exists further parameters can be downloaded from the central network management system.

4.6.3 *Forwarding*

IP forwarding is a user plane operation guided by forwarding table entries. IP routing is often used interchangeably with IP forwarding. The forwarding table associates the destination network with an outgoing interface/next hop router. For the delivery of an IP packet over Ethernet media (within an Ethernet-encapsulated frame), the Ethernet MAC address of the next-hop router is also required. This is obtained via ARP (Address Resolution Protocol).

IP is connectionless, with no prior connection set-up needed. Each IP packet is treated independently from any other IP packet. Packets are forwarded hop-by-hop based on the information of the topology and metrics routers on the way to the destination have. Each router selects the next hop (the best path towards the destination) based on the information of possible routes and other criteria it has.

At the IP layer a better control of routing may be implemented by the use of policy based routing. MPLS Traffic Engineering also addresses this need.

A forwarding table may consist of dynamic and static entries. Routes to remote networks are learned dynamically with routing protocols while static routes are entered by manual configuration. The forwarding occurs the same way with both dynamically learned routes and static routes. The IP header is examined. If a route to the destination network exists, the IP packet is forwarded to the next-hop router. If there is no route, the packet is discarded and an error message is returned to the sender.

Forwarding is unidirectional. For the return traffic another entry in the forwarding table is needed. Traffic may follow the same path in both directions but it may also use a different path in the other direction. Asymmetric routing needs consideration when used with firewalls.

IP header includes a number of fields which need to be examined and processed in addition to the destination address field. Time-to-live field is decremented at every hop along the way to the destination. If its value reaches zero, the packet is discarded. This prevents packets from looping forever. Differentiated Services (DiffServ, DS) field indicates quality of service. Checksum is calculated. Optional headers may be included.

Longest prefix match is commonly used in determining the next hop for the destination network. As an example, if the destination address is 10.1.1.1, and the forwarding table includes entries for 10.1.1.0/24 and for 10.1.1.0/28, the latter one is used for determining the next hop, since the entry is more specific. The entries are separate networks from the routing protocol viewpoint since the prefix length differs.

A default route may be configured to be used when there is no other route. Default route points typically to another router ('a gateway of last resort'). This router should then know more destination networks. Otherwise the packet is dropped.

A routing protocol may be used. Routers communicate and share information of networks with other routers. A metric is used to evaluate the attractiveness of a specific path.

Often static routes also exist in the forwarding table. One example is the default route entered by configuration. There can also be a static route and a dynamically learned route to the same destination, with a network prefix that is equally long. A way of determining the preferred route is needed in this case, e.g. by a 'preference' parameter.

If the next hop router is not available the entry in the forwarding table can be removed. The next hop router unavailability may be caused by a link failure or by a node failure. If the link fails e.g. due to a cable cut, the signal is physically down, and this indicates the failure situation to the router. Link failure may, however, be blocked by an intermediate device, such as a L2 Ethernet switch or a microwave radio link. In this case a routing protocol detects the failure.

If no routing protocol is operated an alternative way for detection is needed. These include BFD (Bidirectional Forwarding Detection) and L2 protocols such as Ethernet OAM. Detecting the failure by a physical signal down indication is often the fastest method, since the least amount of processing by SW is needed.

When a route is removed, another potentially less-optimal route may be found and be taken into active forwarding.

Having equal cost multiple paths load sharing is achieved. In this case two (or more) routes match the destination network prefix. All of these routes may be used simultaneously. Load sharing algorithms use e.g. UDP ports in addition to IP source addresses, to select which of the paths to take. In failure scenarios, load sharing is advantageous since another route to the destination is already known and in use.

Policy-based routing utilizes in addition to (or instead of) the IP destination address also other fields of the IP header. The next hop is selected e.g. based on the IP source address (source routing) or based on the DiffServ field. Source based routing may e.g. guide traffic from an endpoint to exit a certain interface. DiffServ field in the route determination allows the IP network to be developed towards a traffic engineered network.

Service providers have a need to support multiple customers on a single network. Customer specific routes are kept separate from other customers' routes by a virtual routing and forwarding (VRF) function. A routing protocol advertises routes specific to a customer. Routes are isolated and do not leak between VRFs. The routes of the provider network are also kept separate from all of the customer specific routes.

With the BTS first mile access, there is typically only a single physical link available. All traffic is transmitted via that link towards the aggregation tier and to the first hop gateway. IP forwarding is simple. The only entry needed in the forwarding table of the BTS is the IP address of the default gateway. The first IP device then may have multiple alternatives for the next hop. For more complex topologies however, routing protocols provide more benefits and with the aggregation/core network based on IP/MPLS they are commonly used.

4.6.4 Routing Protocols

Routers need information of the next hop router/outgoing interface towards a given destination network. This information is learned dynamically with the help of a routing protocol. A routing protocol learns routes to destination networks from all other routers. These learned routes are stored in a routing table and selected entries from all of the routing information are used in the active forwarding configuration, in a forwarding table or a similar construct.

3GPP has no guidance of the use of routing protocols, whether any routing protocol is used, and if so, which one. An IP host, e.g. a BTS, may have a static pre-configured entry for the default gateway only or it may run a routing protocol. In 3GPP systems other than LTE, all

traffic in the access area from a BTS is addressed to a controller, a BSC (2G) or RNC (3G), so the logical topology in 2G and 3G is hub-and-spoke.

Also, BTS often does not have many routes to the hub site (BSC or RNC), instead, it typically has a single physical link towards the aggregation network. In higher aggregation tiers, more routing possibilities may exist. If the first IP device is at the edge of the aggregation network, BTSs forward all traffic to the default gateway. This gateway may then use routing protocols to learn of available paths to the mobile network hubs (BSC and RNC).

In LTE, the X2 interface when implemented as a direct X2, introduces horizontal traffic between neighbour eNodeBs. A routing protocol may be used to learn routes to the X2 destinations. Often, instead of direct X2 links, X2 connectivity is arranged from a site higher in the network, e.g. a site with an IPsec GW. If X2 is implemented this way again all traffic is forwarded to the hub site.

Even if the mobile network elements would not implement any of the routing protocols, routing protocols are widely deployed in both service provider and enterprise networks. Most likely the core or backbone transport network is using some routing protocol, and if MPLS-based VPNs are used, routing protocol convergence in the MPLS network also impacts recovery within the mobile backhaul.

Routing protocols are classified as link-state or distance-vector protocols. Link-state protocols, such as OSPF and IS-IS, maintain a full topology of an area, and react on network changes. Distance-vector protocols rely on other routers for information of the best route to a given destination network rather than maintain the full topology.

Another classification is based on the role of the protocol in the network: Interior Gateway Protocols (IGPs) are exchanging routing information within an Autonomous System (AS). Exterior Gateway Protocols (EGPs) exchange routing information between Autonomous Systems.

Routing Information Protocol (RIP) is an example of a distance-vector routing protocol. OSPF and IS-IS are link-state routing protocols. OSPF, as well as IS-IS, is built on Dijkstra's shortest path first – algorithm.

BGP is a distance vector EGP protocol. BGP includes extensions defined in multiprotocol-BGP (MP-BGP) related RFCs. MP-BGP is used e.g. with MPLS L3 VPN application.

4.6.5 *Differentiated Services*

Two Quality of Service models, Differentiated Services (DiffServ, DS), and Integrated Services (IntServ), are included in the IETF RFCs [88], [89], [90], [91]. Differentiated Services is based on service separation on a per-hop basis without a need for signaling. Integrated Services model assumes applications communicating their QoS requirements to the network, which then controls whether the required resources are available and reserves them with a signaling protocol.

3GPP defines for the IP transport that the IP DS is used, and that the mapping between traffic categories and DS-CPs shall be configurable.

The QoS field of the IP packet header is shown in Figure 4.20, with the original RFC791 definition (outdated) and the Differentiated Services definition of RFC2474.

The original RFC791 included a 1-octet, type of service field for indication of QoS. Two bits are 0 (unused). Since RFC791, Precedence field (3 bits), and D (Delay), T (Throughput), R (Reliability) bits are replaced by Differentiated Services field.

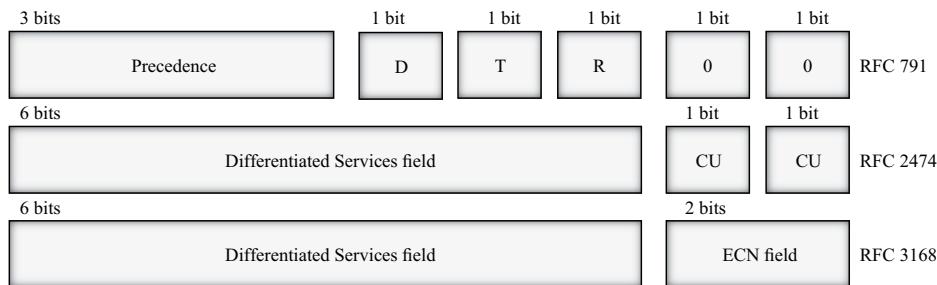


Figure 4.20 QoS field in the IP packet header.

RFC2474 defines six bits of this 1-octet field for Differentiated Services (DS). This DS field, together with two unused bits (CU, currently unused), replaces the previous type of service field. With RFC3168 these two unused bits are allocated for Explicit Congestion Notification (ECN).

4.6.6 Address Resolution Protocol

An Ethernet L2 MAC address of the next hop router is needed before an IP packet can be sent over the Ethernet interface. Before a BTS can send any IP packet over its Ethernet interface, it needs to know the MAC address of the next hop router. The protocol to query the MAC address of a specific IP address is Address Resolution Protocol (ARP), defined in RFC 826 [94].

ARP is a L2 Ethernet broadcast which is asking for the MAC address of the next-hop IP address. The router owning the IP address replies and returns its MAC address. This MAC address is then used as the L2 destination address for the Ethernet frame to be sent. Hosts store the binding of IP to MAC address in order to avoid asking frequently for the same information. On the other hand, after a certain time has elapsed, the entry times out if the MAC address has not been needed. If a new IP packet needs to be sent, a new ARP request is first sent.

A gratuitous ARP allows a device to proactively inform its MAC address without having to wait for a new ARP request to be sent first. The device receiving gratuitous ARP can then update its IP to MAC address binding immediately. So the IP address can be reached again without an added delay.

4.6.7 ICMP

ICMP is defined in RFC 792 [95], Internet Control Message Protocol (ICMP). ICMP is an integral part of the IP Protocol and all routers must support ICMP. With ICMP hosts and routers can report errors and exchange diagnostic, control and status information.

ICMP messages may be grouped into three categories:

- ICMP error messages.
- ICMP request messages, and
- ICMP reply messages.

ICMP error messages include Destination Unreachable, Source Quench, Redirect, Time Exceeded and Parameter Problem. ICMP request messages include ICMP Echo, Router

Table 4.2 Port numbers (www.iana.org).

0..1023	Well known ports, assigned by IANA
1024..49151	User ports (a. k. Registered ports), assigned by IANA
49152..65535	Dynamic ports (Private ports or Ephemeral ports), never assigned

Solicitation, Timestamp, Information Request (obsolete) and Address Mask Request. ICMP reply messages include ICMP Echo reply, Router Advertisement, Timestamp Reply, Information Reply (obsolete) and Address Mask Reply.

Destination unreachable is sent back to the source, when a router cannot find a route to the destination. Redirect informs that another next hop should be used. When TTL expires, Time Exceeded should be sent back to the source.

Ping (Packet Internet Groper) and Traceroute are two common applications that use ICMP Echo Request and Reply messages. Ping tests connectivity at the IP layer, and Traceroute is used to detect a path the packet takes.

ICMP messages are however often blocked for security reasons. This also limits the applicability of the Ping tool. Especially when the ICMP packet crosses an administrative boundary (another operator's network) it is likely that Ping does not work.

4.6.8 UDP

UDP (User Datagram Protocol) is defined in RFC 768.

UDP is referred to as connectionless and unreliable, since it does not guarantee delivery of the packets and has no retransmission mechanism. Packets may also arrive out of order, or be duplicated. The application needs to be able to deal with these characteristics. As an alternative, TCP can be used, when reliable connection is needed.

UDP allows identification of the process within the host of a given IP address via the UDP port number. So a single IP address serves multiple communications, where the UDP port number is used together with the IP address, identifying the communication endpoint.

A number of ports have been assigned by IANA (Internet Assigned Numbers Authority).

With UDP, a port number is often randomly selected from the dynamic port area (Table 4.2).

A UDP datagram consists of a UDP header and data. The header (Figure 4.21) consists of Source and Destination port numbers, a Length field and of a Checksum. Length tells the length of header and data. Checksum is optional, and is calculated over the header and data.

With the mobile backhaul application, UDP is the protocol on top of IP in user plane, and interfacing the radio layer protocols. (Iub frame protocol, Iur frame protocol, Iu user plane protocol and GTP-U protocol on S1 and X2).

The 3G Radio Access Bearers (RABs) are identified with the UDP port number among other information. For example, in the Iu interface, packets of a Radio Access Bearer are sent to an

Source port (16 bits)	Destination Port (16 bits)
Length (16 bits)	Checksum (16 bits)

Figure 4.21 UDP header.

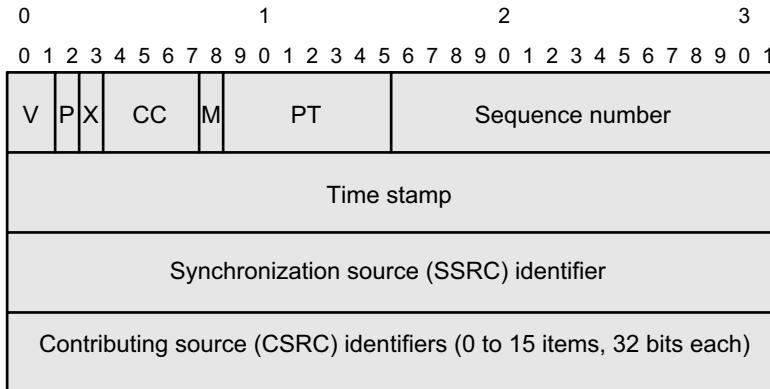


Figure 4.22 RTP header

RNC IP Address and an UDP port. In LTE, EPS bearers are identified by the GTP-U Tunnel Endpoint Identifier (TE ID) and the IP address (source TE ID, destination TE ID, source IP address, destination IP address). GTP-U uses destination UDP port number 2152.

GTP-U response message may have a different UDP port and IP address than the triggering message. If a stateful firewall is used, this has to be taken into account.

4.6.9 RTP

For CS traffic on the Iu-cs interface in 3G, Iu user plane utilizes an RTP/UDP/IP protocol stack. Also, in 2G for A over IP, RTP/UDP/IP is used. Transport bearer is identified using source and destination UDP port numbers and source and destination IP addresses. If multiple IP addresses are available in the RNC and in the Core Network, the address/port combination given in RANAP associates the packets to a particular Radio Access Bearer (RAB).

RTP (Real-time Transport Protocol, defined in RFC3550) is designed for end-to-end delivery of real-time data, e.g. interactive audio and video. UDP allows for multiplexing and possibly for the use of checksum. Services of RTP include payload type identification, sequence numbering, timestamping and delivery monitoring. RTCP (RTP Control Protocol) may also optionally be applied.

RTP header is shown in Figure 4.22.

The header fields and their usage as defined for Iu-cs are [67]:

- V stands for version. The field is 2 bits long. The version defined in RFC3550 is 2. Value ‘1’ is used for the first draft version. Version 2 is used.
- Padding (P, 1 bit field) means that if the bit is set, the packet contains padding octets at the end, which are not part of the payload. Padding is not used.
- Extension (X, 1 bit field) bit set means that the fixed header has to be followed by one header extension. No extension headers are used.
- CSRC count (CC, 4 bits), tells the number of CSRC identifiers that follow the fixed header.
- Marker (M, 1 bit) interpretation is defined by a profile, intended to allow frame boundaries be marked in the packet stream. No contributing sources.

- Payload type (PT, 7 bits) defines the format of the payload. A dynamic payload type is used, with values between 96 and 127. At the receiving peer, the value is ignored.
- Sequence number (SN, 16 bits), is used by the receiver to detect packet loss, and restore packet sequence. The initial value should be random, to mitigate known-plaintext attacks on encryption. Sequence number is supplied by the RTP PDU source. The sink may ignore the sequence number. It may also be used for statistics about quality of the link, or correct out-of-sequence delivery.
- Timestamp is a 32-bit field, indicating the sampling instant of the first octet in the payload. The sampling instant is the reference point, in order to allow synchronized presentation of all media, sampled at the same instant. Timestamp is supplied by the RTP PDU source, and the clock frequency of 16 kHz is used. Timestamp may be ignored by the receiver. It may also be used for statistics of the link quality, or it may be used to correct jitter.
- SSRC identifies the synchronization source. Within an RTP session, no two sources should have the same identifier. Synchronization source is supplied by the RTP PDU source, and the sink may ignore the SSRC if RTCP is not in use.
- CSRC list identifies the contributing sources. An example is that when audio packets are mixed together, talker can be identified at the receiver. No contributing sources.

4.6.10 TCP

In the mobile backhaul protocols stacks, TCP is not included, since the user plane is over UDP/IP, and control plane over SCTP/IP. TCP may still be used in the management plane. Additionally, many end user applications on terminals are built on top of TCP/IP (such as web browsing with http). QoS related topics of TCP are essential. These are addressed in the QoS chapter.

TCP is also used as a reliable transport mechanism for some of the IP and MPLS control plane protocols. For example, BGP and LDP (Label Distribution Protocol) use TCP.

TCP is a connection-oriented and reliable transport protocol. In addition to the basic data transfer, TCP offers reliability, flow control, multiplexing and connections.

TCP header fields (Figure 4.23) are:

- Source and destination ports identify the sending and receiving port.
- Sequence number helps the receiver reorder segments that may be received out-of-order, or may be duplicated.

Basic transfer is octet-oriented. A number of octets are collected into a segment, which is transmitted. A sequence number is assigned to keep track of the octets, so that the TCP receiver can recover the original octet streams from potentially out-of-order or duplicated segments.

- With the acknowledgment number – field, the receiver acknowledges received octets, by indicating the sequence number of the next expected octet.
- Due to variable length Options – field, the data offset – field is used to indicate the size of the TCP header and correspondingly where the data begins.

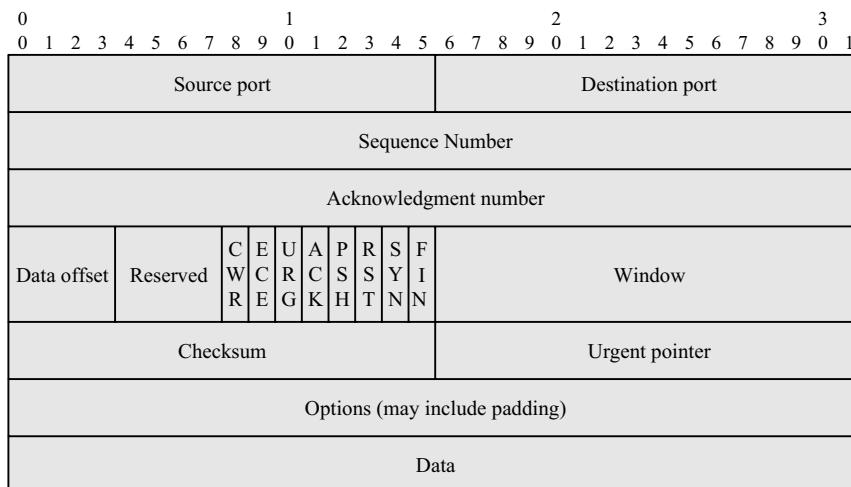


Figure 4.23 TCP header [57].

The next field includes reserved bits and individual flags. Original RFC 793 defined flags are:

- URG (Urgent pointer field is significant).
- ACK (Acknowledgment field is significant).
- PSH (Push Function).
- RST (Reset the connection).
- SYN (Synchronize sequence numbers), and
- FIN (No more data).

RFC 3168 adds the use of two additional bits for explicit congestion notification to the header. These bits are marked:

- CWR (Congestion window reduced) and
- ECE (ECN Echo).

Window tells the number of data octets the sender of this segment is willing to accept.

Use of checksum is mandatory with TCP (as opposed to UDP).

- Checksum is calculated over a pseudo-header. Pseudo-header is shown in Figure 4.24. It includes 96 bits from the IP layer header; namely IP layer source and destination addresses, protocol type, and length of the TCP header and data. In addition to these fields from the IP layer header, TCP header and TCP data fields are included into the checksum calculation.
- Urgent pointer is the current value of the urgent pointer, pointing to the sequence number of the octet following the urgent data. (Only interpreted in segments with the URG set).

Source address		
Destination address		
zero	Protocol	TCP length
TCP header (Variable length)		
TCP data (Variable length)		

Figure 4.24 Pseudo-header for checksum calculation [57].

4.6.11 SCTP

SCTP (Stream Control Transmission Protocol) provides reliable transport service for messages over the IP. One of the design goals of SCTP was to support telecom signalling (such as SS7), although it can be used for other purposes as well.

In the mobile backhaul, it is used for signalling traffic. In 3G Iub, NBAP is carried over SCTP, and in Iu/Iur interface, RANAP and RNSAP similarly rely on SCTP. LTE signalling, S1 and X2 control plane, is over SCTP.

TCP also provides reliable service over the IP. One main difference is that TCP is byte-oriented, while SCTP is message-oriented. TCP essentially transmits bytes reliably, in segments. SCTP indicates the beginning and end points of a message. This is useful in signalling applications, which are based on message transactions. SCTP also has further functionality which does not exist in UDP or TCP: a support for multi-homed hosts, and a support for multiple streams in a single SCTP association.

Key features of SCTP are:

- acknowledged transfer of user data, error-free, and non-duplicated;
- fragmentation of data;
- sequenced delivery of messages within multiple streams;
- bundling of multiple user messages into a single SCTP packet (optional);
- multi-homing at one or both ends of the association.

SCTP is resistant to flooding and masquerade attacks. It also includes congestion avoidance functions.

SCTP is connection-oriented, meaning that SCTP connections need to be opened and closed. A connection is noted by an SCTP association. The SCTP association can be viewed as connectivity between SCTP endpoints. The SCTP endpoints are represented by an IP address and an SCTP port.

Multiple streams means a capability to transport reliably separate sequences of messages. These sequences of messages are called streams, and a stream identifier is used to separate the streams.

SCTP packet format is presented in Figure 4.25.

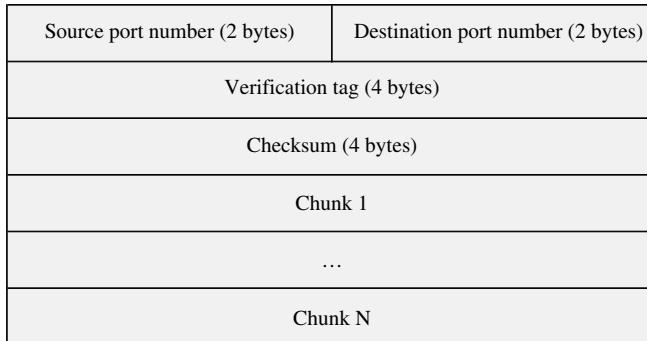


Figure 4.25 SCTP packet format [97].

SCTP common header includes source and destination port numbers, a verification tag, and a checksum. Data portion of the packet consists of chunks, which may be data chunks or control chunks. The amount of chunks varies.

Source port number means the sender endpoint's SCTP port number. Destination port number is the SCTP port number of the intended receiver. Verification tag protects against an attacker injecting data into an existing association. It also prevents packets from a previous association between the same endpoints being assumed valid for the current association.

Checksum is calculated over the entire SCTP packet (SCTP common header and all chunks).

SCTP session start-up consists of four messages: Init, Init-ack, Cookie-Echo, Cookie-Ack. This four-way handshake is designed to address TCP Syn attack. Cookie-Echo and Cookie-Ack can also carry data.

4.6.12 IPv6

One main driver in general for IPv6 is the unavailability of public IPv4 addresses. Public IPv4 addresses have been allocated and the availability (or unavailability) of usable public addresses varies based on geographical area and organization. IPv4 addresses are a scarce resource, and several layers of NAT (network address translation) may be needed, which complicates the network and also introduces issues for many applications.

For the mobile backhaul application IP addresses do not need to be public (apart from some interfaces in the core network), and thus there is in general no pressing need for a larger address space. Other benefits of IPv6, such as an improved header structure, use of link local addresses, no broadcasts, no fragmentation in the intermediate networks, etc., remain.

Even though both IPv4 and IPv6 are IP protocols, they are different protocols. IPv4-only device is not capable of routing IPv6 packets. In practice, many routers today support both protocols. Still, taking IPv6 into use in the network requires preparation and planning. ARP is not used. Separate versions of routing protocols are needed. Checksums are removed from the IPv6 layer. IPv6 basically mandates the use of IPsec, however this also requires key management.

Figure 4.26 shows the IPv6 header. Version marks the version of the protocol, 6 defines IPv6. Traffic class, 8-bit field, is for marking of QoS. Flow label can be used to identify a

Version (4 bits)	Traffic class (8 bits)	Flow Label (20 bits)		
Payload length (16 bits)		Next header (8 bits)		Hop limit (8 bits)
Source address (128 bits)				
Destination address (128 bits)				

Figure 4.26 IPv6 header [50].

sequence of packets that may need e.g. special QoS treatment (real-time). Payload length defines the length of the payload following the header fields. Next header-field identifies the header following the IPv6 header. Hop limit is comparable to TTL in IPv4, it is decremented by 1 in each forwarding node and packet is discarded when the hop limit reaches 0. Source and destination addresses are now 128 bit fields instead of 32 bits in the IPv4.

4.7 MPLS/IP Applications

Multiprotocol Label Switching (MPLS) label is assigned between the L2 and L3 headers, making MPLS a ‘L2.5’ technology. Figure 4.27 shows MPLS protocol stack in an application of MPLS carrying IP and using Ethernet as the L2.

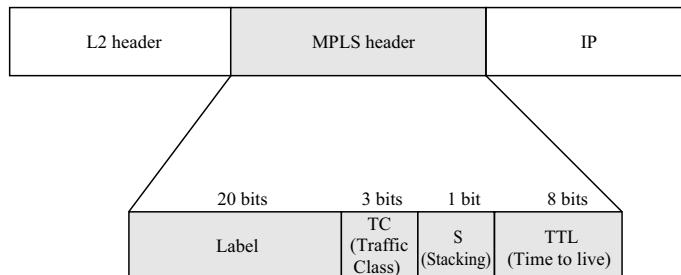
MPLS header fields are Label, Traffic Class, Stacking, and Time-to-Live.

Twenty bits are used for the Label. With cell-mode MPLS, ATM VPI/VCI fields indicate the label. Similarly, frame relay DLCI information can be used for the label. In the subsequent text the focus is on frame mode MPLS with the 20-bit label as shown.

Traffic Class-field is used to mark Quality of Service. The field was originally marked as EXP – Experimental, although the three bits were used for QoS marking. Now the bits have been more correctly named in the IETF (RFC 5462) as Traffic Class.

S- stacking bit: This bit tells whether there are further labels in the stack to be processed.

TTL – time to live. Time-to-live field is used as in IP. The field can also be copied from the IP layer TTL field (TTL propagation). At the egress of the MPLS network, the outgoing IP packet

**Figure 4.27** MPLS header.

should have a TTL value that is equal to the one it would have if it had traversed through the network without label switching.

The initial driver for MPLS was a need to speed up IP forwarding. When a label is assigned, forwarding decisions are done by examining this label, instead of an IP header lookup. Forwarding of packets could be implemented in hardware, as there was no need to go through the IP header fields in every router in the MPLS network. Since then, IP forwarding in routers is mostly performed in hardware and thus the original performance gain does not exist.

What makes MPLS attractive today is the number of applications it supports: IP MPLS VPNs, L2 VPNs, Pseudowire emulation, Traffic engineering with a fast re-route, Transport network behavior (MPLS-TP), etc. From a technology viewpoint, use of an IP control plane is a benefit since common routing protocols can be used. For the mobile backhaul the support of legacy interfaces (TDM-Abis and ATM-Iub) in addition to the native IP interfaces (IP-Iub, and S1/X2) in the same MPLS network is often important as well.

In addition to the RFCs, further reading on MPLS is provided e.g. in [105], [106] and [107].

4.7.1 *MPLS Architecture*

At the MPLS ingress, packets are classified into Forwarding Equivalence Classes (FECs) and are assigned an MPLS label accordingly. This label then represents the FEC. Subsequent forwarding takes place using the label, so there is no need for a complete IP header lookup in the next MPLS routers. At the MPLS egress, the original packet is sent out of the correct interface with the label removed.

MPLS routers are called Label Switch Routers (LSRs). A Label Switched Path (LSP) is the path through the LSRs. The ingress LSR pushes an MPLS label, and forwarding is based on the label until the egress of the LSP. The label is removed (popped) and the original packet is delivered at the egress. In the intermediate LSRs, labels are swapped. MPLS supports multiple levels of labels, so there may be a label stack consisting of n labels. Multiple labels are used in MPLS applications, such as MPLS traffic engineering (TE) and MPLS VPNs.

In the LSP, the MPLS label may be removed (popped) already one hop before the egress LSR, because the egress LSR anyway does not forward the packet anymore based on the label. So the label may be popped by the router prior to the egress LSR. This behavior is called penultimate-hop-popping (PHP). In some MPLS applications however, the label is needed until the egress LSR, because otherwise the LSP is not continuous until the egress LSR. In that case, PHP cannot be used, and the label is kept until the egress LSR.

In downstream-on-request label assignment, an upstream LSR asks for the label from the downstream LSR, which then assigns the label and announces it. In unsolicited mode, labels are assigned without a request, and informed of. An LSR may learn of label bindings that are not the next hop for a certain FEC. In liberal retention mode, these labels are maintained. They are dropped in conservative retention mode. Liberal retention mode has the benefit, that labels for an alternative path already exist, if they are needed for recovery.

Multiple protocols support label assignment; Label Distribution Protocol (LDP), RSVP-TE, and MP-BGP. If MPLS applications are combined (e.g. traffic engineered VPNs), two or more of these protocols may be distributing labels, which results in a label stack. The MPLS forwarding is based on the labels in a similar way, independently of which protocol was used to distribute the labels.

4.7.2 Label Distribution Protocol

With MPLS, the meaning of labels has to be agreed between Label Switch Routers (LSRs). Label Distribution Protocol (LDP), defined in RFC 5036, is one of the protocols that can be used for this purpose. Label Switched Paths (LSPs) are established through the MPLS network. At the MPLS network ingress, IP layer information (in the case of MPLS usage for IP) is used in mapping to an LSP.

LDP uses TCP as the underlying transport protocol for the LDP sessions. UDP is used in the initial discovery phase.

Each LSP corresponds to a Forwarding Equivalence Class (FEC). A FEC defines which packets are mapped to which LSP. For each LSP, a FEC specification must be provided, so that packets can be identified for mapping into the LSP. RFC5036 specifies one FEC element, which is the address prefix. Packets with destination addresses that match the specified prefix, are then mapped to the LSP.

LDP messages fall into four categories:

- peer discovery;
- session management;
- label distribution (advertisement);
- notification messages (errors, advisory information).

A Hello message helps to discover the LSRs. Hello message uses UDP and is sent as a multicast to all routers in the subnet. After discovery, initialization proceeds based on TCP. After initialization, the two LSRs are peers, and can proceed in exchanging advertisement messages.

An extended discovery mechanism helps to discover LSRs that are not directly connected. In this case, a targeted hello using UDP is sent to a specific address, to a well-known LDP discovery port.

After discovery of the neighbours, LDP sessions are established, using TCP, and the session is maintained using session management messages.

Topology of the network is learned via an IGP, such as OSPF. LSRs then locally assign labels to the destination prefixes. The labels can be assigned per-platform or per-interface. Per-platform means that a single label is used on all interfaces. Per interface means interface-specific incoming labels.

4.7.3 BGP

As opposed to OSPF and IS-IS (Intermediate System to Intermediate System, an OSI defined routing protocol), both of which are interior gateway routing protocols, BGP is an exterior gateway protocol. Concerning mobile backhaul, one application of BGP is with MPLS VPNs.

BGP is designed to connect autonomous systems (ASs). BGP can handle a large amount of routes, such as the whole Internet routing table. With BGP, traffic flow is controlled via a number of attributes.

BGP is defined in RFC4271, A Border Gateway Protocol-4 (BGP-4). Multiprotocol extensions for BGP-4 are in RFC4760. BGP applications include e.g. IP MPLS VPN (RFC 4364) and using BGP for VPLS autodiscovery and signaling (RFC4761).

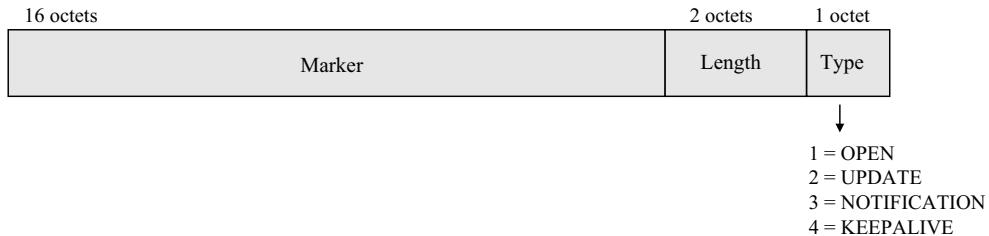


Figure 4.28 BGP header [87].

EBGP (External BGP) means a BGP connection between external peers. External peer means a peer in a different AS. IBGP (Internal BGP) means a BGP connection between internal peers, that is, peers in the same AS.

BGP uses TCP as the underlying transport protocol, with TCP port 179. TCP supports reliable transport, fragmentation and sequencing. TCP connection is opened between the two BGP systems.

BGP exchanges network reachability information with other BGP systems. The IP prefix is included in Network Level Reachability Information (NLRI). Routes contain attributes of the path. An unfeasible route is one which was advertised but is no longer available.

Routes are advertised by the BGP Update message. If multiple routes share the same path attributes, multiple prefixes can be included into a single Update message. Withdrawn routes field of the Update message can be used to inform that the route is no longer in use.

The BGP message format is shown in Figure 4.28.

Marker is 16 octets, all ones. Length means length of the message (header included). Message type is indicated in the Type-field: Open, Update, Notification and Keepalive. Further message types have been added, e.g. RFC 2918 Route Refresh.

A BGP Update message fields are shown in Figure 4.29. BGP header is included in the Update message.

BGP updates advertise feasible routes, or withdraw routes. The Withdrawn routes length tells the length of the withdrawn routes field. Withdrawn routes fields include the length of the prefix and the network prefix itself. The total path attributes length is for the length of the Path attributes field.

Each path attribute is a triple consisting of type, length and value (TLV). Attribute type field consists of two octets: Attribute Flags (1 octet), and Attribute Type Code (1 octet). The highest order bit of the Attribute Flags is the optional bit. If set to 1, it is optional. If 0, it is well-known. The second high-order bit is the transitive bit, which tells whether an optional attribute is carried on to other peers (transitive attribute), or not (non-transitive attribute).

Path attributes may be well-known mandatory, well-known discretionary, optional transitive or optional non-transitive. Well-known attributes all BGP implementations must recognize. Well-known mandatory attributes are included into every BGP update message. Origin, AS path and Next hop are mandatory well-known attributes. Well-known discretionary attributes may be included, but this is not mandated.

Optional attributes do not need to be supported by all BGP systems. Transitive optional attributes are attributes that should be passed on to BGP peers. The mandatory path attributes, Origin, AS Path, and Next Hop are discussed briefly.

Withdrawn routes length	2 octets
Withdrawn routes	Variable length
Total Path Attributes length	2 octets
Path attributes	Variable length
Network layer reachability Information	Variable length

Figure 4.29 BGP update [87].

For Path attribute Origin the value is set to zero, if the NLRI is learned through an interior gateway protocol. Value 1 is reserved for EGP (historical), and value 2 means incomplete. This is used when the NLRI is learned through other means than IGP or EGP.

Path attribute AS path tells the ASs through which the routing information was propagated. Basically each router adds its own AS number. This way, the router can detect a potential loop. The AS path should not already contain its own AS.

Next hop – attribute defines the IP address that should be used for the next hop, for the destinations included in the NLRI field of the update message.

The metric in BGP essentially consists of path attributes, instead of a simple metric such as cost in OSPF. This allows more information and more flexibility for selection of the best path.

BGP, which is used for the public Internet between Autonomous Systems, can be further scaled with route reflectors and confederations.

4.7.4 MPLS Ping

Ping and Traceroute applications are needed with MPLS forwarding, as the IP layer tools alone do not necessarily verify that connectivity exists on the MPLS layer.

RFC 4379 defines an MPLS Echo Request, and an MPLS Echo Reply, that can be used for MPLS-Ping and MPLS-Traceroute applications. The MPLS echo request is an UDP packet, with TTL of 1 in the IP header. The destination UDP port reserved for the MPLS Echo Request is 3503. LSP Ping can be used to test the connectivity of the user plane, with the idea that the LSP Ping travels through the same path as MPLS user plane (FEC).

4.7.5 MPLS L3 VPN and MP-BGP

MPLS L3 VPN, or alternatively IP MPLS VPN, is a flexible tool that can be used for varying connectivity needs. It can be said to provide a ‘point-to-cloud’ connectivity: CE connects to the PE. Via PE, other networks can be reached.

This type of characteristic is useful in the mobile backhaul. When the basic infrastructure is implemented ('the cloud') it is straightforward to add further sites, or arrange connectivity to other sites as needed. One can support connectivity between all sites or a subset of sites depending on the needs. A typical case is shown in Figure 4.30.

In Figure 4.30 an example connectivity from an eNodeB to the GW is drawn with a solid line. eNodeBs connect to the GW either via an aggregation router (CE, in the left hand side), or directly to a PE device (as for PE30). In the first case, CE device peers with the PE. In the second case, eNodeBs peer directly with the PEs (the eNodeBs are then in the role of a CE). Between an eNodeB and the PE, Ethernet devices and microwave radios may exist in the access network. These nodes are not visible in the PE-CE routing peering. A PE-CE routing protocol, for example, OSPF, is used to distribute routes between the PEs and the CEs. Other routing protocols may as well be used as well as static routes.

For resilience, critical sites (such as GW) are dual-homed. Dual-homing of any other site can be supported as well.

The routing protocol used to exchange customer prefixes between the PE devices, is Multiprotocol-BGP (MP-BGP). MP-BGP supports extensions for carrying prefixes other than the basic IPv4 ones. In the IP MPLS VPN application, customer prefixes are appended with a Route Distinguisher (RD), creating a VPNv4 address family. Similarly for IPv6 a VPNv6 address family can be created. VPNv6 address family allows the core routers to be kept unaware of IPv6, so it is one possibility for the introduction of IPv6 in the mobile network, especially in a case where an MPLS network already exists.

IP MPLS Virtual Private Network (VPN) means service provider (SP) infrastructure can be shared for a number of customers. Each customer is offered an IP layer connectivity to her sites, and virtually each customer appears to have her own network. The network devices and links of the service provider are, however, shared for the use of several customers.

In the IP MPLS VPN model, customer edge (CE) router is peering with the service provider edge (PE) device. CE sends routing information to the PE, and the PE routers share VPN

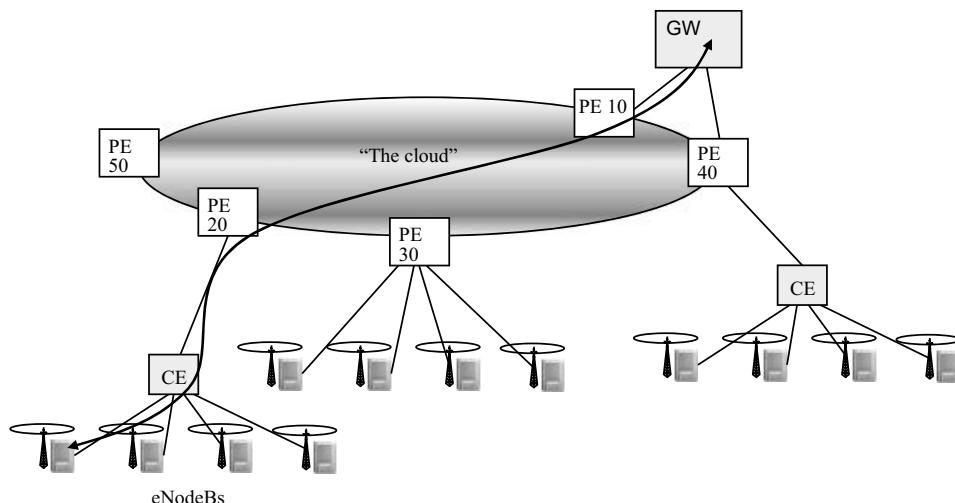


Figure 4.30 Example L3 VPN application for the mobile backhaul.

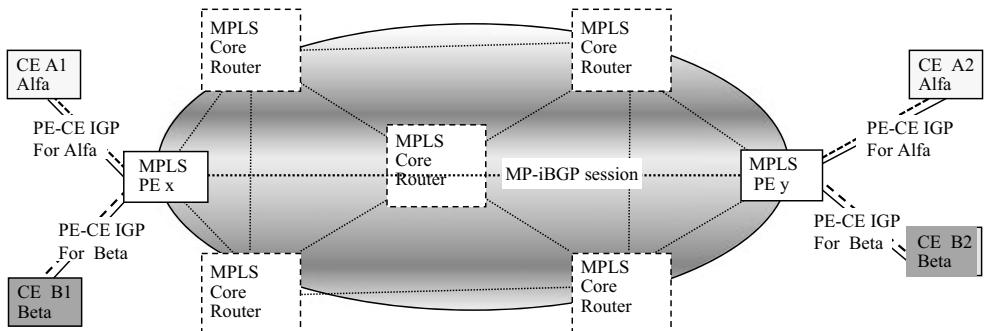


Figure 4.31 MPLS L3 VPN

routing information with the other PE devices of that VPN. Which routes get advertised and accepted, is controlled by BGP export and import route targets.

MPLS core routers (P routers) do not need to know of customer networks. MPLS PE and P routers run an interior gateway protocol (IGP), such as OSPF, within the MPLS network. (This does not interact with the customer IGP).

Figure 4.31 shows two customers, Alfa and Beta. Customer Alfa's edge devices are CE A1 and CE A2, Customer Beta's are correspondingly CE B1 and CE B2. Two separate VPNs are created, one for Alfa and one for Beta (VPN A and VPN B). Provider's edge devices, PE x and PE y, exchange customer routing information using MP-iBGP.

VPN A related routing information is exchanged via MP-iBGP between PE x and PE y, and further redistributed into the PE-CE protocol for Alfa's devices CE A1 and CE A2. Similarly, Beta's devices have reachability to networks known by CE B1 and CE B2 in the VPN B. There is no connectivity between Alfa and Beta, as the VPN routing information is kept within that VPN: VPN A's routes are not advertised to VPN B, and vice versa.

Routing information is kept separate in the PEs by the use of VPN Routing and Forwarding (VRF) tables. VRFs keep customer routing information specific to a customer. A single PE node can thus support multiple customer networks that are isolated from each other.

In the user plane, forwarding is based on labels in the MPLS core. Two labels are used; outer label or the LSP tunnel label is used by the MPLS core. This label is distributed by the LDP. Core routers switch the incoming packet based on this outer label. With Penultimate Hop Popping (PHP) the outer label is removed at the penultimate link, and the inner label is exposed to the egress PE.

The inner label, or the VPN label, is intended for the egress PE, and is distributed by MP-BGP. The egress PE associates the inner label to a VPN, removes the label, and forwards the IP packet out of the correct outgoing interface to the CE.

For the IP MPLS VPN application, a new address family, VPN-IPv4 address family, is defined in MP-BGP. The VPN-IPv4 address consists of 12 bytes: an 8-byte Route Distiguisher (RD) field, and a 4-byte IPv4 address. The new VPN-IPv4 addresses are unique per VPN, even if the same IPv4 address would be used. The routes learned from the CE, are exported into MP-BGP, which then distributes the routes to the PEs who need them.

Route learning is controlled by Route Target (RT) attribute. This allows defining separately which routes are installed to each of the VRFs. Each VRF has one or more RT attributes, and

each route has a set of route targets. In BGP, the route target is carried as an extended community route target. The structure is the same as for the RD.

PE assigns an MPLS label with the MP-BGP protocol, using its own address as the BGP next hop address. The address is in VPN-IPv4 format, with an RD value of 0. Traffic between PE devices then flows with the assigned label, and the label is popped at the other PE.

Resilience is supported at the LSP level by MPLS. The IGP reroutes traffic in the case of failures. Labels for another LSP may already exist due to the liberal retention mode. Additionally MPLS Traffic Engineering fast re-route may be used to forward traffic while a new path computation is ongoing.

In the service provider MPLS network, the devices interfacing the customer network, are PE devices, while other routers are P routers, which do not see customer routes. This simplifies the role of the P routers, as they can then focus on forwarding traffic based on labels, but do not need to exchange and store customer routes.

The CE-PE link is an attachment circuit. The attachment circuit can e.g. be a VLAN, and it is then guiding the selection of the VRF. A routing protocol can be used for the PE to learn the CE prefixes.

Routing protocol may be e.g. BGP, OSPF or IS-IS, or yet some other protocol. Alternatively, static routes can be used. Clearly the selection of the routing protocol needs to be agreed between the customer and the service provider. For the customer it is a benefit if the routing protocol can be the same as the one already used in the customer network. The service provider may be interested in having control of the routes that are redistributed between the PE-CE protocol and the MP-BGP.

Recovery from failures (depending on the location) is now impacted also by the convergence of the MP-BGP and the MPLS core. This topic is shortly revisited in Chapter 7 of this book.

IP MPLS VPN supports large scale deployments. The need to have a full mesh of MP-BGP sessions between the PE devices introduces a limitation. For larger deployments, a Route Reflector (RR) is used. Each PE maintains a session with the Route Reflector instead of a dedicated session with other PEs. Route Reflector is then responsible in distributing routing information between PEs. In order to avoid a single point-of-failure, RR is duplicated.

The characteristics of the application, e.g. as shown in Figure 4.31, appear to have a match with the needs of mobile backhaul:

- flexibly configurable IP layer connectivity;
- support for IPv4 and IPv6;
- scalability for large deployments;
- carrier grade resilience of links and nodes;
- possibility for further MPLS based applications: MPLS TE Fast reroute, etc.

A potential drawback is that due to the peering, customer shares routes with the service provider. So the customer network is at the IP layer not completely under the control of the customer, and recovery depends also on the provider network operation.

4.7.6 Pseudowire Emulation Edge to Edge

Pseudowire emulation architecture components are shown in Figure 4.32.

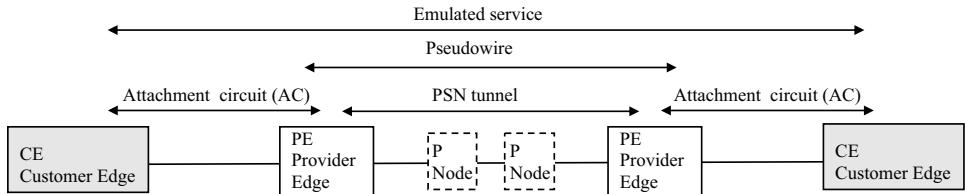


Figure 4.32 Pseudowire emulation edge to edge architecture [100].

Customer equipment, such as a BTS, connects to an MPLS Provider Edge (PE) device, with an attachment circuit to a pseudowire that is terminated at the PE router on the other end. An end-to-end emulated service between the two CEs is provided.

PSN tunnel is used to route the traffic within the MPLS network – via any P routers potentially between the two PE devices. P routers do not need to know about the pseudowire carried within the PSN tunnel. P routers only route the traffic based on the MPLS label.

A pseudowire can carry one or multiple Attachment circuits. Attachment circuits may be of packets (such as Ethernet frames), cells (ATM cells), structured or unstructured bit streams (SDH/Sonet, or narrowband E1/T1/JT1).

An application of PWE3 (Pseudowire Emulation Edge to Edge) is shown below in Figure 4.33.

The attachment circuit is an E1 from a 2G BTS. This circuit is mapped by a cell site gateway to a pseudowire (PW). The pseudowire is terminated at the BSC site using a pseudowire gateway (e.g. a router), and the E1 is delivered to the BSC. Native IP traffic from the eNodeB is carried over the same packet network. The packet network could e.g. be MPLS/IP based. At the PW gateway this traffic is forwarded towards the packet core.

At the BTS site, the PWE function is supported using a cell site gateway. Alternatively, the PW functionality could be integrated to the eNodeB. In that case, the PWE functionality that is supported by the eNodeB, is not part of the 3GPP functionality. Instead it is an additional transport feature that is integrated into the eNodeB element.

For the 2G system, the use of tunneling should not be visible in any way – which means that quality of service, availability, and other characteristics of the tunnel needs to meet 2G system Abis requirements.

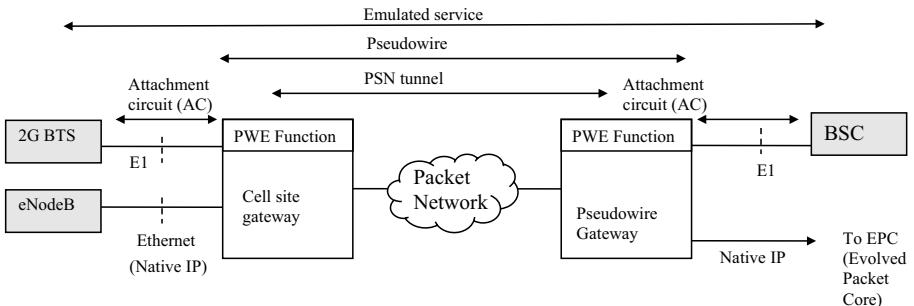


Figure 4.33 Emulating a 2G BTS over the packet network [119].

The pseudowires may be signalled with LDP or BGP (See 4.7.7). The PSN tunnel labels are allocated e.g. by LDP, guided by the IGP as discussed in 4.7.2.

PWE3 solution is an alternative because of the benefits assumed when converging into a packet switched network. Carrying native TDM traffic over a PSN requires that the PSN network supports the required availability, synchronization, security, quality of service, and other requirements. For example, a service break in the PSN is now noticeable also for 2G.

A similar PWE solution can be implemented for the ATM based node Bs. Here the pseudowire carries ATM cells over the PSNs instead of E1s as in the previous example.

Draft-martini has been published as a historical RFC 4905, encapsulation for L2 Frames over MPLS. This has been superseded by RFC4447, and Pseudowire Emulation Edge to Edge Working Group specifications.

4.7.7 MPLS L2 VPN–VPLS

MPLS L2 VPNs include a Virtual Private Wire Service (VPWS) and a Virtual Private LAN Service (VPLS). With VPLS Metro Ethernet services (E-LAN) can be implemented. The underlying technology, MPLS with LDP/BGP and IP control plane, is not visible to the user of the MEF service. Naturally mobile operator self-deployed VPLS is also an option.

VPLS defines multipoint connectivity between customer equipment (CE). The user sees the VPLS as a LAN connecting the CEs, potentially over a WAN. See Figure 4.34. A Virtual Switch Instance (VSI) implements the Ethernet bridging function in the PE device.

VPLS deployment may in general be due to multiple reasons. Some applications in the enterprise area explicitly require L2 connectivity. Similarly, if peering at the IP layer with the service provider is an issue, L2 service is one alternative.

A basic difference to the IP MPLS VPN discussed earlier, is that now the CE devices do not peer with the service provider at the IP layer. The provider is not involved in customer's IP routing. With IP MPLS VPN the customer edge (CE) device exchanges routing information

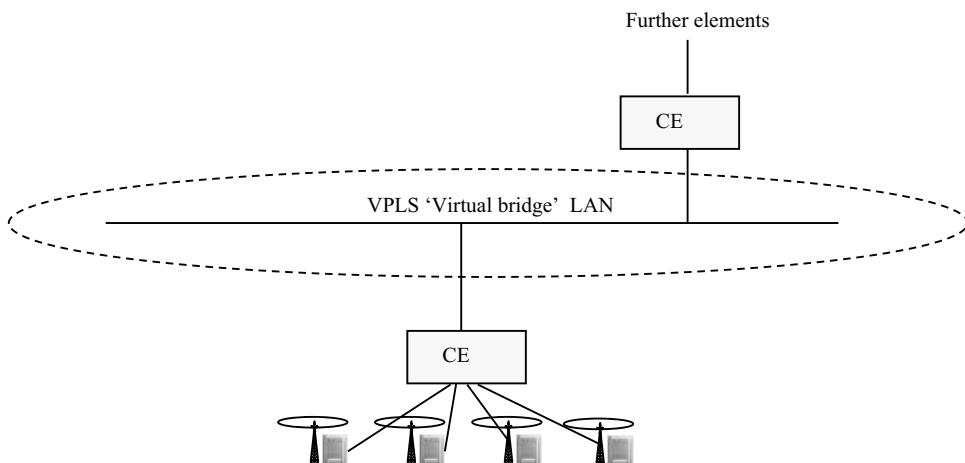


Figure 4.34 VPLS

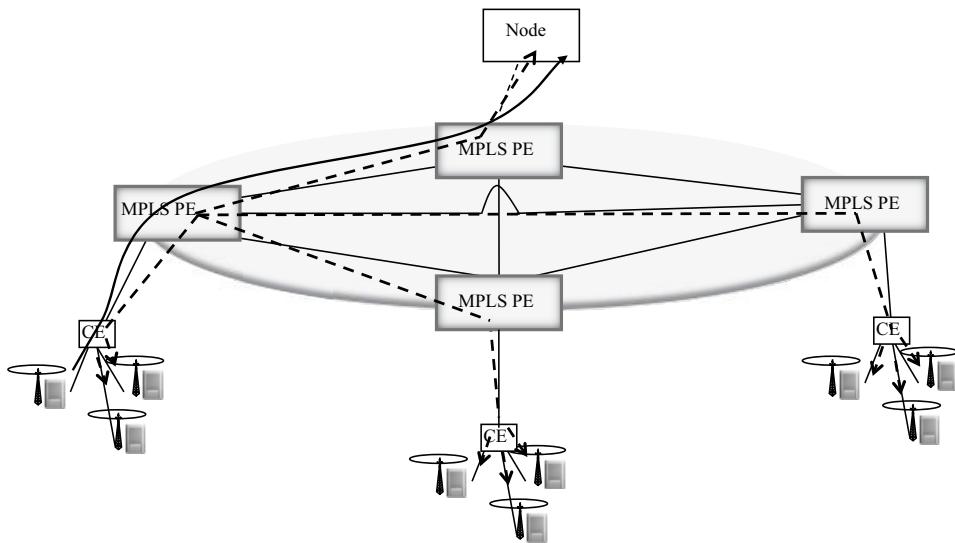


Figure 4.35 VPLS example.

with the provider. With VPLS any customer routing protocol may be used, the choice is independent of the provider. L2 VPN is also transparent to the protocol carried over it. IPv6 can be supported over VPLS.

VPLS emulates a LAN and has the characteristics of a bridge. MAC address learning and unknown unicast/broadcast flooding features are supported and a VPLS instance is also a broadcast domain. Figure 4.35 shows a case for unknown unicast flooding (dashed arrows) and then forwarding based on learned MAC address (solid arrow).

IP packets encapsulated in Ethernet frames arrive from the BTS via an attachment circuit to the MPLS PE device. The attachment circuit connects the customer frames via a pseudowire to the attachment circuit at the other end. Two (or more) labels are used. The inner label is the VC (Virtual Channel) label, identifying traffic within the tunnel. The tunnel is identified by the outer MPLS label, however in general it could also be other than MPLS-based tunnel.

Forwarding is based on MAC address learning at the edges (MPLS PE devices). Unknown unicast and broadcast frames are flooded, and customer traffic is replicated to other MPLS PEs, over pseudowires. The other PEs are not reflecting traffic back to other pseudowires but only towards the CE (split-horizon rule). Pseudowires travel within a PSN (Packet Switched Network) tunnel (an LSP in Figure 4.36). Each PE logically sees a tree topology to every destination and thus there are no loops. Split-horizon rule, together with the full-mesh topology, ensures that there is no loop. Due to these rules, there is no need to run spanning tree in the MPLS core.

When the destination MAC address has been learned at the PE device, traffic can be forwarded accordingly only to the correct pseudowire and further on to the destination (marked as Node). This is indicated by a solid arrow in Figure 4.35.

With the MPLS/IP core, VPLS can extend the L2 service over a larger distance. Physical distance is not limiting the scale. Since nodes share the broadcast domain per VPLS instance, failures and misconfigurations on one site may impact other sites. If for some reason an

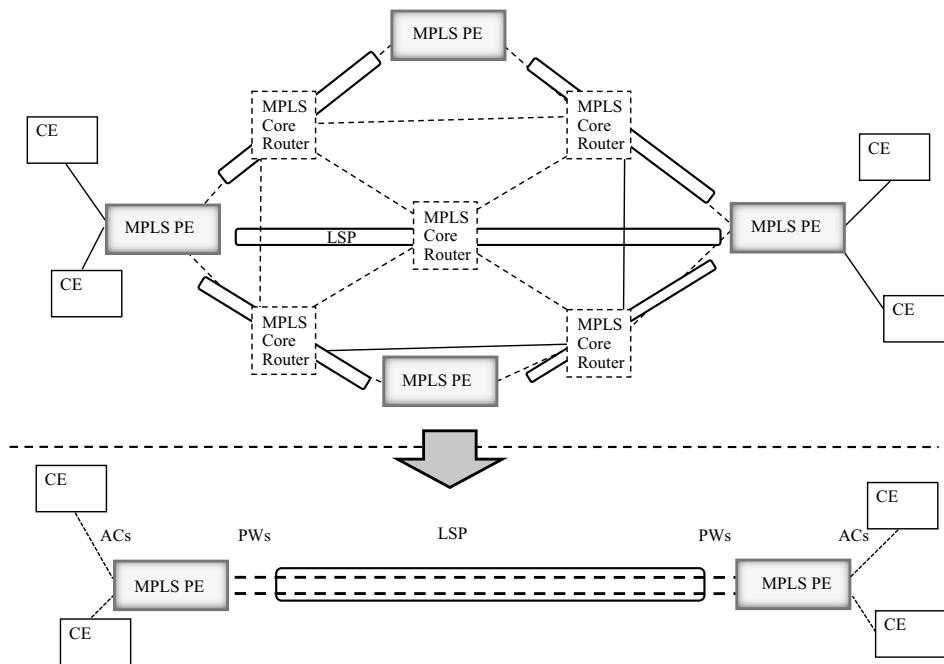


Figure 4.36 LSPs, pseudowires, and attachment circuits.

excessive amount of broadcast or unknown unicast frames is sent, this may cause a disruption of service on all sites. Similarly, if on one site accidentally a L2 loop is created, all sites may be down until the situation is corrected.

The resilience in the provider's MPLS/IP network is achieved with MPLS and IP features. IP control plane (some IGP) operates within the MPLS domain.

In the VPLS implementation, Ethernet frames from the customer are mapped to pseudowires, which are further carried in the LSPs. Each MPLS PE router connects to every other MPLS PE router with an LSP carrying pseudowires, creating a full-mesh of pseudowires.

For pseudowire signalling, either LDP or BGP can be used. RFC 4762 defines the use of LDP. With LDP, targeted LDP sessions are established (a full mesh) between PEs. VC label is then assigned and communicated via LDP. RFC 4761 specifies the use of BGP for auto-discovery and signalling.

Note that a Virtual Private Wire Service (VPWS) uses similarly an inner label (VC label) for pseudowire and an outer label for the PSN tunnel, however supporting point-to-point connection instead of multipoint. Because of the point-to-point nature, there is no need for VSIs. VPWS implements an E-Line type of service.

In larger VPLS networks, an issue comes up because of the full-mesh nature. Having N PE nodes, requires for $N \times (N-1)$ pseudowires to the Virtual Switch Instances on the PEs. A similar amount of control plane LDP sessions is needed, as in the control plane a full mesh of LDP sessions are needed to set up the pseudowires.

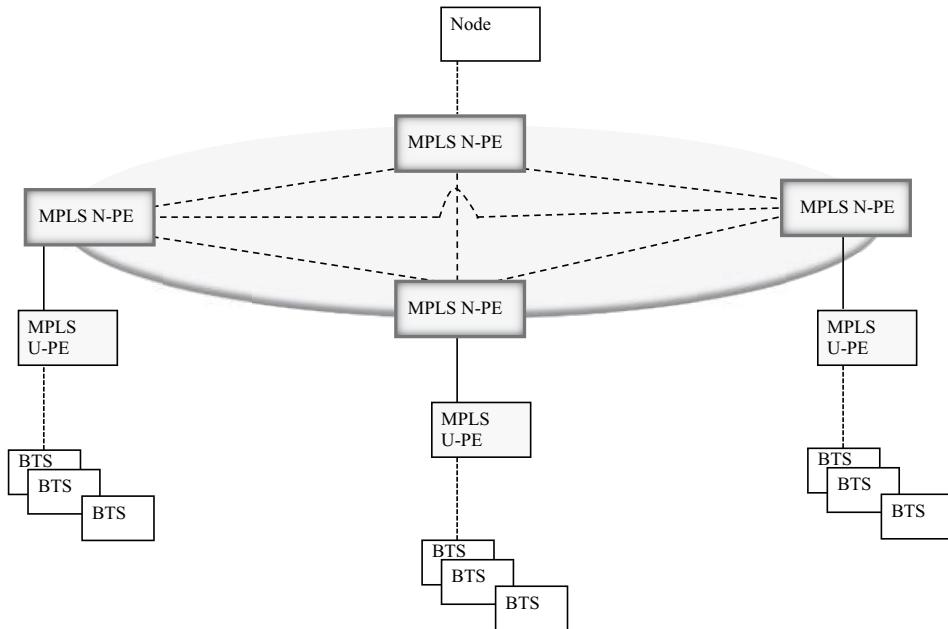


Figure 4.37 H-VPLS.

Scalability can be increased by Hierarchical-VPLS (H-VPLS). With H-VPLS, the topology is changed so that while the core is still full-mesh, each core PE node has a hub-and-spoke topology towards the access.

In Figure 4.37, the VPLS domain is extended by tunnelling technique further into the access tier. MPLS N-PE (Network-PE) connects to the U-PE (User-PE). A single pseudowire (per VPLS instance) is set up between the N-PE and the U-PE.

Another approach is to use Ethernet as the access method. A tag is added to indicate a Provider's VLAN (P-VLAN). Each P-VLAN then maps to a VPLS instance.

For the mobile backhaul application there is no explicit need for connectivity at L2. The application (radio network protocols) are mapped into IP and need IP layer connectivity. Use of any L2 is possible - including of course VPWS or VPLS service.

In general, VPLS-type of multipoint connectivity is not needed. Typical topology is hub-and-spoke with a controller (RNC or BSC) or a GW (SGW, MME or IPSec GW in case of LTE) as the hub. The leaves would only need to communicate in case of direct X2 (LTE). For a potential direct X2 implementation with VPLS, one topic is the size of the broadcast domain, and another one is the allocation of eNodeBs to different VLANs (VPLS instances) with respect to the radio network topology.

4.7.8 MPLS-TE

Traffic Engineering (TE) delivers tools for managing network resources effectively. RFC2702 MPLS Traffic Engineering, documents the goals of Internet Traffic Engineering as to

'facilitate efficient and reliable network operations while simultaneously optimizing network resource utilization and traffic performance.' Traffic engineering can enhance QoS, in terms of packet loss, delay, throughput, and enforcement of Service Level Agreements (SLAs). Resources can be utilized more efficiently, by a more equal distribution of traffic into paths, so that there are less over- or underutilized links.

IP routing with OSPF or other interior gateway protocols mainly uses one metric (cost). This metric is used to calculate the shortest path to the destination. Traffic load, capacity, or congestion of the network, are not taken into account. Multiple traffic streams may be directed by the Shortest Path First (SPF) algorithm to the same link, causing congestion. These shortcomings can be addressed by native IP in varying extent with load sharing and policy-based routing. MPLS TE creates a virtual topology, based on Label Switched Paths (LSPs), and traffic can be steered through these paths according to selected criteria.

A fast protection scheme can also be realized with the traffic engineered LSPs. MPLS TE Fast Reroute (FRR) feature allows traffic to be carried via a pre-provisioned protecting LSP. Two methods are defined: a one-to-one backup and a facility backup. If the detection is fast enough, FRR can reach 50 ms restoration time.

MPLS-TE requires a link state protocol, such as OSPF or IS-IS routing protocol, and related traffic engineering extensions to those protocols. The extensions carry further information than simply the cost to the routers.

An MPLS LSP path is calculated based on the information collected from the network, and if a feasible path can be found, the LSP is set up with RSVP-TE (Resource Reservation Protocol - Traffic Engineering) protocol. RSVP protocol has been enhanced for the MPLS TE application, to include label allocation and further features.

RSVP TE Path message signals the path to the MPLS routers, and RSVP-TE Resv (Reserve) message allocates the labels. RSVP Path message, with a label request, is sent from the ingress node towards the downstream nodes, and RSVP Resv message is sent back from the egress node, with the allocated label information.

MPLS TE LSPs are unidirectional. If an MPLS TE LSP is wished for the return traffic, it needs to be allocated separately.

Explicit routing is supported with the Path message Explicit_route object. Explicit route may be dynamically calculated based on the QoS requirements. When the explicit routing is used, the path can be controlled by the ingress node which is initiating the Path message. The path computation itself may be done by the ingress node or by some other element. Another option is to statically configure the explicit path through the network. Additionally RSVP supports bandwidth reservation.

RSVP is originally a building block for Integrated Services (IntServ) framework, which includes a way to communicate applications QoS requirements to the network elements, which then can control the QoS accordingly. RSVP-TE as a control plane protocol for the MPLS LSP tunnels has also been generalized to support other than MPLS applications. This extension is defined in RFC 3473, Generalized Multi-Protocol Label Switching (GMPLS) RSVP-TE extensions. The control plane is referred to as G-MPLS. This generalized control plane can cover e.g. Sonet/SDH, wavelengths, and spatial switching (e.g. switching an incoming port to outgoing port).

MPLS-TE can be used with other MPLS applications such as pseudowire emulation and MPLS L2 and L3 VPNs for the PSN tunnels. In this case, it imposes the label to the MPLS label stack with RSVP-TE.

With the mobile backhaul, potential applications of MPLS-TE are in the transport network aggregation/core domain, for steering traffic, guaranteeing QoS, and for fast protection switching (Fast re-route, FRR). If the mobile operator has deployed MPLS, these features are potentially of importance. If the mobile operator has no MPLS, and is relying on a service of a transport service provider, MPLS TE features are not directly visible. In the service provider network, MPLS TE may be used to deliver the transport service, even though the technology itself is not visible to the service user.

4.7.9 *MPLS-TP*

Traditional TDM based Sonet/SDH provides for a transport network that is statically provisioned (by a Network Management System) and includes operations, administrations, and maintenance functionality. MPLS-Transport Profile (MPLS-TP) complies with standard MPLS/IP in its main aspects, however, it is also connection-oriented, includes OAM features, and supports fast protection switching. Essentially it bridges Sonet/SDH with MPLS/IP, importing characteristics from both. A control plane protocol is optional so MPLS-TP can be managed statically via an NMS in a way similar to Sonet/SDH. However it may also rely on GMPLS, meaning RSVP-TE and a link state routing protocol with TE extensions. User traffic may be native IP or pseudowires.

Standardization of MPLS-TP is ongoing in the IETF. ITU-T was involved in defining T-MPLS (Transport-MPLS), however after mutual agreement standardization continues within the IETF working groups for MPLS-TP.

With MPLS-TP, OAM travels in-band, together with the user signal in LSPs or pseudowires. OAM can be fault detection, diagnostics, maintenance or other functions. Specifically for MPLS-TP, new functions like Performance Monitoring, Automatic Protection Switching and management and signalling channels have been considered.

In RFC 5586, the pseudowire Associated Channel Header (ACH) is generalized to include MPLS-TP LSPs and MPLS sections, in addition to MPLS pseudowires. One of MPLS reserved label values, label 13, is allocated as a general associated label (GAL) for the purpose of identifying the G-ACH. Previously RFC3032 defined label value 14 as the OAM alert label, for MPLS-TP GAL, value 13 is however used.

4.8 Summary

Transport can be modeled as a service to the radio network: consisting of a mobile network element integrated transport and of a backhaul service. Requirements for the backhaul service originate from the end user services and from the radio network layer functionality. For the backhaul, all traffic types need to be considered: user plane, control plane, O&M, synchronization and transport layer control protocols. Different radio network technology base stations interface the backhaul network with different backhaul protocols: TDM, ATM or IP.

Mobile networks need high capacity access lines for the base stations in order to deliver the data rates supported on the air interface. Flexibility and versatility is needed in the aggregation tier as the services needed are specific to the radio technology used in the base station. MPLS was discussed as one alternative for supporting both the native IP services as well as the pseudowire emulation services for the legacy TDM and ATM interfaces.

Generally in networking protocols operate in local area (LAN) or in Wide Area (WAN). Each base station needs a Metro or Wide area connectivity with the peer element – a controller, a GW or a control plane entity. For the wide area connectivity, high availability, security and QoS are essential. A carrier grade service is needed.

With Ethernet port in the base stations high capacities can be supported with a single physical port. The service delivering the high capacity on the Ethernet port may be implemented e.g. with a microwave radio, a point-to-point Ethernet, next-generation Sonet/SDH, DSL, Fibre/xWDM or IP/MPLS. At the same time, capacity in the TDM networks can be used whenever existing by mapping IP into the TDM structures (PPP over E1/T1s).

References

Chapter 4.1

- [1] Froom, Sivasubramanian, Frahim: Building Cisco Multilayer Switched Networks, 4th Edition. Cisco Press, 2007.

Chapter 4.2

- [2] http://www.ieee.org/index.html?WT.mc_id=hpf_logo, retrieved August 2011
- [3] IETF RFC 2026, BCP9 The Internet Standards Process, Revision 3
- [4] IETF RFC 2029, BCP11 The Organizations Involved in the IETF Standards Process
- [5] IETF RFC 6410 Reducing the Standards Track to Two Maturity Levels
- [6] <http://www.ietf.org/>, retrieved August 2011
- [7] <http://www.iso.org/iso/home.html>, retrieved August 2011
- [8] <http://www.itu.int/en/Pages/default.aspx>, retrieved August 2011
- [9] ITU-T, Resolution 1 – Rules of procedure of the ITU Telecommunication Standardization Sector (ITU-T), October 2008
- [10] <http://metroethernetforum.org/index.php>, retrieved August 2011, October 2011.
- [11] <http://www.broadband-forum.org/index.php>, retrieved August 2011

Chapter 4.3

- [12] IEEE 802.3-2008 IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks. Specific requirements. Part 3: Carrier sense multiple access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications
- [13] ANSI T1.403 Network-to-Customer Installation - DS1 Metallic Interface specification
- [14] ANSI T1.408 Integrated Services Digital Network (ISDN). Primary Rate - Customer Installation Metallic Interfaces. Layer 1 Specification
- [15] ITU-T G.703 Physical/electrical characteristics of hierarchical digital interfaces
- [16] ITU-T G.704 Synchronous frame structures used at 1544,6312, 2048, 8448 and 44 736 kbit/s hierarchical levels
- [17] ITU-T G.775 Loss of Signal (LOS), Alarm Indication Signal (AIS) and Remote Defect Indication (RDI) defect detection and clearance criteria for PDH signals
- [18] ITU-T G.707 Network node interface for the synchronous digital hierarchy (SDH)
- [19] ITU-T G.957 Optical interfaces for equipments and systems relating to the synchronous digital hierarchy
- [20] ITU-T G.652 Characteristics of a single-mode optical fibre and cable
- [21] ITU-T G.653 Characteristics of a dispersion-shifted single-mode optical fibre and cable
- [22] ITU-T G.654 Characteristics of a cut-off shifted single-mode optical fibre and cable
- [23] ITU-T G.655 Characteristics of a non-zero dispersion-shifted single-mode optical fibre and cable

Chapter 4.4

- [24] IETF RFC 1661, STD 51 The Point-to-Point Protocol (PPP)
- [25] IETF RFC 1662 PPP in HDLC-like Framing
- [26] IETF RFC 1334 PPP Authentication Protocols (Obsoleted by RFC 1994)
- [27] IETF RFC 1994 PPP Challenge Handshake Authentication Protocol (CHAP)
- [28] IETF RFC 1990 The PPP Multilink Protocol (MP).
- [29] IETF RFC 2686 The Multi-Class Extension to Multi-Link PPP
- [30] IETF RFC 1332 IP v4 in IPCP
- [31] IETF RFC 1144 TCP/IP Compression for Low-Speed Serial Links
- [32] IETF RFC 3544 IP Header Compression over PPP
- [33] IETF RFC 3241 Robust Header Compression (ROHC) over PPP
- [34] IETF RFC 3095 Robust Header Compression (ROHC):Framework and four profiles: RTP, UDP, ESP, and uncompressed
- [35] IETF RFC 4815 Robust Header Compression (ROHC): Corrections and Clarifications to RFC 3095
- [36] IETF RFC 3843 Robust Header Compression (ROHC): A Compression Profile for IP
- [37] IETF RFC 2507 IP Header Compression
- [38] IETF RFC 2508 Compressing IP/UDP/RTP Headers for Low-Speed Serial Links
- [39] IETF RFC 3545 Enhanced Compressed RTP (CRTP) for Links with High Delay, Packet Loss and Reordering
- [40] IETF RFC 3153 PPP Multiplexing
- [41] IETF RFC 2615 PPP over SONET/SDH

Chapter 4.5

- [42] www.metroethernetforum.org, retrieved October 2011
- [43] IEEE 802.1D-2004 IEEE Standard for Local and metropolitan area networks, Media Access Control (MAC) Bridges
- [44] IEEE 802.1AX-2008 IEEE Standard for Local and metropolitan area networks, Link Aggregation
- [45] IEEE 802.1Q-2005 IEEE Standard for Local and metropolitan area networks, Virtual Bridged Local Area Networks
- [46] IEEE 802.1ah-2008 IEEE Standard for Local and metropolitan area networks, Virtual Bridged Local Area Networks. Amendment 7: Provider Backbone Bridges
- [47] IEEE 802.1ag-2007 IEEE Standard for Local and metropolitan area networks, Virtual Bridged Local Area Networks. Amendment 5: Connectivity Fault Management
- [48] ITU-T Y.1731 OAM functions and mechanisms for Ethernet based networks
- [49] IEEE 802.1ad-2005 IEEE Standard for Local and metropolitan area networks Virtual Bridged Local Area Networks. Amendment 4: Provider Bridges
- [50] IETF RFC 4026 Provider Provisioned Virtual Private Network (VPN) Terminology
- [51] IETF RFC 4664 Framework for Layer 2 Virtual Private Networks (L2VPNs)
- [52] IETF RFC 4665 Service Requirements for Layer 2 Provider-Provisioned Virtual Private Networks
- [53] IETF RFC 4761 Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling
- [54] IETF RFC 4762 Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling

Chapter 4.6

- [55] IETF RFC 791, STD 5, Internet Protocol (IP)
- [56] IETF RFC 2460 Internet Protocol, Version 6 (IPv6) Specification
- [57] IETF RFC 793, STD 7, Transmission Control Protocol
- [58] IETF RFC 768, STD 6, User Datagram Protocol (UDP)
- [59] IETF RFC 1812 Requirements for IP Version 4 Routers
- [60] IETF RFC 1122, STD 3, Requirements for Internet Hosts - Communication Layers

- [61] Comer: Internetworking With TCP/IP Volume 1: Principles Protocols, and Architecture. 5th edition. Prentice-Hall, 2006.
- [62] Stevens: TCP/IP Illustrated, Volume 1: The Protocols. Addison-Wesley, 1994,
- [63] Teare, Paquet: Building Scalable Cisco Internetworks. Third edition. Cisco Press, 2007.
- [64] 3GPP TR25.933 IP transport in UTRAN (Release 5), v 5.4.0
- [65] 3GPP TS 25.426 UTRAN Iur and Iub interface data transport & transport signalling for DCH data streams, v10.1.0
- [66] 3GPP TS 25.434 UTRAN Iub interface data transport and transport signalling for Common Transport Channel data streams, v10.1.0
- [67] 3GPP TS 25.414 UTRAN Iu interface data transport and transport signalling, v10.1.0
- [68] 3GPP TS29.281 General Packet Radio System (GPRS) Tunnelling Protocol User Plane (GTPv1-U), v10.3.0
- [69] 3GPP TS 36.412 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); S1 signalling transport, v10.1.0
- [70] 3GPP TS 36.414 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); S1 data transport, v10.1.0
- [71] 3GPP TS 36.422 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 signalling transport, v10.1.0
- [72] 3GPP TS 36.424 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 data transport, v10.1.0
- [73] IETF RFC 1918, BCP 5, Address Allocation for Private Internets
- [74] IETF RFC 4632, BCP122, Classless Inter-domain Routing (CIDR): the Internet Address Assignment and Aggregation Plan
- [75] IETF RFC 3021 Using 31-Bit Prefixes on IPv4 Point-to-Point Links
- [76] IETF RFC 2131 Dynamic Host Configuration Protocol
- [77] IETF RFC 2132 DHCP Options and BOOTP Vendor Extensions
- [78] IETF RFC 3046 DHCP Relay Agent Information Option
- [79] IETF RFC 2991 Multipath Issues in Unicast and Multicast Next-Hop Selection
- [80] IETF RFC 2992 Analysis of an Equal-Cost Multi-Path Algorithm
- [81] Huitema: Routing in the Internet. 2nd Edition. Prentice-Hall, 1999.
- [82] IETF RFC 2328, STD 54, OSPF Version 2
- [83] IETF RFC 5340 OSPF for IPv6
- [84] IETF RFC 5838 Support of Address Families in OSPFv3
- [85] IETF RFC 1142 OSI IS-IS Intra-domain Routing Protocol (Reprinted ISO 10589: ‘Intermediate System to Intermediate System intradomain routeing information exchange protocol for use in conjunction with the protocol for providing the connectionlessmode network service (ISO 8473)’ ISO/IEC 10589:2002)
- [86] IETF RFC 2453, STD 56, RIP Version 2
- [87] IETF RFC 4271 A Border Gateway Protocol-4 (BGP-4)
- [88] IETF RFC 2475 An Architecture for Differentiated Services
- [89] IETF RFC 2474 Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers
- [90] IETF RFC 3168 The Addition of ECN to IP
- [91] IETF RFC 1633 Integrated Services in the Internet Architecture: an Overview
- [92] Wang: Internet QoS. Architectures and Mechanisms for Quality of Service. Morgan Kaufmann Publishers, 2001
- [93] IETF RFC 3260 New Terminology and Clarifications for Diffserv (Informational)
- [94] IETF RFC 826, STD 37, Address Resolution Protocol (ARP)
- [95] IETF RFC 792, STD 5, Internet Control Message Protocol (ICMP)
- [96] IETF RFC 3550, STD 64, RTP: A Transport Protocol for Real-Time Applications
- [97] IETF RFC 4960 Stream Control Transmission Protocol
- [98] Stewart, Xie: Stream Control Transmission Protocol (SCTP). A Reference Guide. Addison-Wesley, 2002
- [99] IETF RFC 4443 Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification

Chapter 4.7

- [100] IETF RFC 3031 MPLS Architecture
- [101] IETF RFC 3032 MPLS Label Stack Encoding

- [102] IETF RFC 3443 Time To Live (TTL) Processing in Multi-Protocol Label Switching (MPLS) Networks
- [103] IETF RFC 3270 MPLS Support of Differentiated Services
- [104] IETF RFC 5462 Multiprotocol Label Switching (MPLS) Label Stack Entry: ‘EXP’ Field Renamed to ‘Traffic Class’ Field
- [105] De Ghein: MPLS Fundamentals. Cisco Press, 2006
- [106] Minei, Lucek: MPLS Enabled Applications. Second Edition. Wiley, 2008
- [107] Guichard, Le Faucheur, Vassuer: Definitive MPLS Network Designs. Cisco Press, 2005.
- [108] IETF RFC 5036 LDP Specification
- [109] IETF RFC 4760 Multiprotocol Extensions for BGP-4
- [110] IETF RFC 2918 Route Refresh Capability for BGP-4
- [111] IETF RFC 4364 BGP/MPLS IP Virtual Private Networks (VPNs) (Obsoletes RFC 2547)
- [112] IETF RFC 4456 BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)
- [113] IETF RFC 5065 Autonomous System Confederations for BGP
- [114] IETF RFC 4026 Provider Provisioned Virtual Private Network (VPN) Terminology
- [115] IETF RFC 4379 Detecting MPLS Data Plane Failures
- [116] IETF RFC 4576 Using a Link State Advertisement (LSA) Options Bit to Prevent Looping in BGP/MPLS IP Virtual Private Networks (VPNs)
- [117] IETF RFC 4577 OSPF as the Provider/Customer Edge Protocol for BGP/MPLS IP Virtual Private Networks (VPNs)
- [118] IETF RFC 3916 Requirements for Pseudo-Wire Emulation Edge-to-Edge (PWE3)
- [119] IETF RFC 3985 Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture
- [120] IETF RFC 4447 PWE3 Using LDP
- [121] IETF RFC 4448 Encapsulation Methods for Transport of Ethernet over MPLS Networks
- [122] IETF RFC 4553 Structure-Agnostic TDM over Packet (SAToP)
- [123] IETF RFC 4717 Encapsulation for ATM over MPLS
- [124] IETF RFC 5086 Structure-Aware Time Division Multiplexed (TDM) Circuit Emulation Service over Packet Switched Network (CESoPSN)
- [125] IETF RFC 4664 Framework for Layer 2 Virtual Private Networks (L2VPNs)
- [126] IETF RFC 4448 Encapsulation Methods for Transport of Ethernet over MPLS Networks
- [127] IETF RFC 3209 Extensions to RSVP for LSP Tunnels
- [128] IETF RFC 3473 Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions
- [129] IETF RFC 4090 Fast Reroute Extensions to RSVP-TE for LSP Tunnels
- [130] IETF RFC 2702 Requirements for Traffic Engineering Over MPLS
- [131] IETF RFC 5305 IS-IS Extensions for Traffic Engineering
- [132] IETF RFC 3630 Traffic Engineering (TE) Extensions to OSPF Version 2
- [133] IETF RFC 5317 JWT report on MPLS architectural considerations
- [134] IETF RRC 5586 MPLS Generic Associated Channel
- [135] IETF RFC 5654 MPLS-TP requirements
- [136] IETF RFC 5718 An In-band data communication network for the MPLS Transport Profile Virtual Circuit Connectivity Verification (VCCV)
- [137] IETF RFC 5085 Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires
- [138] IETF RFC 4385 Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN
- [139] IETF RFC 3429 Assignment of the ‘OAM Alert Label’ for Multiprotocol Label Switching Architecture (MPLS) Operation and Maintenance (OAM) Functions (informational)

5

Backhaul Transport Technologies

Jouko Kapanen, Jyri Putkonen and Juha Salmelin

Mobile backhaul (MBH) is a transport network connecting mobile base stations to radio network controllers and its gateways and servers. Its building blocks are often the same as in fixed access and transport networks, especially in the backbone and aggregation tiers. Also in the access tier similar systems and equipment as in fixed access are used, when they suit the backhaul requirements. Convergence of fixed and mobile networks is a reality in backhaul. Typically wire-line MBH access uses equipment developed in the first place for fixed access and adapted for MBH use. In wireless access, i.e. in microwave radio solutions, the mobile backhaul is, however, the main market and is thus driving MWR developments – this means that MWR solutions are often genuinely made for MBH use and need less adaptations.

In this chapter we first look at some generic characteristics of transport systems and networks, mainly for the benefit of radio network specialists. Then we discuss typical systems used in MBH networks, starting from wireless access, and continuing with wire-line access systems (both copper line and optical systems), and then, briefly, systems used in upper tiers of MBH networks. The chapter is rounded off by considering other types of building blocks for MBH solutions, namely leased line and other solutions based on services provided by another (telecom) operator and shared with other users.

5.1 Transport Systems

5.1.1 OSI-Model

In networking, and nowadays also in all transport, discussions often start with considering solutions in different layers, for example a Layer 2 solution for access or Layer 3 solutions for some part of backbone. These layers refer directly, or approximately, to the layer model developed originally for data transport.

Open System Interworking (OSI) model was formulated in International Standardization Organization (ISO) decades ago (Table 5.1). It was developed to define how data moves in open systems and it is used widely to abstract telecommunication/data communication systems. Systems are split into seven layers, each of which communicates with one above or below, and with similar layers elsewhere in the system. Functionalities of all equipment along the data path are mapped to a layer. The protocol standards that the ISO and other standards organizations like IEEE develop are associated with these layers.

One may argue that the original OSI model is today out of date and some protocols are hard to locate in a specific layer. It is, however, still a valid tool in conceptualization of networks. For example Ethernet switching is considered as L2 (Layer 2) functionality while IP routing takes place in L3. Some protocols like IP/MPLS fall in between two layers.

Also in MBH networks the roles of different layers are commonly discussed as well as pros and cons of locating certain functions in a specific layer.

5.1.2 Access Schemes

Frequency division duplex (FDD) is a way of communicating by separating transmission directions (uplink, downlink) in frequency domain. Time division duplex (TDD) divides transmissions in time, but so quickly from end-user point of view that connection can be called duplex rather than simplex.

Duplex refers to the communication that takes place over a channel to both directions at the same time. Simplex communication also takes place in both directions; not at the same time but one direction at a time. Sometimes broadcast type of one-way communication is also referred to as simplex and one-way at the time communication as half-duplex.

Table 5.1 ISO OSI-layers [25].

Layer	Data unit	Function	Example
L7	Application	Data	User interface, displaying data HTTP, Firefox, Angry Birds, Network Manager
L6	Presentation	Translating data between formats in different applications (computers)	.wav to .mp3, EBCDIC to ASCII
L5	Session	Manage and organise multiple connections per application	Web-browser, Telnet
L4	Transport	Segment	Making data streams, data recovery, retransmission IP tunneling, TCP, UDP
L3	Network	Packet, Datagram	Transporting data in datagrams, managing connectivity, routing Internet Protocol, IGMP, X.25, RRC, Q.931 (ISDN)
L2	Data Link	Frame	Arranges bits to fit physical circuitry, flow control, bit-error detection MAC, Ethernet, HDLC, ATM, LLC
L1	Physical	Bit	Physical and electrical specifications, transmission media IEEE 802.3, FDDI, QPSK-modulation, G.703, SDH

Multiple Access is a method by which several users can share the capacity of a transmission channel. In Time Division Multiple Access, TDMA, multiple users share the same channel so that each user gets a short time frame at a time. Multiplexing takes place so quickly that each traffic seems to be continuous. This applies especially for digital transmission where data can be split into discontinuous stream of packets or frames. Higher OSI-layer protocols then take care to combine data for application needs. Nowadays TDMA is mainly used for digital transmission.

Frequency Division Multiple Access (FDMA) splits traffic from different users or services in frequency domain. Time-wise all streams can take place simultaneously, but traffic can be based on packets. This method is used by the majority of legacy microwave radio links and some optical fiber transmission systems.

Code Division Multiple Access (CDMA) is best known from WCDMA radio interface. All users communicate at the same time in the same radio frequency, but transmissions are separated by different orthogonal codes.

One important categorization is connection-oriented versus connectionless. The latter means communicating between two nodes without prior agreement or signaling. An example of the former is Plesiochronous Digital Hierarchy (PDH), described later, and of the latter Internet Protocol (IP). PDH can also be called circuit switching, because the communicating nodes form a channel (circuit), usually at Layer 1, before communication starts. Connections in IP are packet switched. Packets are transmitted independently and the same L1 channels may carry packets from many sources and to many destinations. There are lots of intermediate protocols that can be called virtual circuit switching. ATM is an example of protocol that uses packets (cells) but establishes a connection.

5.1.3 *Plesiochronous Digital Hierarchy (PDH)*

Digital multiplexing is used to connect digital signals to the desired transmission segments and to the end user. Almost all digital transmission networks are based on standardized digital hierarchies.

Plesiochronous digital hierarchy is a TDM technology used most commonly in access and aggregation transport networks. Plesiochronous means ‘almost synchronous’. Different data streams in a large PDH network have the same nominal rate but are not exactly synchronous. [1]

PDH is the legacy of practically all telecommunications networks today. PDH basic unit is one digitized voice channel. Using pulse code modulation technique (PCM) analog signal is digitized by 8 bits taking 8000 samples per second. The bits are encoded to line signal of 64 kbit/s with HDB3 coding. These voice channels are seldom transmitted alone over long hauls. They are byte-by-byte multiplexed (i.e. TDM) together forming the primary rate of PDH system. In Europe (CEPT/ETSI-area) primary rate is called E1 having a rate of 2048 kbit/s and in Japan J1 and USA T1 having a rate of 1544 kbit/s, see Figure 5.1. E1 contains 30 basic 64 kbit/s channels and two time slots e.g. for signaling, frame alignment and alarm indication. The length of every frame is then 32×8 bits = 256 bits. A whole E1 frame must be sent at the same rate as the basic signal (Level 0) is sampled which sets the frame length at 125 µs. T1 (DS1) frame contains 24 basic channels and every frame has one extra bit for framing and synchronization.

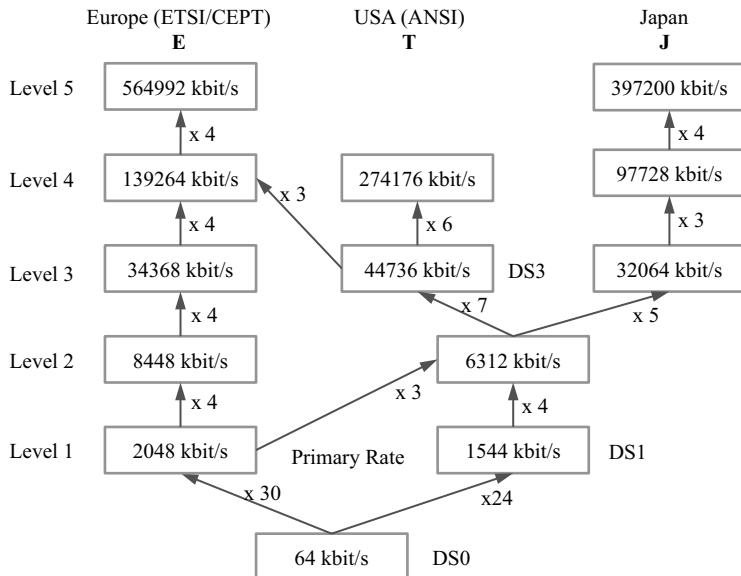


Figure 5.1 PDH bit rate hierarchy in different standards.

Higher order PDH rates are generated by multiplexing lower level streams. Next rate is slightly higher than the product of input bit rate and number of channels. Stuffing or positive justification is needed because streams (bits) do not arrive to a PDH network element exactly at the same moment and there may a slight shift between frames, too. Sub-frames must also be transported transparently without changing their plesiochronous clocks.

One of the advantages of PDH is that continuous bit stream carries synchronization not only to the transmission nodes but also to the mobile base stations. The data rate is controlled by a clock in the equipment generating the data. The free-running rate is allowed to vary by ± 50 ppm of 2.048 Mbit/s which by itself is not good enough for BTS synchronization. More about synchronization in Chapter 6.

One disadvantage of TDM systems is the fixed frame structure. Unlike IP and Ethernet packet networks that utilize statistical multiplexing PDH reserves the pre-configured channels regardless of the need. Especially in mobile data traffic where geographical location, users and services (sources) change continuously this causes congestion or waste of capacity. More about the benefits of packet backhauling can be found in Chapter 4.

Cross-connection equipment is used in older 2G systems down to 8 kbit/s level to improve transmission efficiency when 2 Mbit/s frames are only partially filled.

5.1.4 Synchronous Digital Hierarchy (SDH)

SDH is a mechanism that allows PDH rates to be transported synchronously and digital hierarchy to be extended to higher rates. SDH is the international version of the standard published by the ITU while SONET is the United States version. The standards are very similar in implementation, making it easy to interoperate between SDH and SONET at any rate.

Advantages of SDH transmission compared to PDH:

- High enough transmission rates for current mobile core data network needs.
- Synchronous network provides accurate timing.
- More simple add/drop functionality compared to PDH.
- Reliable in node and network level.
- Global standardization and interconnectivity.
- Support of many client protocols (with NG-SDH), future proof.

The basic transmission format for SDH is STM-1 (Synchronous Transport Module, Level 1). Each SDH rate is an exact multiple of the lower level signal. STM-1 frame is transmitted in $125\ \mu\text{s}$ time period. The corresponding signal in SONET is STS-3c (Synchronous Transport Signal, Level 3, concatenated) or OC3c (Optical Carrier, Level 3, concatenated). Basic rate in SONET is called STS-1/OC-1, 51.840/50.112 Mbit/s, that is sometimes called also STM-0. Overhead is 3.4% of data rate.

There are three ways of mapping 1544 and 2048 kbit/s primary rate signals into the VC-11 and VC-12, respectively, as defined in ITU-T Recommendation G.707: asynchronous, bit synchronous and byte synchronous.

Incoming PDH signal in Figure 5.2 is synchronized to SDH rate and stuffing bits are used. Pointers are used to adjust varying PDH rates to SDH rate. In SDH multiplexing stuffing is not needed. PDH rates are mapped to STM frames by fixed size containers (C). To carry a container over a SDH path overhead is needed for alarm and management information. If the stream is slower than the rest of the container is stuffed with extra bits. This filling is called mapping. Path Overhead (POH) is used to define a path between edges of SDH network. A container together with POH forms a virtual container (VC) that travels over the SDH network to the point where (PDH) data is extracted. Administrative Unit (AU) is the next step in SDH framing hierarchy when AU-pointer is added to the frame. AU-pointer indicates the location of the first VC byte and allows the VC ‘float’ inside STM-frame. This turns VC into AU. Tributary Units (TU) are used to multiplex lower order PDH signals to VC-4. Each TU-n has a corresponding VC-n that is then mapped to VC-4 with the help of TU-n pointer. Administrative Unit Group (AUG) is used to multiplex a STM-signal either from the same level AU or from several lower level AUs.

Here ‘c’ refers to concatenated or clear channel. This implies that the entire payload rate may be used by a single data stream. It is not tied to E- or T-rates. The rest of the transmission

Table 5.2 STM, SONET and Optical Carrier data rates.

SDH signal	SONET signal	Optical Level	Line rate [Mbit/s]	Data rate [Mbit/s]
STM-256	STS-768	OC-768	39813.120	38486.016
STM-64	STS-192	OC-192	9953.280	9621.504
STM-16	STS-48	OC-48	2488.320	2405.376
STM-4	STS-12	OC-12	622.080	601.344
STM-1 (STM-0)	STS-3 STS-1	OC-3 OC-1	155.520 51.840	150.336 50.112

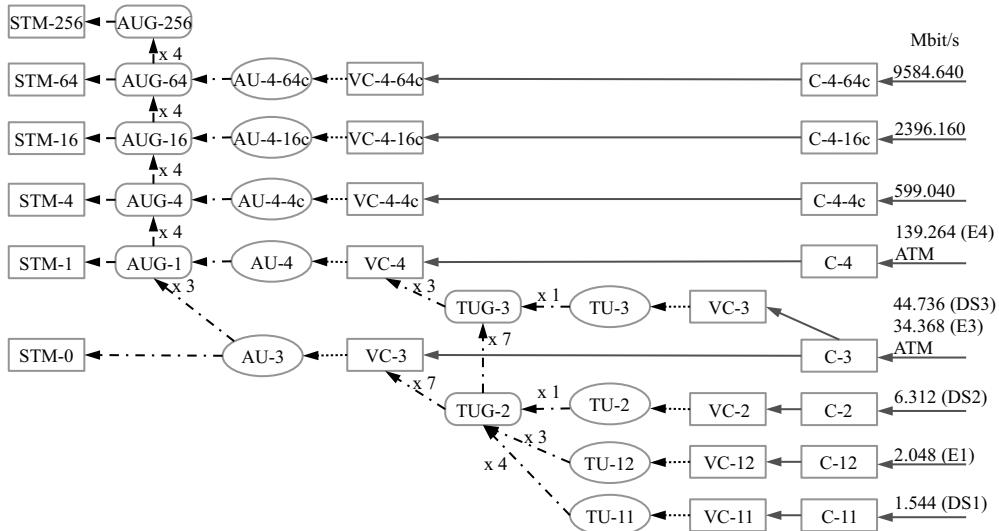


Figure 5.2 SDH/SONET rates and multiplexing hierarchy.

links are channelized. In a channelized link the payload rate is divided into multiple lower rate channels originating from standardized frame rates and structures. For example, the payload of an OC-48 link may be subdivided into four OC-12 channels. In this case the data rate of a single cell or packet flow is limited by the bandwidth of an individual channel.

In Ethernet traffic payload header is transmitted first after the whole frame is built. SDH header is called the overhead and it is interleaved with payload data. Interleaving enables very low latency. Data passing through SDH node can be delayed by at most $32\ \mu\text{s}$. SDH frame is presented in Figure 5.3.

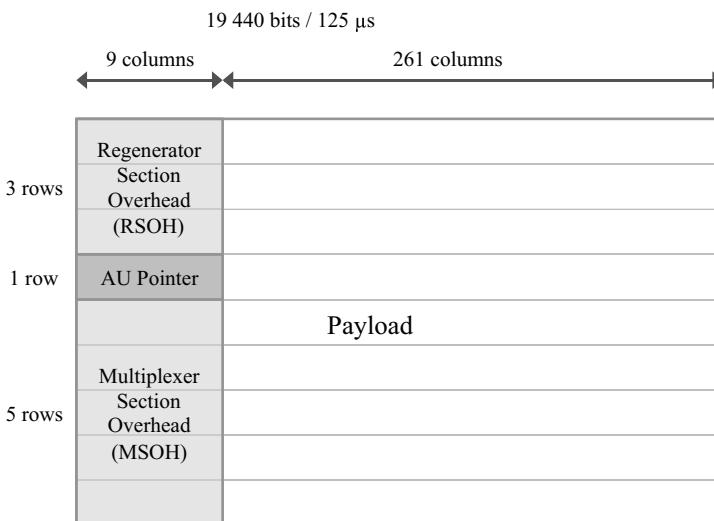


Figure 5.3 SDH frame structure.

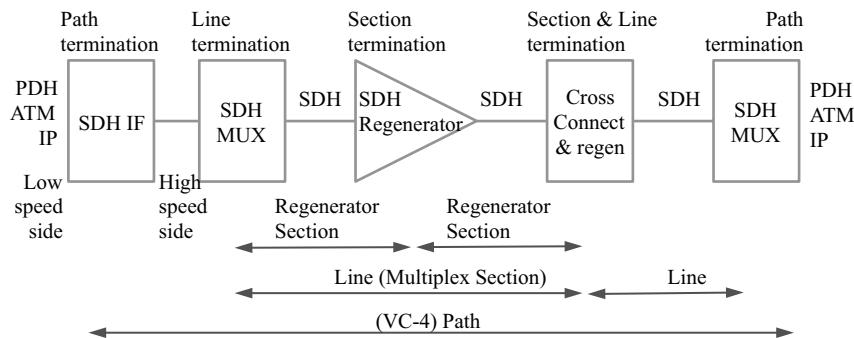


Figure 5.4 SDH network sections.

Figure 5.4 depicts the structure of SDH network sections. The smallest manageable entity of SDH network is the regenerator section. Each node having repeater function monitors the performance of the regenerator section. Each generator terminates the Regenerator Section Overhead (RSOH) part of the SDH frame (Figure 5.3) and recalculates the new content. The next broader manageable entity is multiplex section, the part of the SDH network between two multiplexers. Multiplex Section Overhead (MSOH) is used for OAM&P information related to multiplex sections. Section overhead also includes a pointer defining the position of the containers in the payload area. It is used for aligning non-synchronous payload data.

In SDH all data streams are synchronous. Data from multiple tributary streams is interleaved to a higher level in byte level. In a multiplexed SDH frame channels are in a fixed location in a frame. Demultiplexing or dropping a single channel is possible by picking up the wanted bytes from multiplexed stream without dropping the rate to lower levels as with PDH.

SDH works even if there are two different clocks in the network. SDH provides payload pointers to permit differences in the phase and frequency. It keeps the data in synchronous even when the network becomes asynchronous. The payload pointer indicates in frame level the offset between the VC payload and the STM frame. It identifies the location of the first byte of the VC in the payload and allows it to float within the STM frame. All the nodes in SDH network over the path do pointer processing and there is a separate timing reference network. The synchronization network is the network that is responsible for distributing synchronization information to network elements that need to operate synchronously.

5.1.5 SDH Protection

One strength of SDH is its multitude protection, especially ring-protection mechanisms. Term 1 + 1 protection means that two transmission channels are continuously occupied to protect each other. If the main channel fails the far-end (sink) selects the protective channel using a protection switch. In 1:1 (or in general M:N) protection source sends the data only to the working channel (or N working channels), and the protective channel (or M spare links) is in reserve. Protective channels may be used for some other (secondary) traffic. When failure is detected both ends switch to the protective channel. In general M:N protection is more efficient

in using transmission resources while $1 + 1$ may operate faster. M:N protection also requires signalling channel between protective elements to agree which path to use. Protection can be implemented as link-based or path-based. Also SDH interfaces may be protected.

Protection techniques can be categorized in several dimensions: linear protection or ring protection, trail or sub-network protection and bidirectional or unidirectional. Linear protection can operate between any pair of points within the network. It can protect against a defect in an intermediate node, a section, a line or an end-to-end transport path (Figure 5.4). Rings are a very efficient way to provide resiliency and they have also other benefits.

Often resiliency switching is considered bidirectional, but unidirectional switching may be advantageous for faster switching time, avoiding unnecessary disruptions for extra traffic and for easier implementation. The benefits of bidirectional switching are easier management and no delay unbalance between directions.

Alarm Indicator Signal (AIS) is a mechanism in PDH and SDH systems to indicate loss of signal (LOS) or loss of framing. In the case of LOS node replaces missing data with ‘1’ (‘All Ones’) towards a higher level system that is the indication of AIS. Remote Defect Indicator (RDI) is a path level indicator that an AIS or a signal failure condition is received.

Automatic Protection Switching (APS) automatically detects failures on a working channel, switches traffic to a protection channel (in source) and selects traffic reception from the protection channel (in sink). It may also revert back to the working channel once failure is repaired. APS can be used on any path where an alternative (physical) route exists. APS functionality in SDH/SONET involves reserving a protection channel with the same capacity as the protected channel. SDH APS is unidirectional.

Subnetwork connection Protection (SNCP) is a $1 + 1$ protection mechanism for SDH network. It can be deployed in ring, point to point or mesh topologies. SNCP is complementary to Multiplex Section Protection (MSP) used in physical handover interfaces. SNCP’s functional equivalent in SONET is called Unidirectional Path Switched Ring (UPSR). Multiplex Section Shared Protection Rings (MS-SP ring) offers shared protection mode. In Multiplex Section Dedicated Protection ring protection (MS-DP ring) main channels have dedicated protective channels.

In shared protection each multiplex section carries equally working channels and any section can use the protection channels in case of failure. This way the protection capacity is shared between sections in the ring.

A two-way ring has two shared fibers transmitting to opposite directions. In order to take advantage of the additional capacity of the protection channels under normal operating conditions, this capacity can be used to carry low priority traffic. Protected traffic is the traffic that will be switched to protection channel in case of fault or failure, or it can be forced. Non-preemptive Unprotected Traffic (NUT) is non-critical traffic that does not require protection mechanism and is not affected by protection mechanism. Extra (preemptive) traffic is best effort background traffic that runs on the protection channel and may be blocked when the protection channel is needed.

5.1.6 Optical Transport Hierarchy (OTH)

ITU-T G.709 defines a means of communicating data over Optical Transport Network (OTN). It is a standardized method for transparent transport of services in DWDM systems (Section 5.3.2.1) and is also known as Optical Transport Hierarchy (OTH) standard. [3]

Table 5.3 G.709 OTN rates.

Client signal	OTN Line Signal	OTUk Line Rate [Gbit/s]	OPUk Payload Rate [Gbit/s]	Frequency accuracy [ppm]
STM-16/STS-48	OTU1	2.666057	2.488320	±20
STM-64/STS-192	OTU2	10.709225	10.037629	±20
10GBASE-R/10GFC	OTU2e	11.095727	10.356012	±100
STS-768/STM-256	OTU3	43.018413	40.150519	±20
Up to 4 10GBASE-R	OTU3e2	44.583355	41.611131	±100
100GBASE-R	OTU4	111.809973	100.376298	±20

Transparent means that payload can carry all kinds of client signal (SDH/SONET, Ethernet, SAN etc.). An overview and general framework of OTNs is defined in ITU-T G.871/Y.1301. [10]

OTH provides multiplexing with full clock transparency, overhead for performance monitoring (PM) and Forward Error Correction (FEC) enhances signal reach with single string. In OTH network there is client signal independent end-to-end management, supervision and protection, and also non-intrusive monitoring of client signal.

OTH signals are named as OTUk ($k = 1, 2, \dots$) and they are serving (carrying) several kinds of client signals listed in Table 5.3. Recommendation defines payload (OPUk) and overhead structure as well as advanced error correction and OAM&P for OTN signals.

10GBASE-R defines Ethernet framing at a rate of approximately 10.3 Gbit/s. The rate does not match the rate used by SDH/SONET but the match is perfect for OTN (Table 5.2). Fiber Channel (FC) is a Gigabit speed network technology primarily used for Storage Area Networks (SAN).

The main advantage of OTN in mobile backhauling is the huge capacity of fiber.

5.1.7 Next Generation SDH (NG-SDH)

SDH/SONET is today considered as a legacy system and not optimal for packet data transmission because of TDM technology, connection-oriented nature and inflexible (re)configuration. When increase of data traffic was visible initiatives were made to utilize this legacy for bursty packet data more efficiently. There was a need to enhance flexibility of large high-capacity (metropolitan area) networks' bandwidth management and service provision, increase scalability and operational efficiency, but at the same time maximally utilize the huge spendings made to existing networks. SDH/SONET capability to provide high-bandwidth capacity is the primary reason to use it in the internet networks.

Next generation SDH bring three features to existing networks: virtual concatenation (VC), generic framing procedure (GFP) and link capacity adjustment scheme (LCAS).

Generic Framing Procedure (GFP) is defined in ITU-T G.7041. It gives standard mappings for many data services into SDH, and replaces proprietary schemes. The best known application is Ethernet up to 10 Gbit/s. Virtual Concatenation (VCAT, ITU-T G.707) configures pipes of variable bandwidth in increments down to 2 Mbit/s through an existing SDH network, with no changes needed to the network infrastructure. Each pipe can have its capacity distributed across multiple fibers. Link Capacity Adjustment Scheme

(LCAS, ITU-T G.7042) allows in-service variation of the bandwidth. It also allows protection bandwidth to be re-used for traffic. The same feature is also provided in Resilient Packet Ring technology.

Packet over SONET (PoS) is one initiative to carry data over SONET/SDH. IETF PoS specifies the use of PPP encapsulation over SONET/SDH links. PPP was designed for use on point-to-point links and is suitable for SONET/SDH links, which are provisioned as point-to-point circuits even in ring topologies. PoS frames are mapped into SONET/SDH frames and basic data rate has bandwidth of 149.760 Mbps. In most WAN applications today, routers with PoS interfaces are connected to carrier SONET rings via ADMs.

Multiservice provisioning platform (MSPP) is one of the oldest initiatives to carry Ethernet traffic over legacy SDH/SONET. MSPP nodes offer multiple interfaces of different technologies as well as ADM functionalities and also include cross-connect functionality to handle multiple backbone fiber rings in a single piece of equipment.

5.1.8 Asynchronous Transfer Mode (ATM)

Asynchronous Transfer Mode (ATM) is a circuit switching technique using asynchronous packet transmission. It uses asynchronous time-division multiplexing and it encodes data into small fixed-sized 53 bytes cells. This differs from approaches such as the Internet Protocol or Ethernet that use variable sized packets or frames. ATM systems mostly work in the data link layer and use PDH or SDH as physical layer protocol. Speed of ATM networks can reach up to 10 Gbit/s.

ATM provides various data link layer services, but it has lost its position for more favorite IP and Ethernet networking. It was designed for a network that must handle both traditional high-throughput data traffic and real-time, low-latency content such as voice and video. ATM uses a connection-oriented model in which a virtual circuit must be established between two endpoints before the actual data exchange begins.

Originally ATM was selected as a transport protocol in 3G WCDMA Radio Access Networks (RAN). Nowadays more and more new 3G BTS backhaul connections are IP based. Also B-ISDN networks were defined to use ATM transport, but only a few such networks exist anymore.

ATM uses ATM adaptation layers (AAL) to support different non-ATM services. ATM establishes virtual circuits (VC) over the network using Virtual Paths (VP) and Virtual Channels. Every ATM cell has a Virtual Path Identifier (VPI) and a Virtual Channel Identifier (VCI) that together identify the virtual circuit used by the connection. As the cells traverse an ATM network, switching takes place by changing the VPI/VCI values. The process is called label swapping. This way the traffic from one end to the other always passes the same route.

When an ATM circuit is set up each switch on the circuit is informed of the traffic class of the connection. This is called traffic contract and it is the basis for ATM Quality of Service (QoS) mechanism.

5.1.9 Hybrid TDM/Packet

Mobile backhaul networks are in transition from TDM era to fully packet-based technology. There are several reasons why this transition takes place gradually:

- Utilize legacy network investments.
- PDH/SDH carries accurate synchronization.
- PDH/SDH guarantees certain bandwidth and latency for critical traffic.
- Utilize the packet network cost benefit of statistical multiplexing with best-effort traffic.

To utilize current physical layer investments IP/Ethernet data can be mapped into PDH/SDH frames as discussed in the NG SDH section, 5.1.7. It is also possible to establish virtual PDH circuit over packet network. In this case circuit emulation terminals are needed in data sinks and sources (BTSs and controllers/gateways).

PseudoWire Emulation (PWE) is a method that provides the transparent transport of a TDM signal (E1 or T1) via a packet switched network (IP/Ethernet/MPLS). Circuit Emulation Service over Packet (CESoP) is a PWE-technology for carrying TDM traffic over IP/Ethernet/MPLS transport network. Pseudo-wire Emulation Edge-to-Edge (PWE3) is IETF initiative to define architecture for service provider edge-to-edge pseudowires (RFC 3985). Pseudo wire emulation is also discussed in Chapter 4.

5.2 Wireless Backhaul Technology

Wireless in transport has long been synonymous with microwave point-to-point (PtP) radio link. For decades microwave radios (MWR) have been used in all levels of transport networks, for short and long hauls. In the digital era main data structures have been PDH and SDH, today it is Ethernet. The majority of MW radios use frequency bands allocated for fixed services between 2 ... 15 GHz, while newer urban installations use frequency bands 18...38 GHz. Typical capacities used to be $2 \dots 16 \times 2$ Mbit/s in access section and 140 Mbit/s, STM-1 or even STM-4 in national trunk lines. The backbone network installations used to be huge in size: long and stable masts, large parabolic dish antennas, heavy radio and baseband rack installation in air-conditioned special premises.

In the upper level of (core) networks the radios are yielding to optical fiber transmission due to a huge increase of transmission rates. Wireless transport systems, however, have their benefits and are widely used in the access part of the mobile backhaul networks. Mobile and fixed network convergence is reality and this is true also in mobile backhaul.

Over 55% of all MBH physical connections worldwide are on microwave, and a total of 64% of MBH equipment revenue in 2010 was from TDM, dual TDM/Ethernet and packet microwave (Chapter 2). Due to increasing mobile data rates in LTE/LTE-A era the mobile hand-sets' distance to the BTS (eNB) will shorten. Much more capacity is needed per square-kilometer, especially in urban areas. High transport capacities are transmitted to the buildings and cabinets by optical fiber but it is not feasible to wire every small base-station. The capacity of microwave radio technology is increasing to 1 Gbit/s level while new millimeter wave bands offer wider bandwidths and opportunity up to 10 Gbit/s capacities and very dense radio networks.

5.2.1 Radio Wave Propagation

In wireless transport connections radio waves may face several anomalies that must be considered and requires planning. Issues related to radio wave propagation are:

- attenuation due to free path loss, i.e. due to distance;
- attenuation due to atmospheric gases;
- diffraction fading due to obstruction of the path obstacles;
- fading due to atmospheric multipath or beam defocusing;
- fading due to multipath from ground and other surface reflections;
- attenuation due to precipitation (weather) or solid particles in the atmosphere;
- variation of the angle-of-arrival at the receiver terminal and angle-of-launch at the transmitter terminal due to refraction;
- reduction in cross-polarization discrimination (XPD) in multipath or precipitation conditions;
- signal distortion due to frequency selective fading and delay during multipath propagation.

When radio wave propagates from transmitter to receiver its intensity decreases. Signal attenuation as a function of distance in line-of-sight link is called free space path loss (FSL) and it can be calculated from Equation 5.1.

$$FSL = 32.45 + 20 \cdot \log_{10}(f) + 20 \cdot \log_{10}(d) [dB] \quad (5.1)$$

Unit of f is MHz and d is km. One can see that when distance increases a decade (10-times) signal weakens 20 dB. In multipath non-LOS or near-LOS cases with wider than pencil beam antennas degradation is typically 35-40 dB per decade.

Radio link between two ends is not actually just a line but the radio-frequency power propagated in an area (ellipsoid) defined by Fresnel-zone (Figure 5.6(a)). Link design must be done so that approximately half of the Fresnel zone radius remains clear from obstacles in normal conditions. In lower frequencies and longer hops some margin is needed for the change in atmospheric refraction (b).

Radius of Fresnel zone in distance d_1, d_2 can be calculated from Equation 5.2.

$$r_F = 17.3 \sqrt{\frac{d_1 \cdot d_2}{f \cdot d}} [m] \quad (5.2)$$

Radius r is in meters, distances d in kilometers and frequency Gigahertz. If there is an obstacle inside the zone the additional attenuation can be estimated from Equation 5.3:

$$L_{ad} = -20 \frac{c}{r_F} + 10 [dB] \quad (5.3)$$

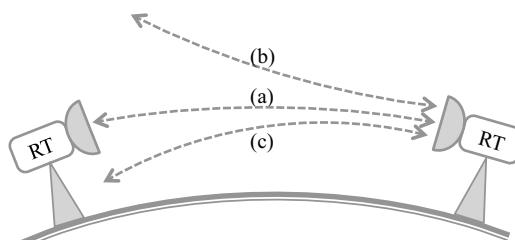


Figure 5.5 Microwave bending (de-focusing) due to changes in atmospheric refraction, a) normal atmosphere, b) sub-refractive, c) super-refractive.



Figure 5.6 MW radio hop with Fresnel zone. a) obstacle at distance d_1, d_2 limits clearance c and causes diffraction attenuation. b) sub-refraction in atmosphere makes obstacle to bulge to the line-of-sight.

A single relatively short obstacle reaching the line-of-sight causes 10 dB additional attenuation. Radio wave bending, de-focusing and diffraction attenuation are the main causes of signal degradation and outages in lower MW frequencies and over long hops (Figure 5.5).

Multi-path propagation causes selective fading that varies when atmospheric conditions change. Selective fading distorts the signal and can only be cured by proper radio modem techniques and link hop design that avoids multipath reflections. Calculation methods for radio hop parameters as well as for estimated outage times to dimension radio hops are available in [24].

Radio wave is attenuated by gases and water vapor in atmosphere, see Figure 5.7. Water attenuation is negligible below 10...15 GHz and reaches local maximum at 22 GHz where the first water molecules emission peak is. Gas attenuation is moderate in lower MW bands and peaks at 60 GHz where the oxygen attenuation is about 15 dB/km. Gas attenuation is constant and varies slightly when water vapor content changes.

Radio signal is also attenuated by rain that is a varying statistical phenomenon. Rain attenuation below about 10 GHz is negligible, but at higher frequencies and over short hops

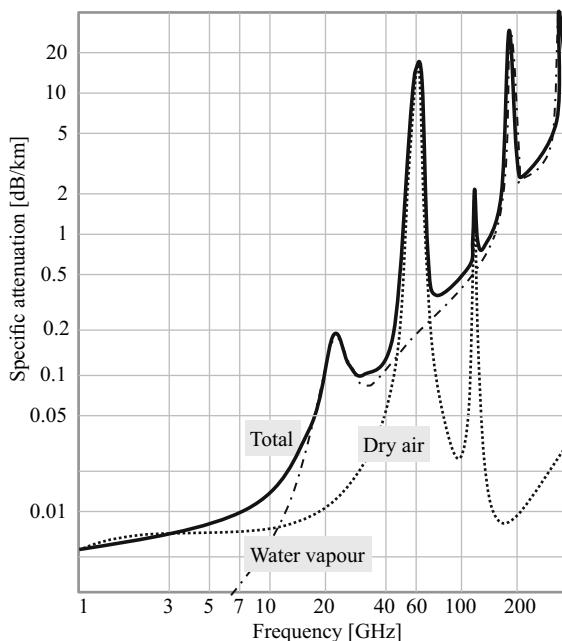


Figure 5.7 Specific attenuation due to atmospheric gases. [ITU-R P.676-8]

rain is the main cause of outages. Recommendation ITU-R P.837-4 contains maps of rain fall rates (mm/h) exceeded 0.01% of time (53 minutes per year or 4 minutes per month) for different geographical areas. Figures can be scaled for other percentages, too. Once we know the attenuation caused by certain rain-intensity as a function of frequency the needed flat fade margin can be calculated. The margin is taken to link budget tool to optimize radio hop length, antenna sizes, transmit power and other parameters.

Typical rain rates for 99.99% availability range from continental Europe 40 mm/h to Far East 100 mm/h. Radio link is to be dimensioned to withstand this high additional attenuation. Also other time-percentages can be calculated if lower or higher availability is required.

The effects of slow flat fading and faster frequency-selective fading must both be taken into account in link design. There are a number of techniques available for alleviating these effects, most of which alleviate both at the same time. The same techniques often alleviate the reductions in cross-polarization discrimination also. They can be categorized as techniques that do not require diversity reception or transmission, and techniques that do require diversity. It is desirable for economic reasons to avoid diversity whenever possible.

In order to reduce the effects of multipath fading without diversity there are several techniques that can be employed:

- increase of path inclination;
- reduction of effect of surface reflections;
- shielding of the reflection point;
- moving of reflection point to poorer reflecting surface;
- optimum choice of antenna heights;
- choice of vertical polarization;
- use of antenna discrimination;
- reduction of path clearance.

There are also installation related impairments like antenna misalignment, mast or pole vibration, RF cable and connector attenuation. Wet snow accumulating on antenna radom also cause attenuation in higher MW bands.

5.2.2 Frequencies and Capacities

Microwave radio transmits in a selected frequency band and occupies a bandwidth at that frequency. Both frequency band and bandwidth are technical design parameters, but also local spectrum regulation has an effect on frequency usability and (license) cost.

Figure 5.8 shows achievable data rates as a function of modulation used. Data rate is the higher the wider signal bandwidth and the more bits per transmitted symbol (i.e. the higher the modulation schema). More advanced modulation requires a stronger signal and is also more prone to interference. BPSK is a robust modulation using only two signal levels. 1024QAM is a highly advanced modulation method only recently introduced in MW radios. Decibels show how much signal can attenuate to achieve the same quality as with 1024QAM. So, 64QAM requires roughly 12 dB stronger signal than QPSK (4QAM).

Channel widths 7...28 MHz are widely available for major MW bands globally. The number of 56 MHz channels is quite limited and they may already have been allocated

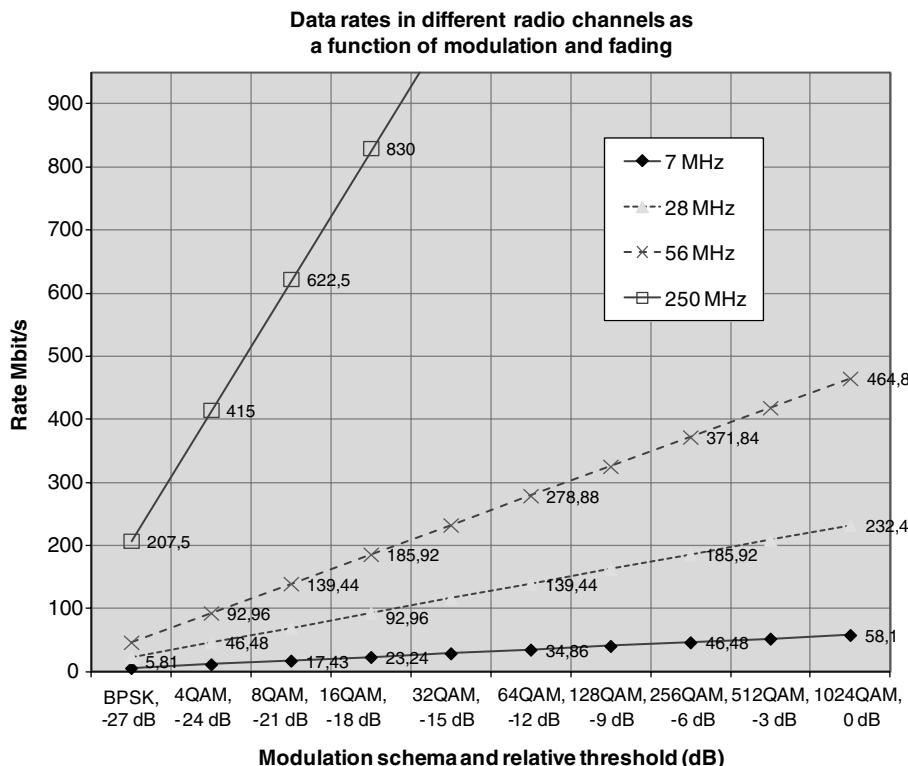


Figure 5.8 Radio data rate as a function modulation with different radio channel bandwidths.

for prevailing operators. 250 MHz channels are available in the newest high capacity bands at 40...90 GHz but products for those capacities are practically non-existent.

The figure also indicates the effect of adaptive modulation (AM) that is a common feature of modern MW radios. When radio link degrades, for example, due to rain, AM algorithm changes the modulation to keep a communication link on with decreased capacity. This is beneficial in mobile networks because dropped connection also drops the base-stations' control. When the link returns it takes a while before BTS is synchronized again, but if at least minimum capacity is available, the control can be maintained. This would be tricky with TDM backhaul, but packet-base (Ethernet/IP) backhaul with proper QoS procedures can handle changing capacity. Lower service classes (best-effort) are dropped first.

Traditionally most PtP MW links are made using FDD duplex schema and frequency regulation in most countries is based on FDD. Each hop occupies a pair of frequencies and bandwidth is equal in both directions (uplink and downlink). TDD is also recognized and it is mainly in use in point-to-area links. One technique to increase capacity with existing channels is to use dual-polarization. This requires that MW radio have special technology, cross polarization interference cancellation (XPIC), in use. Dual-polarization can be used for polarization diversity as well as for Multiple Input Multiple Output (MIMO) purposes.

Table 5.4 shows frequency allocations for fixed wireless service. It is a generic table for ITU Region 1 (Europe etc.) and does not specify exactly how frequencies are used in a specific

Table 5.4 Frequency allocations for fixed service.

Frequency band [GHz]		Bandwidth, [GHz]	Channel-widths available	Use in CEPT- countries	Typical hop
5,900	7,100	1,200	3.5/7/14/20/30/40/600/ 80 MHz	A lot	55 km
7,125	8,500	1,375	7/14/28/56 MHz	A lot	53 km
10,000	10,680	0,680	3.5/7/14/28/56 MHz	Medium	40 km
10,700	12,500	1,800	28/40/56/80 MHz	Medium	40 km
12,750	13,250	0,500	3.5/7/14/28/56 MHz	A lot	25 km
14,500	15,350	0,850	3.5/7/14/28/56 MHz	A lot	23 km
17,700	19,700	2,000	13.75/27.5/55 MHz	A lot	20 km
21,200	22,000	0,800	3.5/7/14/28/56/112 MHz	Few	
22,000	23,600	1,600	3.5/7/14/28/56/112 MHz	A lot	14 km
24,200	24,500	0,300		Few	
24,500	26,500	2,000	3.5/7/14/28/56/112 MHz	A lot	9 km
27,500	29,500	2,000	3.5/7/14/28/56/112 MHz	Medium	4 km
31,000	31,300	0,300	3.5/7/14/28 MHz	Few	
31,800	33,400	1,600	3.5/7/14/28/56/112 MHz + block	Medium	3 km
37,000	39,000	2,000	3.5/7/14/28/56/112 MHz	A lot	5 km
40,500	43,500	3,000	7/14/28/56/112 MHz + block	Few	3 km
48,500	50,200	1,700	3.5/7/14/28 MHz		
51,400	52,600	1,200	3.5/7/14/28/56 MHz		2 km
55,780	57,000	1,220	3.5/7/14/28/56 MHz		
57,000	59,000	2,000	50/100 MHz	Few	
59,300	62,000	2,700	-		
61,000	61,500	0,500			
64,000	66,000	2,000	50...2500 MHz	Few	3 km
71,000	76,000	5,000	250...2250/4500 MHz	Few	2 km
81,000	86,000	5,000	"	Few	
92,000	95,000	3,000		Few	
Total		46,325			

country. Frequency regulation is carried out under internationally agreed principles, but under national authority. Principles of how frequency licenses are granted vary. Table 5.4 also gives information about band usage in CEPT-area and typical MW radio hop lengths.

Lower frequency bands 6...13 GHz are deployed for long trunk lines usually in rural areas. Hop-lengths are tens of kilometers. In some countries there may even be minimum allowable hop length limitations for below 10 GHz hops to ensure efficient use of frequency spectrum. Old long-haul bands below 6 GHz are too narrow for modern transmission and many of them have already re-allocated to mobile service.

Frequency bands 23...38 GHz are very popular for urban installations. Hops are 5...15 km long. Antennas are smaller and these radios fit very well to urban BTS installations. The number of wide channels (28/56 MHz) is very limited and in many cities these channels start to be congested.

New bands 32 GHz and 42 GHz are being opened in many countries to ease the need for high capacity backhaul (LTE). Also millimeter wave (mmW) bands in the range 50...90 GHz are coming onto the scene. For example, bands 57...66 GHz and 71...94 GHz are now widely under consideration for future small cell wireless backhauling. Frequency regulation is done or under preparation in several countries for these bands. Commercial semiconductor technology is available and research projects ongoing on how to make that technology more inexpensive. According to ITU-R frequencies up to 400 GHz are considered as radio frequencies. Between 40 – 275 GHz there are frequency allocations for fixed traffic 131 GHz altogether. There are several 5...10 GHz continuous (paired/un-paired) bands available that enable over 10 Gbit/s transmissions.

Also, frequencies below 10 GHz have gained interest in mobile backhaul. They offer an alternative for backhauling in non- or near-line-of-sight environment using low-cost hardware known from WLANs and mobile BTSs. A drawback is limited spectrum and increasing usage; for example 5.8 GHz RLAN band have a bandwidth of about 250 MHz. Even BTSs/eNB's themselves and the frequencies they employ for mobile traffic could be used for backhauling. This technique is called relaying or in-band/out-band backhauling. It makes mobile network roll-out and network planning easier when there is only one technology to consider. The drawback is that backhauling eats up the valuable mobile spectrum.

One key factor when evaluating the total cost of radio link is the frequency license fee. It is typically paid annually and for traditional point-to-point radios it is paid per every hop. Level of license fees varies a lot between countries, from a small fee to auctioned market price that creates a significant cost for the MWR user. License fees are typically associated with frequency (interference) co-ordination done by a national regulator for each and every radio transmitter. From the fee point of view unlicensed bands may be an attractive alternative. But, on the other hand, lack of co-ordination also increases the risk of costs due to non-functioning backhaul. In some bands and in some countries a system called light licensing is introduced. A licensee makes him/herself a reservation using a web-application on a first come first served basis.

5.2.3 Network Topologies

The basic topologies for radio link technology are point-to-point, point-to-multipoint and multipoint-to-multipoint. Point-to-point links establish radio connection between two points – near-end and far-end – with narrow beam called pencil-beam (Figure 5.9). With MW radios there must be line-of-sight (LOS) between the ends. Duplex communication takes place with FDD. Point-to-point is by far the most common MWR topology used in the MBH networks – the point-to-point MWR links can then form various network level topologies depending on the density and location of base station sites; sometimes rings or other ‘closed’ topologies are used to enable network level protection (have an alternative route for the base station connection).

Point-to-multipoint topology or architecture is also called point-to-area because PMP access-point (AP) provides coverage in a similar way as a mobile base-station. Typically economies go so that one AP is several times (3...5) more expensive than a terminal equipment (TE). Access technology is often TDD that enables different temporary capacity needs to be adapted per TE. The main difficulty in adopting PMP MW system is often the line-of-sight requirement – it can be difficult to have LOS for so many base station sites within the (theoretical) coverage area of the AP that it grants PMP system feasibility.

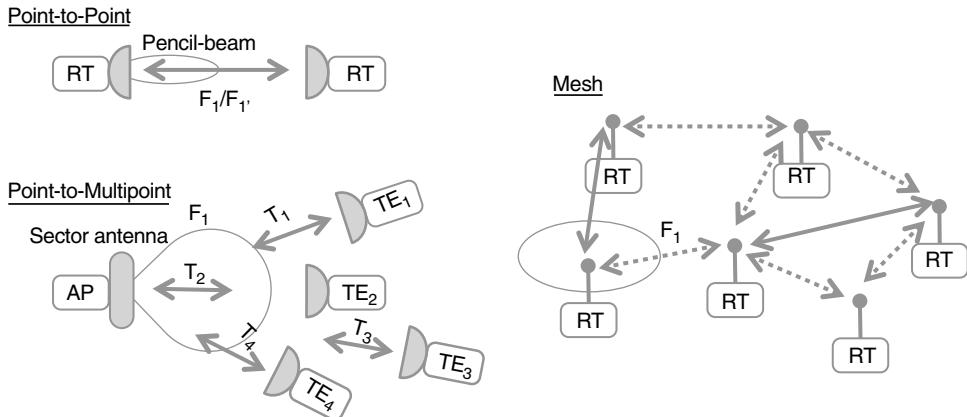


Figure 5.9 Examples of wireless network topologies: Point-to-point, Point-to-Multipoint, and Multipoint-to-Multipoint/Mesh. RT = Radio Terminal, TE = Terminal Equipment, AP = Access Point.

Multipoint-to-multipoint or (partial) mesh topology enables complex networks that can utilize packet transmission protocols to make advanced load sharing, routing and protection. Today mesh is not much in use in MBH while the technology and its benefits are under study.

When looking at larger microwave radio network topologies they can be classified as tree, chain, star and ring, or a mixture of all these. The most common microwave radio link network topology is tree. In tree the capacity is carried to the access leaves through higher capacity trunk lines. It is a well proven way to build networks. Tree sets higher reliability requirements to trunk lines. Some branches of the tree may contain a chain of several links. In star topology capacity is distributed from one point to several directions. Ring is a kind of chain with both ends connected. Ring topology is an efficient way to bring redundancy to the network, but it also requires special protocols.

Availability of frequency channels, access method (TDD/FDD) and geography sets also limits to the network topology selection.

Point-to-multipoint topologies are typically used in broadband wireless access (BWA) systems that provide fixed wireless access for a certain geographical area.

Mesh topologies are used commonly in consumer wireless access systems like WLAN.

5.2.4 Availability and Resiliency

To protect against hardware failures microwave radios can be installed in stand-by configuration (Figure 5.10). It is typically a 1:1 configuration where protective radio (TRX) is ready to take over the traffic if the working radio fails. Both the working and redundant radios occupy

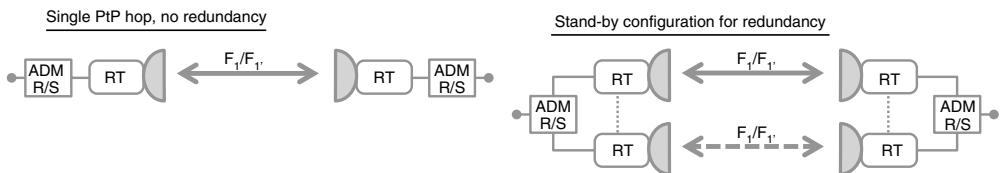


Figure 5.10 Radio link protection.

the same frequency channel. In the other end change-over is not needed, but control channel is used to indicate the loss of signal.

Typically the Mean Time to Failure (MTTF) is shorter and repairing takes more effort for radio equipment located outdoors, if compared to switching equipment in indoor premises.

Radio hop protection methods are called diversity methods and they operate as 1 + 1. Most common are (Figure 5.11):

- Space Diversity (SD): two antenna separated vertically.
- Frequency Diversity (FD) or Polarization Diversity (PD): two frequencies or polarizations in one antenna.
- Route Diversity (RD): protective path is horizontally separated.

Only route diversity is an effective mean against rain attenuation. Other methods are mainly used against multi-path diffraction attenuation in lower frequencies (Section 5.2.1). Equipment and hop protection can be combined, for example stand-by mode with diversity. There is more about backhaul resilience in Chapter 9. Diversity and resiliency configurations usage in MBH access network is decreasing due to installation and cost requirements, and also because the mobile network itself provides resiliency. N + 1 type of redundancy is still used in trunk networks.

Ethernet ring provides reliable multi-point connectivity, but at the same time protocol must avoid loops. Two basic principles to make Ethernet rings are:

- open loop: when protection is required block the failed link;
- closed loop: when protection is required, steer traffic regardless of spanning tree protocol.

There are numerous Ethernet ring protocols; some of them are standardized, some proprietary. Typically, support for different loop types is required: single loop, overlapping loops and conjoined rings (loops with common nodes). Ethernet Ring Protection (ERP) is a L2 mechanism. It uses control VLAN for monitoring and control and Payload VLAN for data. Each payload VLAN is assigned to the control VLAN which is managed by one of the switches that act as a master. During normal operation ERP master blocks the ERP payload traffic preventing loops. Link failures are detected by 1) loopback packets stop circulating around the ring, or 2) ERP-aware switches report detected link failure to the master.

Ethernet ring protection recommendation (G.8032/Y.1344) defines automatic protection switching (APS) for Ethernet ring topologies. Loop protection is achieved by ensuring that traffic may flow in all but one direction at a time. Ethernet rings support a multi-ring network that consists of conjoined rings. APS protocol is used to control projection over the ring. Under specified conditions the switching time in case of failure in the Ethernet ring shall be less than ‘magical’ 50 ms.

5.2.5 Performance

One step in radio link dimensioning is to set the target availability level for the whole backhaul network. Some recommendations for large telecommunication networks can be found from [T-REC-G.826 or G.828]. In data networks and mobile backhaul the last miles target availability levels are key parameters to optimize network cost and customer perceived

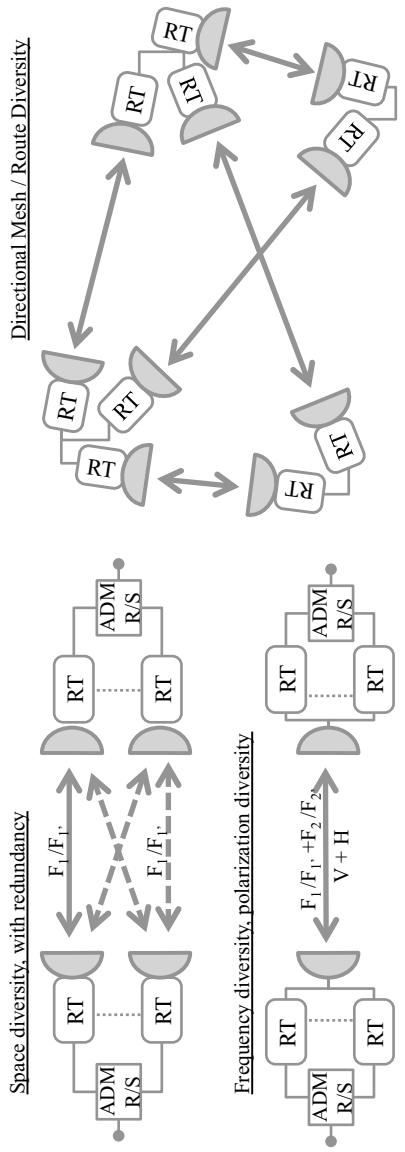


Figure 5.11 Radio link diversity configurations.

quality. It is always a balancing act between the total availability in network level taking into account the customers, mobile services and business case.

Radio link performance indicators are defined in ITU-T G.826. Errored Block (EB) is a block in which one or more bits are in error. Consecutive bits associated with a path form a block. Errored Second (ES) over the whole path is a one-second period during which there is one or more errored blocks or at least one defect. Over a single connection ES period also occurs when Loss of Signal (LOS) or Alarm Indication Signal (AIS) is detected. Severely Errored Second (SES) is a one-second period which contains $\geq 30\%$ errored blocks or at least one defect. SES is a subset of ES. Over a connection during SES bit-error ratio $\geq 10^{-3}$ or during which LOS or AIS is detected. Background Block Error (BBE) is an errored block not occurring as part of an SES, i.e. a sporadic error under good signal condition. [9]

These standard performance figures are typically recorded in every node and used to characterize the wireless backhaul quality. They are also monitored for end-to-end paths and may be used to verify SLA conformance.

Respective performance parameters related to the measurement time are Errored Second Ratio (ESR), Severely Errored Second Ratio (SESR) and Background Block Error Ratio (BBER). There are also error performance definitions for SDH and ATM cell-based transmission.

5.2.6 Other Wireless Technologies

One wireless technology often mentioned when high capacities are needed is Free Space Optics (FSO). It means using high intensity modulated laser beams for transmission. FSO products and application have existed for a long time for high-capacity data networking while they have not gained a market share in mobile backhauling. The main reason is because of still quite clumsy and expensive technology and low availability due to atmospheric (fog) attenuation.

Satellite connections are also used for mobile backhauling in some cases. Typical applications are in sparsely populated areas where other transport systems are even more expensive than satellite. Transport from geostationary satellites is not optimum because of ground station requirements (for backhaul capacities) and long latency. Recently there have been proposals to provide backhaul capacity from Medium Earth Orbit (MEO) satellites with link budget and latency that meets MBH requirements.

5.3 Wire-Line Backhaul Technology

5.3.1 DSL Technologies

Digital Subscriber Line (DSL) technologies utilize the existing copper lines originally installed for Plain Old Telephone Service (POTS) for data transmission. DSL is used for access transmission to provide IP services for home computer users, but it is very common backhauling technology, too. The DSL subscriber line is called loop and it is typically a cable pair, but most DSL-variants can also combine (bundle) several pairs. The reach and maximum capacity of different xDSL technologies varies according to the loop rate, noise conditions and complexity of modulation and coding, see Table 5.5.

The most common technology is Asymmetric Digital Subscriber Line (ADSL). It has different standards for Europe and North America, while the newest ADSL+ is more

Table 5.5 Properties of some DSL techniques [ITU G.991, G.992, G.993].

Name	Reference standard	Maximum speed down/up	Symm/ Asymm	Loop length	Pairs	Bonding
ADSL	ANSI T1.413 Issue 2	8.0/1.0 Mbit/s	A	... 5 km	1	Yes
	ITU G.992.1 Annex A	12.0/1.3 Mbit/s				
ADSL2 +	ITU G.992.5 Annex M	24.0/3.5 Mbit/s	A	... 9 km	1 ...	32
SHDSL	ITU G.991.2	2.3/2.3 Mbit/s	S	... 3 km	2	4
VDSL	ITU G.993.1	52/16 Mbit/s	A	... 1200 m	1	
VDSL2	ITU G.993.2	250 ... 4 Mbit/s	S + A	50 m ... 5 km		

harmonized. It uses the same line as POTS but a DSL filter allows the same copper line to be used for both voice and data transmission. Capacity is asymmetric; downlink direction to the BTS or customer premises is greater than uplink capacity. The configurations are star-like. The customer premises equipment (CPE) is small and cheap while the network end, called Digital Subscriber Line Access Multiplexer (DSLAM), usually terminates several tens or hundreds of access lines and connects to the data network by ATM/IP/SDH interface. With G.bond extension capacity can be doubled to 48/6 Mbit/s.

Single-Pair High-speed Digital Subscriber Line (SHDSL) is a symmetric DSL technology. SHDSL occupies the whole bandwidth for data transmission and POTS speech service is not possible in the same copper line. SHDSL has been popular in backhaul installations when copper pairs have been available and capacity requirement reasonable, because it provides directly capability for E1/T1, ATM and Ethernet transmission. SHDSL standard is ITU-T G.991.2 and it is also known as G.SHDSL. The updated version is known as G.SHDSL.bis or SHDSL.bis. SHDSL features symmetrical data rates up to 2,304 kbit/s for one-pair installations. The two pair feature may alternatively be used for increased reach by keeping the data rate low. Halving the data rate per pair will provide similar speeds to single pair lines while increasing the error/noise tolerance. Higher data rates may be achieved using two or up to four copper pairs.

Very high bit rate digital subscriber line (VDSL) is the fastest technology available. It enables configurable up/downlink speeds over 100 Mbit/s some hundreds of meters loop. VDSL is commonly used as ‘fiber extension’ to deliver FTTB further to the home or to the BTS site. It is enough for LTE eNB rates in the beginning when site rates are not very high, or for small cell sites with only a few users per eNB.

Copper line is neither stable nor interference-free signal path. Usually the line rate is negotiated in the beginning of the connection, but electro-magnetic interferences can disturb the signal later causing errors. It depends on DSL modem implementation how well it can adapt to the situation and changing traffic patterns. Due to heavy interleaving there may also be challenges to keep packet time delay variation in control for synchronization purposes.

Figure 5.12 presents measured uplink and downlink rates from an IP-DSLAM node using VDSL2 link to CPE. Measurements are performed with and without cross-talk. The degradation caused the interference is worse at shorter distances.

There are several advanced methods to further improve the performance of DSL. Dynamic Spectrum Management includes methods like rate adaptation and near/far-crosstalk cancellation. Simple methods balance the signal spectrum in one cable pair. More complex methods do the processing for all parallel signals in a node and bunch of cables. They require heavy signal processing in central office end and typically some processing also in terminal end.

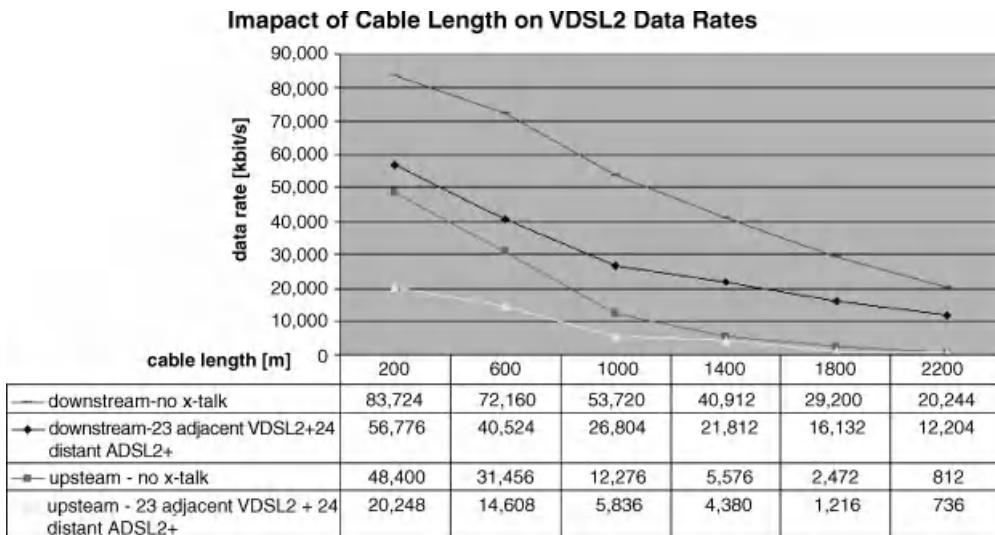


Figure 5.12 Impact of cable length on VDSL2 data rates.

VDSL vectoring is a good example of how bit-rates have increased from 100 Mbit/s level to over 200 Mbit/s level. DSL rates up to 1 Gbit/s over a copper-pair have been demonstrated in laboratories.

Pair bonding is the technique of combining two or more parallel copper lines to provide greater bandwidth. In typical copper cable installations there are usually several spare pairs. More and more household and office building users are moving to wireless leaving pairs for MBH use. These pairs can be utilized as long as they are valid. With current MBH capacity requirements it is a viable alternative to newly laid optical fiber. Factors such as the accessibility of the pairs and the requirements of central-office equipment are also important when evaluating feasibility.

5.3.2 Optical Technology

Optical fiber is a thin highly transparent coated glass-tube inside which modulated laser light transfers. Modern fibers have attenuation about 0.2 dB/km meaning 100 km run attenuates signal only 20 dB which sits in the link budget of many optical systems. Fibers constitute a cable.

Multi-mode optical fiber has a larger core (typical 50 μm) that allows use of inexpensive connectors, optical transmitters and receivers. Multi-mode fiber material is more expensive and causes dispersion that limits signal bandwidth and causes attenuation. Single-mode fiber is thinner (typically $< 10 \mu\text{m}$). It allows wider bandwidth and lower attenuation, but requires more expensive components and interconnection methods.

Connecting plain optical fibers requires special equipment and skills. Once optical connectors are installed in the fiber ends connecting can be performed as with any other cable. Optical systems strength is the long reach that enables concentrating processing intensive equipment in mobile networks into a few central offices.

Fiber optic transmission utilizes certain wavelength ranges with low attenuation, called windows and avoids those wavelengths which have naturally high attenuation. The first

window is at 800–900 nm. At the second window 1300 nm fiber attenuation and dispersion are much lower, enabling long distances. The third window at 1500 nm is most used nowadays due to widely available amplifiers and low fiber attenuation. Hydroxide causes high attenuation peak at about 1400 nm if OH molecules are present in the medium.

5.3.2.1 Wavelength Division Multiplexing

Wavelength Division Multiplexing (WDM) is a technique that utilizes several optical wavelengths in the same fiber. Basically, it is FDM technique, but in the optical spectrum usually word ‘wavelength’ or lambda is used instead of frequency. Each wavelength is a channel dedicated to a certain user or service (Figure 5.13).

Without wavelength (optical) multiplexing the whole spectrum must be converted to electrical signal and multiplexed in a conventional way (for example SDH). Optical-Electrical-Optical-conversion (OEO) decreases the length between active nodes and increases cost and complexity. Optical multiplexing is a passive operation. The only active components needed through the whole transmission path that can be even thousands of kilometers are optical amplifiers.

Coarse WDM (CWDM) uses wide wavelength range 1270...1610 nm with a coarse grid. ITU-T G.694.2 defines 18 channels. Due to relaxed accuracy requirements the reach of CWDM is limited to about 60 km and it is suitable for 2.5 Gbit/s rates. SFP TRX-modules are available for CWDM to allow upgrading existing systems to this technology. CWDM is also used to transmit upstream and downstream optical signals in a single fiber using bi-directional transceiver (BiDi SFP). Typically transmitters (lasers) are quite frequency coherent but receivers are more wide band. That is why frequency selective filters are needed in demultiplexers to pick up just the wanted channel.

Dense WDM (DWDM) uses C-band and L-band (1530...1620 nm) wavelengths that enable a much longer reach. DWDM systems range from 40 channels or ‘colors’ with 100 GHz grid to 160 channel with 25 GHz channel spacing. ITU-T G.694.1 defines frequency grid for DWDM. Systems use a fiber pair with transponders and multiplexers capable of handling TX and RX directions. Transceiver is a combination of optical transmitter and receiver. It can convert the electrical signal (OEO) directly to the wanted wavelength or to the ‘grey light’ containing all wavelengths. Transponder is a wavelength converter tuned to the wanted channel in the fiber network side and receiving either grey optical or electrical signal from the client system.

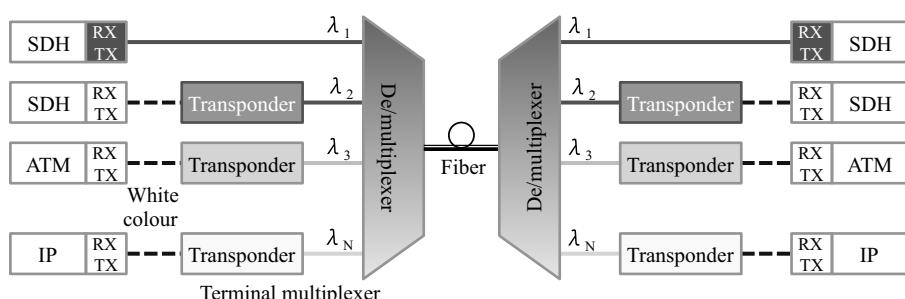


Figure 5.13 WDM de/multiplexing.

WDM is used in very long haul high capacity core and international connections. WDM is used in MBH aggregation and usage also in access is increasing.

Reconfigurable optical add-drop multiplexer (ROADM) is an optical add-drop multiplexer that can be remotely configured to switch the individual of multiple wavelengths' traffic from optical signal. This allows data channels to be added or dropped from a WDM system without the need to OEO.

5.3.2.2 Passive Optical Network (PON)

Passive optical network (PON) is a point-to-multipoint fiber access technology that uses unpowered optical splitters to serve several end points. Hub-point or central office for PON is called optical line terminal (OLT) and end points are optical network units (ONU). In mobile backhaul use ONU may also be called Cellular Backhaul Unit (CBU). The point-to-multipoint fiber tree and branch options are called optical distribution network (ODN).

With PON less fiber and less central equipment is needed than with point-to-point optical links (Table 5.6). Typical splits are 16...128 ONUs per one OLT. Downstream signals are broadcast, i.e. every ONU can 'see' the signal. This may require coding/encryption. Upstream signals are combined using a multiple access protocol, usually time division multiple access (TDMA) that causes higher upstream delay. The OLTs measure the range to the ONUs in order to provide time slot assignments for upstream communication.

GPON (ITU-T G.984) networks have now been deployed in numerous networks across the globe. ITU G.987 defines 10G-PON (or XG-PON) with 10 Gbit/s downstream and 2.5 Gbit/s upstream. Framing is close to GPON and designed to coexist in the same network. Due to attenuation of splitters the range is typically limited to couple of tens of kilometers and split to 32.

Ethernet PON (EPON or GEPON) and 10G-EPON is included as part of the Ethernet in the first mile (EFM) definition. EPON uses standard Ethernet frames with symmetric 1 Gbit/s upstream and downstream rates. 10 Gbit/s EPON supports simultaneous operation of 10 Gbit/s on one wavelength and 1 Gbit/s on a separate wavelength. The 10G-PON wavelengths differ from GPON and EPON, allowing it to coexist on the same fiber with either of the Gigabit PONs. EPON is the most widely deployed PON technology globally. EPON is also part of the DOCSIS Provisioning of EPON (DPoE) specifications. DPoE makes the EPON OLT act like a DOCSIS cable modem. In addition DPoE supports MEF 9 and 14 Ethernet services.

Table 5.6 Comparison of PON technologies [Standards mentioned].

	EPON	GPON	10G-PON	WDM-PON
Standard	IEEE 802.3ah	ITU G.984	ITU G.987 IEEE 802.3av	No
Users/PON	32768 max	128 max		32 typical
System BW	1 Gbit/s symmetric	2.5 Gbit/s down, 1,25 Gbit/s up	100 Gbit/s down, 10/1 Gbit/s up	<1 Tbit/s
Average user BW	67 Mbit/s	120 Mbit/s	250 Mbit/s	1 Gbit/s
Frame type	Ethernet	GPON Encapsulation	Ethernet	Agnostic

Table 5.7 Fast Ethernet (FE, 100 Mbit/s) standards. MM=Multi-mode, SM=Single-mode.

Abbreviation	Medium	Reach up to	Description
100BASE-T	Copper		Basic abbreviation for FE over twisted pair cables (up to Cat 5)
100BASE-TX	Copper, 2 pairs	100 m/segment	Dominating FE standard for two-pair Cat 5 and above cables
100BASE-FX	Fiber, Pair MM	400 m simplex	1300 nm near-infrared (NIR) wavelength transmitted via two strands of optical fiber, one for receive and the other for transmit
100BASE-BX	Fiber, Single SM	2 km duplex 40 km	Multiplexers split signal into transmit and receive wavelengths 1310/1550 nm. The downstream terminal uses 1550 nm and upstream 1310 nm wavelength
100BASE-LX10	Fiber, Pair SM	10 km nominal	1310 nm wavelength

Wavelength Division Multiplexing PON (WDM-PON) utilizes several wavelengths. It allocates one or more dedicated wavelength(s) for each ONU or concatenates wavelengths to vary transport capacity. There is no WDM-PON standard available but some vendors are working on it.

5.3.3 Ethernet Interfaces

Most common data access connection is 10/100/1000 Mbit/s Ethernet. Usually all speeds are supported by the same interface card. In contrary to DSL cabling Ethernet uses special cables categorized based on the maximum speed they can carry. The key for cables is to minimize cross-talk between pairs. Summary of Ethernet standards are in tables 5.7 and 5.8.

Cat 5 is the Fast Ethernet (100 Mbps) cable up to 100 m lengths. The cable, termination and verification are specified in ANSI/TIA/EIA-568 A or B. Standard connector is called RJ-45 or 8P8C modular connectors. Category 5e is the enhanced version. Basic cable type in today's

Table 5.8 Giga Ethernet (GE, 1 Gbit/s) standards. MM=Multi-mode, SM=Single-mode.

Abbreviation	Medium	Reach up to	Description
1000BASE-T	Copper, 4 pairs	100 m	Basic abbreviation for GE over twisted pair cables (Cat 5e and up)
1000BASE-TX	Copper, 2 pairs	100 m	GE over Cat 6 and up
1000BASE-SX	Fiber, Pair MM	220...550 m	770...860 nm laser
1000BASE-LX	Fiber, MM/SM	550 m MM 5 km SM	1270...1355 nm laser
1000BASE-BX10	Fiber, Single SM	10 km	1490 nm down and 1310 nm upstream
1000BASE-ZX	Fiber, SM	70...100 km	Non-standard interface using high-quality single-mode fiber at 1550 nm

installation is Cat 6 or higher. Cat 6 cable is for Gigabit Ethernet speeds and cable length up to 50 m. Length depends on the rate and total quality of the installation, and the bandwidth must be verified if targeting the maximum speeds. Using Cat-6a cable 10GBASET connections up to 100 m are possible. Even wider bandwidth Cat 7 cable standard have been defined but not widely used today.

10 and 100BASE-T use two pairs and with a splitter two 100BASE-T lines can be transferred over one cable. 1000Baset uses four pairs. 100BASE-FX use SC, ST, LC, MTRJ or MIC connectors. LC and SC connectors are the most commonly used ones. Telecom field installations and outside plants (OSP), however, require hardened fiber optic connectors like BX5. Telecommunication installations tend to use low-attenuating high-quality cables while the short range datacom typically use lower cost cables.

It is possible that Ethernet cards have separate detachable Gigabit Interface Converters (GBIC) that performs the interfacing functionality to copper or fiber. It makes it easier to change transport media, swap broken interface and manage spare parts. De-facto interface module nowadays is a small form-factor pluggable (SFP) module. Modules are not fully compatible and also some vendors may have vendor a lock-in feature that forces only proprietary modules to be used.

5.3.4 Ethernet in the First Mile

Ethernet in the First Mile (EFM) is the name for a set of Ethernet standards in amendment IEEE Std 802.3ah-2004. It is a set of additional specifications, allowing users to run the Ethernet protocol over various media, such as a pair of telephone wiring and single strands of fiber (Table 5.9).

Other extensions are 100BASE-LX10, 100BASE-BX10 and 1000BASE-BX10. These extensions make the EFM port types better suited for use in access networks and mobile backhauling.

5.3.5 DOCSIS

Data Over Cable Service Interface Specification (DOCSIS) is a method of transporting data over cable-TV network using modulated RF-signal. It was first standardized for the USA and

Table 5.9 Ethernet in the First Mile (EFM) additions to 802.3.

Abbreviation	Medium	Reach up to	Description
2BASE-TL	Copper, POTS	2.7 km	Up to 5.696 Mb/s (varying) over POTS wires at distances in the order of 2.7 km based on SHDSL
10PASS-TS	Copper, POTS	100 m	Up to 10 Mb/s (varying) over POTS at distances in the order of 750 m based on VDSL
1000BASE-LX10	Fiber, Pair SM	10 km	Single wavelength 1270...1355 nm, high-quality cabling
1000BASE-PX10	Fiber, SM	10 km	Ethernet passive optical network (EPON) between up to 16 users.
1000BASE-PX20		20 km	Ethernet passive optical network (EPON) between up to 16 users.

there is also a version for Europe called EuroDOCSIS. First versions of DOCSIS were specified mainly for cable-only TV signal transmission, but later included improvements for internet-traffic. The service allows bi-directional transfer of data between the cable system headend (central-office or hub) and customer locations over an all-coaxial or hybrid fiber-coax (HFC) cable network. The headend interface is called as Cable Modem Termination System (CMTS) and customer premises equipment as Cable Modem (CM). Cross-version compatibility has been maintained across all versions of DOCSIS. [19]

A maximum optical/electrical distance between the CMTS and CM is 160 km (100 miles) in each direction. It uses 6 MHz RF-channels and speed is 38 Mbit/s down and 27 Mbit/s uplink. Channels can be bonded so that next generation 8 carriers bonding yield up to 304 Mbit/s.

DOCSIS is a potential last mile mobile backhaul access method because of large penetration of cable-TV lines in populated areas and high speed. However, it is not widely used.

5.4 Aggregation and Backbone Tiers

As mentioned in the beginning of the chapter, transport systems used in MBH upper tiers, i.e. in aggregation and backbone networks, are similar to those used in fixed networks. Characteristics of MBH aggregation network is that it is not only transferring bits, but also processes and combines traffic streams from high number of sources.

In legacy networks aggregation site may be the start-point of several micro-wave radios having a large common PDH/SDH multiplexer. In 2G or low-capacity 3G mobile networks even the first hops in the aggregation level are realized by wireless links. In upgraded networks they have been replaced partially (or wholly) by new NG-SDH/MSPP nodes which include packet switching capabilities. In legacy aggregation networks bit rates are from STM-1 to STM-16 and in backbone networks typically STM-16, in a few cases also STM-64; in high capacity networks these signals may be optically multiplexed, i.e. DWDM is used to increase fiber capacity. Topologically these networks are usually rings, so that alternative path is always available and fast protection switching can be applied.

In newer packet-based aggregation and backbone networks the main elements are routers (IP or IP/MPLS), Ethernet switches and layer 1 transport links connecting them together. In the simplest case the connecting links are just dark fibers between the optical interfaces of routers and/or Ethernet switches – in both cases most often high speed Ethernet interfaces (1G, 10G or higher bit rate). Typical routers and switches have tens of Gbit/s interface cards, can handle several protocols and are highly redundant. They form Points of Presence (POPs) for transmission network connections to IP services and also work as edge router in the edge of network clouds.

In other cases there may be extensive optical networks, an optical layer, connecting the sites where the routers and switches are located; in these cases the connection may be based on a wavelength in DWDM systems, routed through that optical network to connect the wanted router/switch interfaces. Capacities in these networks can be significantly higher than in the above mentioned legacy networks; in aggregation networks already up to 10 Gbit/s and in backbone networks n*10G or n*40G or even higher. Also in these networks alternative routes usually exist between all important nodes; however, due to packet switching, topologies are not limited to rings but are of partial mesh type.

5.5 Leased Line Services for Mobile Backhaul

Leased Line Services are one alternative for the Mobile Operator to implement the transport connections in the MBH network. Conventional TDM leased lines (i.e. PDH E1s and T1s) have been used right from the beginning of first Mobile Networks (2G, GSM), but the transition to packet based technologies and higher MBH bit rates has led to TDM leased lines gradually being superseded by Ethernet services. Based on the forecast of Figure 2.10 (packet-based technologies in MBH networks) and by assuming that the majority of wire-line connections and ca. 5–10% of wireless (i.e. MWR-based) are leased, we may conclude that more than 50% of all MBH connections today, and in the near future, are leased.

There are several leasing cases as shown in Figure 5.14 below. If the leasing provider is incumbent and connection is TDM, i.e., E1/T1, it is often quite easy to lease the whole connection towards the mobile core starting from a BTS site. The access link can use DSL, fiber or MWR (case 2 in Figure 5.14). An independent transport provider can also offer leasing services starting from the BTS site, but they have to hire the copper line from the incumbent provider's access network. These non-incumbent providers also have access fiber, but often with a very limited coverage compared to the incumbent providers. A mobile operator may also build a partially owned/partial leased network, for example such that he only owns access and leases the rest (cases 3 and 4).

Optical fiber renting may be the most traditional way of leasing connections. So called dark fiber providers and data-bases where free fiber strands can be searched exist. In some countries there may be regulatory constraints in sharing parts of the mobile network. Obligations may be to give capacity (fibers) to other operators, too, while in some countries sharing may be limited to maintain competition.

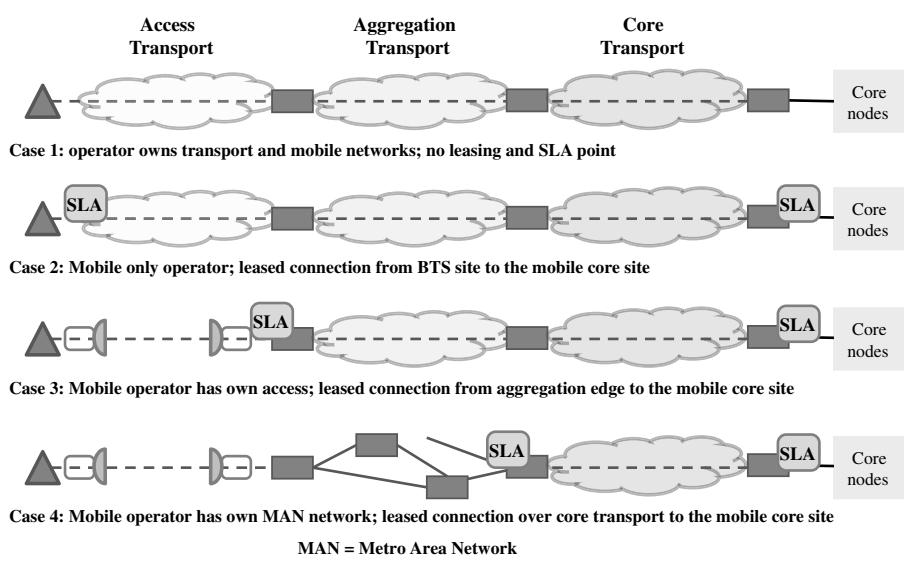


Figure 5.14 Examples of MBH leasing cases.

5.5.1 Ethernet Services and SLA's (MEF)

MEF (Metro-Ethernet Forum) is a global industry alliance whose mission is to accelerate the worldwide adoption of Carrier-class Ethernet networks and services. MEF has defined Carrier Ethernet as a ubiquitous, standardized, carrier-class service and network that is defined by five attributes that distinguish Carrier Ethernet from familiar LAN based Ethernet:

- Standardized Services.
- Scalability.
- Reliability.
- Quality of Service and
- Service Management.

The argument is that with this set of attributes the service can be guaranteed to be good enough for mobile backhaul while costs stay in control.

MEF develops technical specifications and implementation agreements to promote interoperability and deployment of Carrier Ethernet worldwide.

5.5.1.1 User Network Interface (UNI) and Ethernet Virtual Connection (EVC)

User to Network Interface (UNI) is the physical interface (port) between the service provider and customer, and it realizes the demarcation between the customer and service provider.

Ethernet Virtual Connection (EVC) is an association of two or more UNIs. EVCs are service containers connecting the customer sites (UNIs) (Figure 5.15).

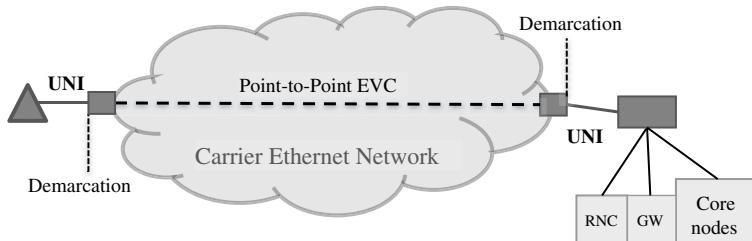


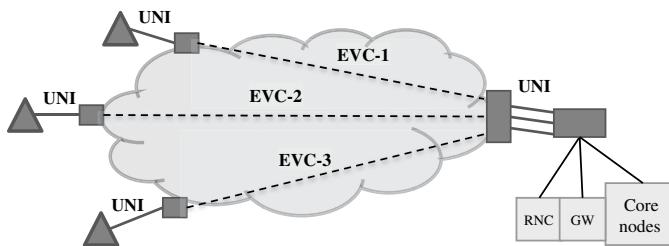
Figure 5.15 Example of point-to-point (E-Line) Ethernet Service.

5.5.1.2 Ethernet Service Definitions

MEF has specified and defined the Ethernet Services in two Technical Specifications; MEF 6.1 'Ethernet Services Definitions – Phase 2' [20] uses the service attributes and parameters defined in the MEF 10.2 'Ethernet Services Attributes Phase 2' [21 and 22] to create in the first the three generic Ethernet Service types (E-Line, E-LAN and E-Tree) which are further used to define the actual Ethernet services (like EPL, EVP-LAN etc.) (Table 5.10).

Table 5.10 MEF Ethernet Service Types and related Ethernet Services.

Service Type	Port-Based (All-to-One Bundling)	VLAN-Based (Service Multiplexed)
E-Line (Point-to-Point EVC)	Ethernet Private Line (EPL)	Ethernet Virtual Private Line (EVPL)
E-LAN (multipoint-to-multipoint EVC)	Ethernet Private LAN (EP-LAN)	Ethernet Virtual Private LAN (EVP-LAN)
E-Tree (rooted multipoint EVC)	Ethernet Private Tree (EP-Tree)	Ethernet Virtual Private Tree (EVP-Tree)

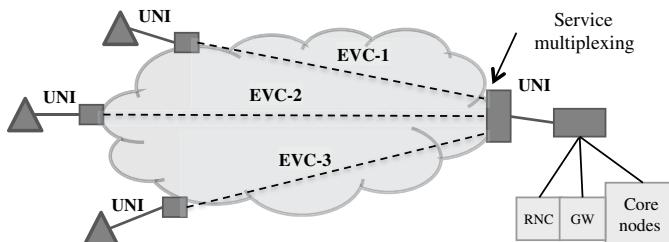
**Figure 5.16** Ethernet Private Line (EPL) Service.

Ethernet Privat Line Service:

The Ethernet Private Line (EPL) is a Port based point-to-point service between two UNIs. There is no service multiplexing possibility and therefore each individual EVC requires own port at the UNI (Figure 5.16).

Ethernet Virtual Private Line Service:

The Ethernet Private Line (EVPL) is a VLAN based service between two UNIs. Because of service multiplexing possibility at the UNI, less ports are needed e.g. at the controller/GW site, see Figure 5.17 below.

**Figure 5.17** Ethernet Virtual Private Line (EVPL) Service.

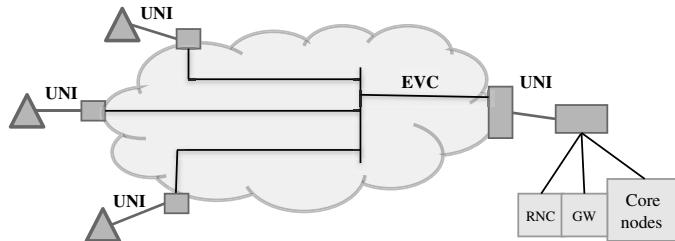


Figure 5.18 Ethernet Private LAN (EP-LAN) Service.

Ethernet Private LAN Service:

The Ethernet Private LAN (EP-LAN) is a Port based multipoint-to-multipoint LAN (Local Area Network) service between two or more UNIs of an EVC. There is no service multiplexing possibility and therefore each individual EVC requires own port at the UNI (e.g. in case there are own EVCs for user-, management- and control- planes) (Figure 5.18).

Ethernet Virtual Private LAN Service:

The Ethernet Virtual Private LAN (EVP-LAN) is a VLAN based multipoint-to-multipoint LAN service between two or more UNIs of an EVC. Because of service multiplexing possibility at the UNI, fewer ports are needed e.g. at the controller/GW site.

Ethernet Private Tree Service:

The Ethernet Private Tree (EP-Tree) is a Port based point-to-multipoint service between one root UNI and one or more leaf UNIs of an EVC. There is no service multiplexing possibility and therefore each individual EVC requires own port at the UNI (e.g. in such a case there are own EVCs for user-, management- and control- planes) (Figure 5.19).

Ethernet Virtual Private Tree Service:

The Ethernet Virtual Private Tree (EVP-Tree) is a VLAN based point-to-multipoint service between one root UNI and one or more leaf UNIs of an EVC. Because of service multiplexing possibility at the UNI, less ports are needed e.g. at the controller/GW site.

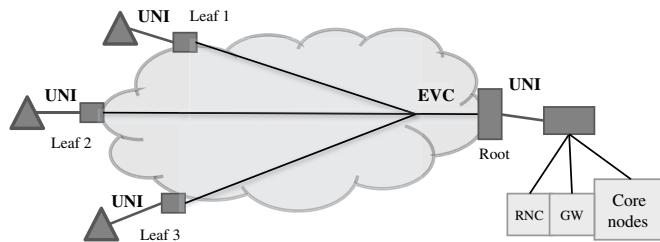


Figure 5.19 Ethernet Private Tree (EP-Tree) Service.

5.5.1.3 Service Level Agreement and Service Attributes

Service Level Agreement (SLA) is a commercial and juridical contract between the Service Provider and the customer (subscriber) specifying the service level (QoS) commitments and related business agreements (e.g. rental and compensation for SLA violation).

Service Level Specification (SLS) defines the service attributes (e.g. Bandwidth Profile and Service Performance) to be met for SLA conformance.

5.5.1.4 Bandwidth Profiles (BWP)

Bandwidth profile is a set of parameters defining e.g. how the traffic is limited and shaped at transport network ingress point (UNI). BW Profile parameters are Committed Information Rate (CIR), Excess Information Rate (EIR), Committed Burst Size (CBS), Excess Burst Size (EBS), Color mode (CM) and Coupling Flag (CF).

The ingress UNI Bandwidth Profile Enforcement (i.e. metering, color marking and policing operations) is implemented by so-called ‘two rate, Three Color Marker’ (trTCM) token bucket algorithm. The trTCM operating mode is determined with parameters CM (Color Mode) and CF (Coupling Flag):

- If Color Mode (CM) is set, the incoming user packet must be pre-colored as green or yellow. If CM is not set, the incoming user packet color is not checked.
- If Coupling Flag (CF) is set, the yellow user packets may utilize the unused green tokens. If CF is not set, the green and yellow token buckets work independently.

There are two practical operating modes (CM/CF-combinations):

- Color Blind Mode (mandatory), when CM is not set (making CF negligible): user packets are declared to CIR as long as there are green tokens available, then respectively to EIR, or discarded in case the total BW profile (CIR + EIR) is exceeded.
- Color Aware Mode, when both CM and CF are set: user packets are pre-allocated (-colored) either to CIR or EIR, unused green tokens may be utilized for EIR.

The CIR defines the rate up to which traffic is transmitted by the network with the required QoS, while the EIR defines the rate up to which the traffic exceeding the CIR is forwarded in the network without any SLA guarantees, i.e. the Performance Objectives are not valid for EIR and packets may be dropped in case of congestion inside the carrier’s network. The CBS is the maximum size of packet burst considered as conforming traffic and EBS is the size of packet burst above CBS that is transmitted with no guarantees. Based on these parameters the metering function at the UNI can take three different decisions about each packet (Color Blind Mode):

- If the rate of traffic is below CIR then the packet is considered as bandwidth profile conformant and it is marked as green and forwarded in the network.
- The packets exceeding the CIR but below EIR are marked as yellow and are forwarded in the network with no SLA guarantees (as ‘best effort’ traffic).
- Finally the packets exceeding even the EIR are marked as red and usually these packets are discarded immediately.

Bandwidth Profiles can be defined per-UNI (port), per EVC (VLAN) or per Class of Service (VLAN and CoS) basis (Figure 5.20).

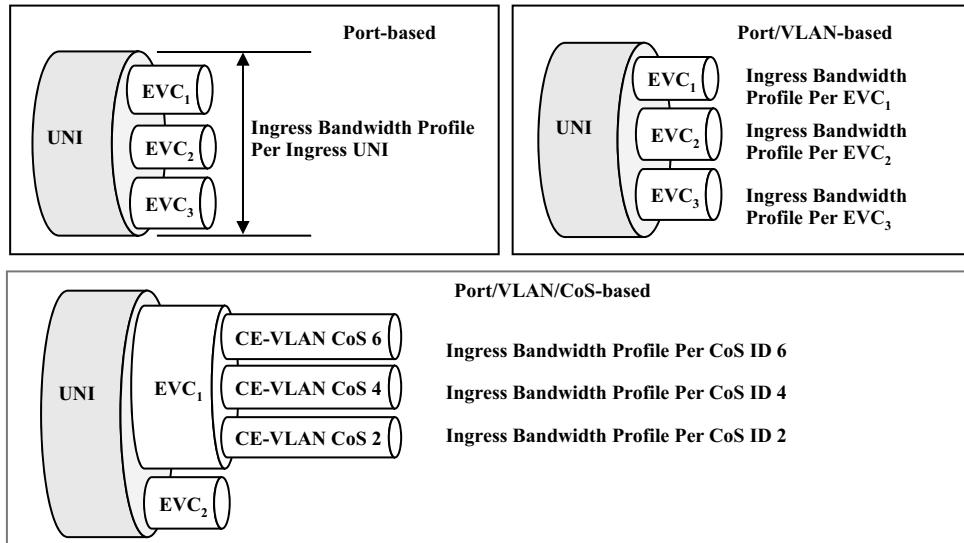


Figure 5.20 Three Types of Bandwidth Profiles Defined in MEF 10.2.

For more details about BW profile parameters and their use, see Section 8.3.10 (QoS with MEF Services).

5.5.1.5 Service Performance

Service Performance for a certain service class is defined with CoS Performance Objectives and in terms of Frame Delay (FD), Frame Delay Variation (FDV) and Frame Loss Rate (FLR).

In Table 5.11 Bandwidth Profile and Service Performance attributes have been combined to a four Service Class Model proposed for MBH [23].

Table 5.11 Service Class Model proposed for Mobile Backhauling [MEF22].

Service Class Name	Bandwidth Profile	CoS Performance Objectives		
		FD	FDV	FLR
Very High (H+)	CIR > 0 EIR = 0	A _{FD}	A _{FDV}	A _{FLR}
High (H)	CIR > 0 EIR ≥ 0	B _{FD}	B _{FDV}	B _{FLR}
Medium (M)	CIR > 0 EIR ≥ 0	C _{FD}	C _{FDV}	C _{FLR}
Low (L)	CIR ≥ 0 EIR ≥ 0*	D _{FD}	D _{FDV}	D _{FLR}

Notes:

A ≤ B ≤ C ≤ D and A_{FDV} is as small as possible

(*) both CIR = 0 and EIR = 0 is not allowed as this results in no conformant Service Frames

5.5.2 *Leased Ethernet Service Offering*

Ethernet services are being offered in Point-to-point (E-Line), Point-to-multi point (E-Tree) and in Any-to-any (E-LAN) configurations by all of the main service providers.

Today, the service providers are using a mixture of platforms to provide their Ethernet services, including legacy SONET/SDH platforms (for Ethernet Private Line services) and MPLS-based platforms (used in some cases just for E-LAN services). Many operators intend to migrate from SONET/SDH to MPLS-based platforms over the next three to five years.

The Ethernet SLA is typically based on the same performance objectives as the carrier's IP VPN service – mainly as Ethernet interworks with IP VPN – including Frame Delay, Frame Loss Rate, Frame Delay Variation service attributes. In many cases the offered SLA's performance objectives are relatively loose, probably indicating the currently often missing ability for end-to-end SLA verification. The trend, however, seems to be that the service providers are now deploying intelligent network interface devices (NID) on the customer site to improve site to site (end to end) SLAs with better measurement and monitoring.

The number of different CoS (Class of Service) types offered vary by country and Ethernet service provider. In the majority of cases three or four CoS types are available, but in some cases there is just one CoS or even up to seven ones. One example of four-CoS offering:

- Real Time CoS – e.g. for low latency VoIP.
- Critical CoS – for other low latency data applications such as video with jitter and latency guarantees.
- Priority CoS – for priority data applications, and
- Standard CoS type – suitable for low priority data applications such as e-mail and web browsing.

The availability of Metro Ethernet services is still limited in their geographical reach which is dependent on the amount of fiber available. Although the service providers are deploying more fiber in urban areas, they are also expanding the types of access method, including Copper (EoC, Ethernet over Copper), Microwave, SDSL and wireless broadband services as a low cost substitute for fiber deployment.

5.5.3 *IP as a Backhaul Service*

In addition to buying their own equipment or leasing Ethernet service, a mobile operator has an option to lease IP-connection as a backhaul service. Commercial IP transport services are mainly intended for high-capacity large-scale (WAN) networks, connecting LANs and international connections, but could be used for backhauling too. They may provide a feasible alternative for core-network connections. Access to IP service typically takes place with wire-line techniques, like OC-3/12/48 or 100/1G/10G Ethernet. Service providers are also introducing other access methods like copper (DSL) and microwave radio to make the offering more usable. The IP VPN service level agreement (SLA) is typically based on the same metrics as the carrier's Ethernet service.

One benefit of IP connections is that it is agnostic for lower level transport. The lower layer can be SDH/SONET, ATM or native Ethernet that gives flexibility for the provider. This flexibility is seen in capacity, in service provisioning time and in reaction time for customer

(backhaul) change requests. Easier management is beneficial for operating expenses. Agreements can be based on either fixed capacity or based on consumption. Service classes can be utilized to keep costs for best effort traffic at a reasonable level, while also providing carrier-class routing infrastructure on demand. When the connection is based on IP routing and there are numerous link techniques beneath the service is faster to configure and reconfigure.

5.6 Summary

In this chapter we have considered transport technologies, systems and other building blocks used in building a MBH network. Wireless solutions, i.e. MWR radios, are very common in lower access and especially in the last mile links to the base station sites where their share is over 50% of all transport solutions. Wire-line solutions are used for the rest of the base station sites Optical access is gradually increasing its role also in access – speed of transition is limited by the big investments needed in cable laying. In aggregation tiers and in the core of MBH network optical transport is the most common solution, and where capacities are high, DWDM system or a native packet optical layer is built. In new packet based MBH networks in all network tiers traffic aggregation, routing and protection (resilience) is based on Layer 2 or Layer 3 nodes, i.e. Ethernet switches and/or IP/MPLS routers; in the highest capacity backbones, optical layer may also partly take care of these functions.

Backhaul network is a great investment for a mobile operator. Utilization of legacy and making the right technology selections that also support the future 4G mobile broadband networks are the key for success. Backhaul technology must also provide opportunity for optimizing operating expenses and setting the right level of user experience. The last mile is the main differentiator between fixed transport and mobile backhaul networks, especially in the future when mobile cell sizes shrink.

References

- [1] ITU-T G.703, Physical/electrical characteristics of hierarchical digital interfaces
- [2] ITU-T G.707, Network Node Interface for the Synchronous Digital Hierarchy (SDH)
- [3] ITU-T G.709, Interfaces for the Optical Transport Network (OTN)
- [4] ITU-T G.781, Structure of Recommendations on Equipment for the Synchronous Digital Hierarchy (SDH)
- [5] ITU-T G.783, Characteristics of Synchronous Digital Hierarchy (SDH) Equipment Functional Blocks
- [6] ITU-T G.803, Architecture of Transport Networks Based on the Synchronous Digital Hierarchy (SDH)
- [7] ITU-T G.804, ATM cell mapping into plesiochronous digital hierarchy (PDH)
- [8] ITU-T G.821, Error performance of an international digital connection operating at a bit rate below the primary rate and forming part of an Integrated Services Digital Network
- [9] ITU-T G.826, End-to-end error performance parameters and objectives for international, constant bit-rate digital paths and connections
- [10] ITU-T G.871/Y.1301, Framework of Optical Transport Network Recommendations
- [11] ITU-T G.872, Architecture of optical transport networks
- [12] ANSI T1.105, SONET – Basic Description including Multiplex Structure, Rates and Formats
- [13] ANSI T1.105.02, SONET – Payload Mappings
- [14] ANSI T1.105.04, SONET – Data Communication Channel Protocol and Architectures
- [15] ANSI T1.105.06, SONET – Physical Layer Specifications
- [16] IETF RFC2615, PPP over SONET/SDH
- [17] IETF RFC1661, The Point-to-Point Protocol (PPP)
- [18] IETF RFC1662, PPP in HDLC-like Framing

- [19] Data-Over-Cable Service Interface Specifications: Cable Modem to Customer Premise Equipment Interface Specification. 2008
- [20] MEF 6.1, Ethernet Services Definitions - Phase 2, April 2008
- [21] MEF 10.2, Ethernet Services Attributes Phase 2, October (27) 2009
- [22] MEF 10.2.1, Performance Attributes Amendment to MEF 10.2, January (25) 2011
- [23] MEF 22.1, Mobile Backhaul Implementation Agreement Phase 2, January 2012
- [24] ITU-R P.530, Propagation data and prediction methods required for the design of terrestrial line-of-sight systems, (10/2009)
- [25] ISO/IEC 7498-1 : 1994(E) <http://standards.iso.org/ittf/>

Part II

Mobile Backhaul Functionality

6

Synchronization

Antti Pietiläinen and Juha Salmelin

Synchronization has been always one of the key issues in cellular networks and a lot of effort has been put to this single topic. Even conferences are arranged yearly in the area of synchronization in telecom. In this context, the word telecom refers ultimately to cellular networks in almost all cases.

6.1 Cellular Networks Synchronization Requirements

6.1.1 Frequency Accuracy

In cellular networks the frequency accuracy requirement is clearly higher than needed for eliminating interference between neighbouring radio bands. The user equipment (UE) measures the reception of neighbouring base stations in preparation of a possible handover from one base station to another. For this purpose, the user equipment has to tune to another radio frequency for a very short while to avoid interrupting voice and data streams. There is no time for fine tuning so the frequency references used by the base stations have to be almost exactly aligned. In the case of soft handover in WCDMA, UE receives signals from neighbouring base stations simultaneously at the same frequency, but with different codes. Consequently, the frequencies of the base stations need to be almost the same. The probability of receiving an undistorted signal from both base stations decreases as the frequency difference increases, dropping signal-to-noise ratio. When the difference grows further, the handovers fail completely.

When UEs move, the frequencies of the base stations seen by the UE shift due to the Doppler Effect. When approaching or receding from a base station, the frequency increases or decreases, correspondingly. In the present GSM systems the UE has to work with Doppler shift up to speeds of 250 km/h at 900 MHz, and 130 km/h at 1800 MHz. This corresponds to a frequency offset of around 250 Hz in both cases [1]. WCDMA systems have been designed to work at 250 km/h at 2100 MHz, which corresponds to a frequency offset of 591 Hz [2]. In all cases, 50 ppb (parts per billion, 10^{-9}) has been allocated for the wide area base station

Table 6.1 Frequency offset toleration and allowed velocities.

Type	GSM 900 MHz	GSM 1800 MHz	WCDMA 2100 MHz
Mobile must tolerate offset	295 Hz	340 Hz	591 Hz
Wide Velocity & Doppler	250 km/h & 250 Hz	130 km/h & 250 Hz	250 km/h & 486 Hz
0.05 ppm contrib.	45 Hz	90 Hz	105 Hz
Local Velocity & Doppler	205 km/h & 205 Hz	80 km/h & 160 Hz	196 km/h & 381 Hz
0.1 ppm contrib.	90 Hz	180 Hz	210 Hz

accuracy and 100 ppb for local area base station. From this information, one can create Table 6.1. Even though most of the budget is allocated for frequency shift caused by the Doppler Effect, a fair proportion is left for the inaccuracy of base station frequency.

The jitter and phase noise requirements for the signal are rather strict since the frequency must satisfy the 50-ppb requirement for as short time periods as 0.67 ms and 1 ms for WCDMA and LTE, correspondingly. The 50 ppb limit is also used by CDMA variants.

6.1.2 Time Accuracy

6.1.2.1 CDMA

CDMA variants are FDD (Frequency Division Duplex) systems that, nevertheless, require time synchronization. The different signals at the same frequency can be separated from each other by means of orthogonal Walsh codes. The orthogonality is preserved if the code sequences are aligned in time.

6.1.2.2 TDD Systems

TDD (Time Division Duplex) systems such as Mobile WiMAX, WCDMA TDD, and LTE TDD require time synchronization. There are several reasons to synchronize downlink and uplink transmissions in the network. Base stations may interfere with each other if one is listening for mobiles and simultaneously the neighbouring base station is transmitting at high power. User equipment (UE) might interfere with each other in a similar way if two UEs are near each other but served by different base stations. Due to the proximity, the power received from the nearby UE may be large enough that the use of different WCDMA code or OFDM channel would not create enough separation. Handovers are, of course, easier if the UE knows when to measure prospective base stations and does not need to interrupt transmission to do so. For decoding signal simultaneously from different base stations the base station transmission needs to be synchronized.

6.1.2.3 Single-frequency Network

Single-frequency network (SFN) technique is used in downstream direction for example in Multimedia Broadcast Multicast System (MBMS). More than one base station transmits the same signal. The transmission is synchronized in time and frequency so the same symbols are received at the same time from all transmitters.

6.1.2.4 Enhanced 911 Phase II OTDOA Location Determination

The E911 Phase II is scheduled to come into force in September 2012. E911 applies to both the USA and Canada. The location accuracy requirement is 50 meters for 67% of calls and 150 meters for 95% of calls for handset-based technologies. In OTDOA (Observed Time Difference of Arrival) the handset measures the timing of signals sent from the base stations making it handset-based (for network-based technologies the requirement is 100 meters for 67% of calls and 300 meters for 95% of calls). Electromagnetic fields travel 50 m in 170 ns and 150 m in 510 ns. It is difficult to give a definite time inaccuracy budget but allocating ± 200 ns for the time error of the reference in the base station could be within reasonable bounds.

Note, an ever increasing fraction of handsets is equipped with GPS receivers. The handsets can tell their location without OTDOA. Further, the location information of the GPS-equipped handset could be used to calibrate the OTDOA measurement. In this case it is adequate that the time error of the time reference remains stable to within e.g. 100 ns between updates from GPS capable handsets. It means that a stable frequency reference could be enough for reaching adequate accuracy.

6.1.2.5 Requirements Summary

Table 6.2 summarizes the synchronization requirements of various systems. A common way of defining time errors is to give symmetric error bounds around a nominal value, for example $\pm 3 \mu\text{s}$. 3GPP, on the other hand, defines time errors as maximum time differences between any pair of cells on the same frequency that have overlapping coverage areas. For making the error bounds commensurate with the common practice, the 3GPP numbers have been divided by 2 and prefixed by a \pm sign.

The requirements of CDMA, WiMAX, and OTDOA requirements are given for completeness sake and are not discussed elsewhere in this book.

As seen in Table 6.2 the most widely used cellular systems require ± 50 ppb accuracy at the air interface. Most of the budget is consumed within the base station. A typical value left for the frequency reference used for the radio frequency synthesis is about one third, 16 ppb.

6.2 Frequency Synchronization in TDM Networks

There are three TDM (time-division multiplexing) technologies currently in use, PDH (Plesiochronous Digital Hierarchy), SDH (Synchronous Digital Hierarchy) or SONET (Synchronous Optical Network), which is the North-American equivalent of SDH, and OTN (Optical Transport Network). The SONET requirements have been incorporated in the SDH specifications.

6.2.1 Synchronization Architecture in TDM Networks

Figure 6.1 depicts synchronization network architecture. SDH/SONET carries synchronization on the physical layer, i.e. the clock frequency of the physical signal is synchronized to a central reference. In the case of PDH, the different bit rate signals are not synchronized to each other but the client signal, for example E1, 2 Mbit/s, is carried across the network so that the

Table 6.2 Synchronization requirements of various systems [3–9].

	Frequency accuracy in Air Interface	Time of Day requirement in Air Interface
GSM	$\pm 50 \text{ ppb}$ Pico: 100 ppb	n.a.
WCDMA/TD-SCDMA (incl. I-HSPA)	$\pm 50 \text{ ppb}$ Medium/Local: 100 ppb, Femto 250 ppb	FDD: n.a. MBMS: $\pm 20 \text{ ms}$ TDD/TD-SCDMA: $\pm 1.5 \mu\text{s}$
CDMA IS-95A, 1x, 1xEV-DO	$\pm 50 \text{ ppb}$	$\pm 3 \mu\text{s}$ normal operation $\pm 10 \mu\text{s}$ holdover 8 h
Mobile WiMAX*	$\pm 15 \text{ ppb}$	$\pm 0.5 \dots \pm 0.7 \mu\text{s}$
LTE FDD	$\pm 50 \text{ ppb}^*$	n.a. MBMS SFN: up to a few μs $\pm 1.5 \mu\text{s}$ for $\leq 3 \text{ km}$ cell radius and home BS not using network listening as synchronization source $\pm 5 \mu\text{s}$ for $\geq 3 \text{ km}$ cell radius Home BS using network listening for synchronization Source, small cell $< 500 \text{ m}$ distance $\pm 1.5 \mu\text{s}$ Source, large cell $> 500 \text{ m}$ distance $\pm (1.33 \mu\text{s} + \text{propagation delay})/2$
LTE TDD	$\pm 50 \text{ ppb}$	$\pm 1.5 \mu\text{s}$ for $\leq 3 \text{ km}$ cell radius and home BS not using network listening as synchronization source $\pm 5 \mu\text{s}$ for $\geq 3 \text{ km}$ cell radius Home BS using network listening for synchronization Source, small cell $< 500 \text{ m}$ distance $\pm 1.5 \mu\text{s}$ Source, large cell $> 500 \text{ m}$ distance $\pm (1.33 \mu\text{s} + \text{propagation delay})/2$
OTDOA, E911 Phase II Sep 2012**	n.a.	In the order of $\pm 200 \text{ ns}$

* WiMAX parameters: 1024 OFDM carriers, BW 10 MHz, Cyclic prefix ratio 1:8, RF carrier 3.5 GHz

** If the location determination system can be calibrated using handsets equipped with GPS receivers, then the time accuracy requirement reduces into frequency stability requirement where the phase drift between calibrations may not be more than about 100 ns.

frequency of the signal is preserved. OTN is similar to PDH, where the client TDM signal preserves frequency when carried over the network.

Usually networks are synchronized centrally by a primary reference clock. The mobile network nodes like base stations and controllers receive synchronization from the transport network. Ring structures can be used for protection purposes. For instance in Figure 6.1 the synchronization chain on the right side ring ends before the bottom node of the ring. However, if a link breaks in the left side of the ring the synchronization chain coming from the right side will be forwarded to the nodes in the left side that would otherwise lose synchronization. PDH and SDH are described below.

6.2.2 PDH

Traditionally the clock of one end of the link is free-running, that is, it is not synchronized by the network. The other end is locked to the first end. However, the maximum allowed

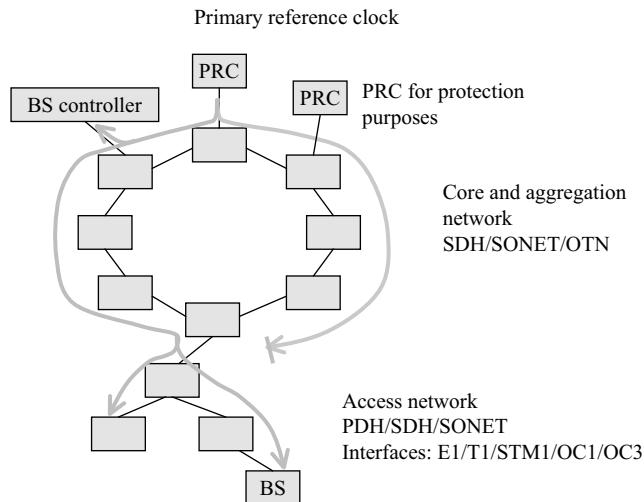


Figure 6.1 Synchronization network architecture.

frequency error is limited so that the receiver can lock in to the signal for being able to receive the data. For example, the frequency accuracy limit is ± 50 ppm for 2.048 Mbit/s signal and ± 32 ppm for 1.544 Mbit/s signal. Another and more important reason for frequency accuracy limit is the requirements of the multiplexing process where tributaries are combined into larger bit rate signals. The larger rate signal is framed so that there are time slots reserved for the tributaries. For filling almost exactly the reserved times slots and never exceeding the maximum bit rate reserved for each tributary, the frequency error limits need to be accurately defined. Exceeding this limit would cause loss of data. Today, most of the PDH networks obtain the frequency reference from a central reference as in SDH, see below. In cellular networks, this is a requirement. ITU-T recommendations related to PDH synchronization are G.823 and G.824. Especially in the case of 2 Mbit/s PDH hierarchy multiplexing to higher PDH bitrates has disappeared to a large extent. Instead, the signals are multiplexed directly into SDH containers.

6.2.3 SDH/SONET

When designing SDH (SONET in North America) networks, one target was easing up multiplexing and demultiplexing. In each PDH multiplexing level the upper level knows what bits in the lower bit-rate time slots are used for demultiplexing the tributaries at the other end. However, the next multiplexing level no longer knows this. Thus, each node needs a multiplexer mountain to add or drop a small bit rate signal. However, each SDH/SONET multiplexing level knows exactly the used bits by the client signals. By synchronizing all SDH/SONET layers tightly together, the frequency of time slips compared with the nominal frequency is very small. The primary reference clock PRC (North America: primary reference source, PRS) may deviate from the nominal frequency only by ± 0.01 ppb. PRC is defined in ITU-T recommendation G.811.

In the case of 2.048-Mbit/s hierarchy, clocks in SDH nodes are called SECs (SDH equipment clock) and are specified in G.813. There may be maximally 20 in a row. The total number of SECs over a synchronization chain may be 60 if there is an SSU (Synchronization supply unit), G.812 Type I, clock between chains of maximally 20 SECs. The total numbers of SSUs is 10. The SDH equipment clocks in the case of 1.544-Mbit/s hierarchy are SEC option 2 clocks. The jitter and wander network limits are specified in G.823, wander for 2.048-Mbit/s hierarchy, G.824, wander for 1.544-Mbit/s hierarchy, and G.825, SDH jitter specifications. 0.171 and 0.172 describe testing equipment for PDH and SDH, respectively.

The ring protection occurring for synchronization trail, as discussed in relationship with Figure 6.1 requires a messaging mechanism where a clock losing synchronization indicates to the next clock that the quality of the clock is not acceptable. Consequently the next clock will know to select another clock input if available. For indicating clock quality, ITU has defined SSMs (synchronization status messages) and clock selection principles in G.781. Functional blocks are defined in G.783. The mapping of SSMs in SDH frames is specified in G.707/Y.1322.

6.2.4 ATM

ATM (Asynchronous Transfer Mode) was chosen as the transport protocol for WCDMA and it dominated for the first ten years. For transporting ATM cells over transport network they are encapsulated in PDH or SDH/SONET frames and these layers also carry synchronization. In this sense there is no difference compared with SDH or PDH synchronization. Since ATM cells are usually encapsulated directly in SDH/SONET frames at the RNC (radio network controller), there is no end-to-end PDH layer left. This leaves out an option available in PDH where accurate frequency of a mobile operator could be transported asynchronously over a SDH/SONET network of a transport provider even if the SDH/SONET network is not synchronized well.

6.2.5 OTN

Optical Transport Network (OTN) was developed to carry 2.5-Gbit/s and higher data streams over optical wavelength carriers. Other transport protocols are then clients of OTN. Since a synchronization hierarchy was already defined for SDH, instead of defining a new synchronization hierarchy for OTN, the layer was made transparent for synchronized client signals. As a result, SDH clients could pass tens of OTN nodes without accumulating too much jitter and wander.

6.3 Frequency Synchronization in Packet Networks

Since the dawn of internet protocol (IP), IP packets have been transported over telecom networks. Until the end of the last century, the physical layer of inter-office connections was one of the traditional telecom technologies, analog, PDH, or SDH/SONET. However, the introduction of optical Ethernet interfaces towards the end of the 1990s started to change the scene. Now, the most common interface in telecom routers is Ethernet. Mobile networks long

remained the last fortress of legacy transport since the bandwidth requirement of cellular base stations remained low. However, then HSPA (high-speed packet access) combined with fixed monthly fees exploded the bandwidth demand and PDH access did not scale up. Ethernet came as the saviour in the price per bit problem – but associated with another problem: synchronization.

Ethernet has been designed as a low-cost plug and play technology. Combined with IP friendly adaptation to higher layers, Ethernet has become the dominant link layer and physical layer technology in packet networks. Within the Ethernet scene bridges have almost completely overcome repeaters. Correspondingly, Ethernet itself spans just a single link, terminating the clock at the receiving port. The idea of connecting the received clock to outputs (as in SDH/SONET) had been around since the early years of the past decade and standardizing Synchronous Ethernet began mid-way through. However, it takes a long time from starting a standardization project until all links in a network segment support the standardized protocol. Luckily, regarding the use of Ethernet for cellular transport, the development of packet timing technologies for telecom had already started earlier, such as adaptive clock recovery (ACR), precision implementations of Network Time Protocol (NTP), and telecom oriented options for Precision Time Protocol (PTP). These protocols can operate independently of the physical layer and thus carry synchronization without on-path support.

6.3.1 ACR (*Adaptive Clock Recovery*)

Ethernet was first introduced into cellular backhaul when the endpoints, for example base stations were still PDH based. As a consequence there was a need to carry PDH frames over packet network. Pseudowire specifications such as SaToP [10] and CESoPSN [11] are used. In both cases, the packet rate is fixed, typically 1000 pps, which is enough information for recovering synchronization at the receiving end. Both specifications have the option of using RTP (Real Time Protocol) for time stamping the packets. However, this option is not necessary for clock recovery because of the fixed packet rate.

6.3.2 NTP

NTP is a quarter of a century old protocol designed for synchronization of time over the internet. It has been updated several times. Version 4 was published in June 2010 [12]. NTP has been designed to run in software that has similar access to the network layer as other software applications. According to the latest version the new algorithms extend the potential accuracy over LANs to tens of microseconds.

This accuracy would consume a meaningful proportion of the whole uncertainty budget considering frequency synchronization of base stations. Therefore, the traditional NTP is a slightly too inaccurate method with which to consider synchronization of mobile networks. However, NTPv4 mentions the possibility of accessing the physical layer and defines a location in the frame where the timestamps are associated. In recent years, NTP equipment with hardware assist has emerged.

The protocol stack of a telecom implementation is shown in Figure 6.2. The base stations have, in practice Ethernet interfaces. NTP utilizes UDP transport layer over unicast IP. The

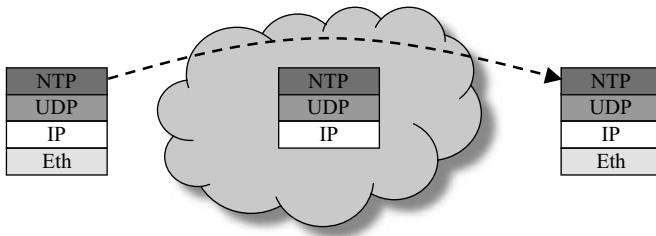


Figure 6.2 NTP protocol stack.

three top layers are carried end to end while the Ethernet layer is typically terminated in IP and MPLS routers.

Figure 6.3 depicts the messaging. If the NTP algorithm is used, the slave polls the master in bursts of eight messages where the individual messages are separated by two seconds from each other. The algorithm selects the best information available from the burst for maximizing accuracy. The interval between the bursts can vary between 16 s and 36 h. The time stamps $t_1 \dots t_4$ correspond to the transmit and receive moments of the two messages.

Assuming that the forward and reverse delays are equal the time error of the slave is

$$\Delta = t_{slave} - t_{master} = \frac{(t_1 - t_2) + (t_4 - t_3)}{2}. \quad (6.1)$$

The time error is used by the clock tuning algorithm. As the packet rates are low, there may be proprietary implementations that use higher packet rates for achieving adequate accuracy for base stations.

6.3.3 PTP Protocol

PTP (Precision Time Protocol) [13] was originally designed for reaching very high time synchronization accuracy, for example, for industrial automation and measurement technology. NTP could not adjust to the requirements and therefore PTP protocol was designed. The

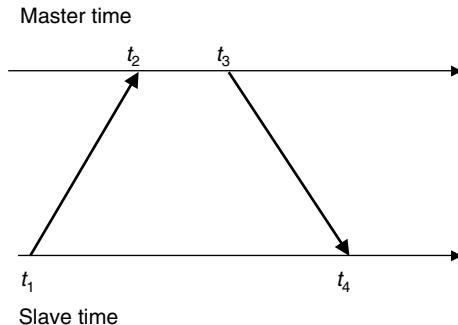


Figure 6.3 NTP timing.

first version of PTP included the necessary definitions for hardware time stamping (essentially time stamp point in the message), on-path support (boundary clocks, BC), automatic clock hierarchy build-up (best master clock algorithm, BMCA), and faster maximum packet rate than in NTP, 1 pps.

When the telecom synchronization industry started to develop packet timing for cellular backhaul, NTP was a relatively good option. However, for remaining truly conformant with the NTP standard, the packet rate and algorithm had constraints. Therefore, several telecom companies participated in the development of PTPv2. On the one hand, PTP protocol messaging could be tailored for high packet rate required for optimal performance. On the other hand, PTP has no restrictions on the actual clock algorithm, allowing the development of algorithms specifically designed for example for base station synchronization. Having said that, proprietary NTP algorithms could still achieve the same accuracy targets, however, as PTP standard has the means to also reach high time accuracy by using on-path support defined by the standard, PTP proved to be more future proof than NTP.

PTPv1 used a reserved multicast address, which is convenient when all nodes support the protocol (full on-path support). The nodes do not need to know their neighbours' IP addresses and clock hierarchy can be created based on physical connectivity instead of protocol addresses. This single-link multicast scheme is familiar from several other protocols that include hierarchy build-up, such as routing protocols on Layer 3 and spanning tree protocol on Layer 2. At the time when a packet timing protocol was needed for telecom, on-path support could not be dreamt of. As a consequence, IP unicast operation was developed and the packet rate was increased to cover all telecom needs, typically up to 64 pps. As there were many different interests associated with PTPv2, the standard became flooded with options, almost doubling the page count compared to version 1. For helping telecom implementers, an informative Annex A.9, 'Recommendation for implementations in unicast networks or networks with non-PTP bridges and routers', was prepared. It was a last-minute addition, but successful, as the first telecom implementations to reach mass roll-outs of PTP clocks were based on the annex.

Figure 6.4 depicts an example of the protocol stack in the case of frequency transport without on-path support. In this case PTPv2 Annex D, 'Transport of PTP over User Datagram Protocol over Internet Protocol Version 4', is utilized. As in NTP, the three topmost protocols

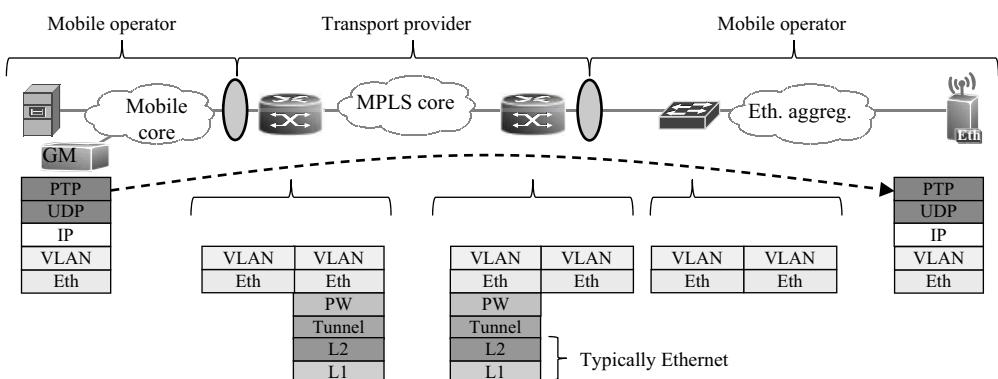


Figure 6.4 Protocol stack for frequency synchronization without on-path support.

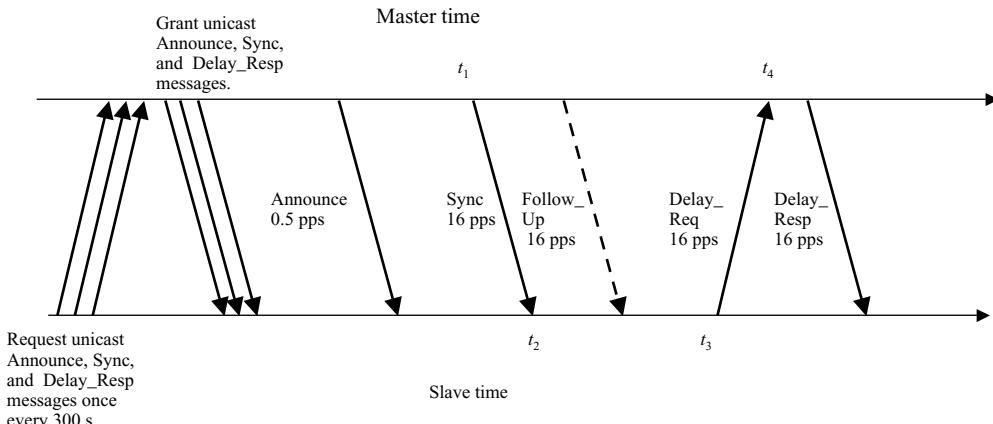


Figure 6.5 PTP protocol messaging.

are carried end-to-end. In the example, the base station transport of the mobile operator is based on Ethernet. The mobile operator has leased L2VPNs from a transport provider, which runs an MPLS network. Often the mobile operator has a small number of IP routers on the way, in which case the Ethernet layer does not span end-to-end.

The synchronization principle in PTP is the same as in NTP. There is one time stamped message upstream and one downstream, see Figure 6.5. From the beginning, NTP has utilized a client-server scheme of operation whereas PTP has not. Originally, instead of waiting for a request from a slave, a PTP master sends Sync packets to the multicast address as long as it does not encounter another master with higher priority. For enabling client-server operation for unicast PTP, a new message class, signaling messages, was utilized to request unicast messages from the master. The Announce messages advertise the clock class of the master, and reception of these messages is necessary before a slave's state machine may enter the slave state. After entering the slave state, the slave starts to use the timing messages for synchronization.

The three unicast requests may be contained in a single signaling message if found appropriate since the requests are in the form of 10 byte TLVs (type length value) and a number of them can be inserted in a single message.

The Follow_up messages are needed if the master does not have a real-time HW timestamper. When utilizing Follow_Ups, it is adequate to detect when the timing packet transmission occurred and then enter the information into a Follow_Up message by means of a software process. There is no follow-up for the Delay_Req message since it is adequate that the *slave* knows the accurate transmission time of the Delay_Req by the time the algorithm is ready for the time error calculation. The two messages whose transmission and receiving times have relevance are event messages and the rest are general messages. Even though the Follow_Up messages contain time stamps they are not event messages since the time they themselves are sent or received is not needed.

The protocol allows some creativity in using the timing messages. One can decide to use all four time stamps or just two time stamps contained in one of the two event messages. The least bandwidth and processing power is consumed if only Sync messages are used. Using both directions usually yields the best result but consumes the most resources. Using upstream

Table 6.3 PTP message types and protocol layer information.

Message type	Typical rate used by PTP clocks	Destination IP address	Destination UDP port	Source IP address	Source UDP port*
Signaling message for requesting unicast transmission	e.g. once every 250 s	Master port's unicast addr.	320	Slave port's unicast addr.	320
Announce	1/2 pps	Slave port's unicast addr.	320	Master port's unicast addr.	320
Sync	16 ... 64 pps	Slave port's unicast addr.	319	Master port's unicast addr.	319
Follow up	16 ... 64 pps	Slave port's unicast addr.	320	Master port's unicast addr.	320
Delay request	16 ... 64 pps	Master port's unicast addr.	319	Slave port's unicast addr.	319
Delay response	16 ... 64 pps	Slave port's unicast addr.	320	Master port's unicast addr.	320

* Not specified in IEEE 1588 but usually the same port number is used for the source as for the destination.

only may be attractive since this direction usually has a substantially lighter load than the downstream direction.

The error functions that a phase feedback servo tries to minimize to zero in downstream, upstream, and two-way cases are given in Equations 6.2, 6.3, and 6.4, respectively.

$$\Delta = t_2 - t_1 \quad (6.2)$$

$$\Delta = t_3 - t_4 \quad (6.3)$$

$$\Delta = \frac{(t_2 - t_1) + (t_3 - t_4)}{2} \quad (6.4)$$

In the one-way cases a constant time error of one-way delay remains when the error function is minimized. This does not matter, as long as only frequency synchronization is targeted.

Table 6.3 summarizes the messages used in the unicast model. The UDP port number 319 indicates the event messages and port number 320 maps to general messages.

6.3.4 ITU PTP Telecom Profile for Frequency Synchronization

The telecom profile is very similar to Annex A.9. The messaging is the same as shown in Figure 6.5. The protocol stack is PTP/UDP/IPv4. There are some differences, which are listed in Table 6.4.

In the default profile described in IEEE 1588-2008 the redundancy of masters is based in a system where only one Grandmaster is active at a time. If the preferred GM fails, then the second one in priority order takes over. In the unicast model where each slave separately loads

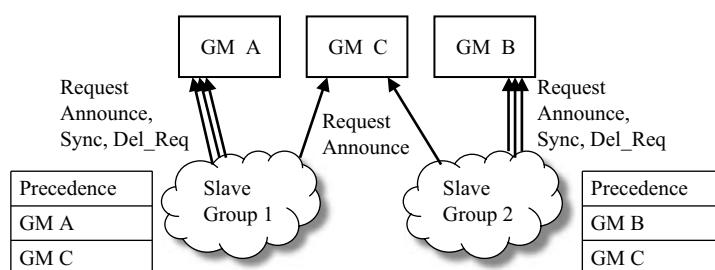
Table 6.4 Differences in PTP messages between the two PTP profiles used in telecom.

	Annex A.9	G.8265.1
profileIdentifier (first six octets of the sourcePortIdentity field in all PTP messages)	00-1B-19-00-01-00*, Default PTP profile for use with the delay request-response mechanism. Version 1.0**	00-19-A7-00-01-00, ITU-T PTP Profile for Frequency distribution without timing support from the network (unicast mode) Version 1.0**
Domain number (in the header of all PTP messages)	Default: 0, Configurable range 0–127	Default: 4, Configurable range 4–23
clockClass (first byte of the grandmasterClockQuality field in the Announce messages)	6–58	80–110 G.8265.1 defines mapping of G.781 Quality Levels to the PTP clockClass values
Request unicast transmission, duration of grant	10–1000 s, default 300 s	60–1000 s, default 300 s
Announce message	1/8–8 per second, default 1/2 per second	1/16–8 per second, default 1/2 per second
Sync message	1/2–128 per second, default 16 per second	1/8–128 per second, no default specified
Delay_Resp message	1/64–128 per second, default 16 per second	1/8–128 per second, no default specified

* Annex A.9 also allows the use of peer-delay mechanism with profile identifier 00-1B-19-00-02-00.
This option is probably not implemented in Annex A.9 compliant equipment.

** Note, this version number refers to the profile. The PTP version number is 2 in both cases.

a Grandmaster, there is a need to operate multiple Grandmasters simultaneously to serve all base stations, see Figure 6.6. Further, it might be decided that an extra Grandmaster operates as standby for the active Grandmasters. G.8265.1 defines a redundancy architecture where the slaves have a precedence table of GMs. They poll all GMs for Announce messages in order to know whether the clock quality of each potential master is adequate. If the preferred GM fails, then the slave starts to request timing messages from the redundant Grandmaster.

**Figure 6.6** Redundancy architecture.

The polling of Announce messages by two groups of slaves may burden a GM too much if the groups are large. Therefore, the standard allows the option of not polling the spare master unless the preferred master fails. The redundancy mechanism described in G.8265.1 can be implemented in Annex A.9 slaves. Thus, functionally, the two versions could be identical.

The telecom profile specifies exclusively the use of IPv4. Thus, IPv6 is not allowed. However, it is straightforward to implement the protocol using IPv6 based on Annex E of IEEE 1588-2008.

6.3.5 Synchronous Ethernet

Synchronous Ethernet has been designed to a certain extent along the same principles as SDH. Therefore, similar synchronization network structures as in the case of SDH, see Figure 6.1, can be expected.

IEEE 802.3 working group standardizes Ethernet and has specified limits, for example 100-ppm clock accuracy for most Ethernet interfaces. ITU has worked within this space when specifying increased accuracy and forwarding synchronization from input to output. For indicating clock accuracy, synchronization status messages are needed as in the case of SDH. Using IEEE802.3 slow protocol family is a suitable option since 802.1 bridges do not forward these messages. Nowadays, IEEE802.3 allows the creation of organization specific slow protocols (OSSP). ITU has specified one for carrying the status messages and it is called ESMC (Ethernet Synchronization Messaging Channel). The ESMC principles are the same as for SSM for SDH. The synchronization protection switching and possible spanning tree protocol protection switching for the data traffic run independent of each other.

G.8261 defines network limits in terms of MTIE and TDEV, which are described later. EEC-Option 1 clock (Synchronous Ethernet equipment clock) has the same specification as SEC (SDH equipment clock). EEC-Option 2 clock has the same limits as SEC Option 2. Recommendation G.8262 Specifies the EEC clock and G.8264 describes the ESMC messages.

The allowed number of consecutive synchronous EEC Option 1 nodes in the synchronization path is the same as in the case of SEC Option 1 clocks, 20. Then, in principle an SSU should be added to connect to the next EEC.

6.3.5.1 Over Optical Transport Network, OTN

Similar to SDH as a client signal, OTN was made transparent for Ethernet clients and defined in such a way that Synchronous Ethernet timing signals remained within acceptable jitter and wander limits. These aspects are discussed in [14]. It has turned out that Synchronous Ethernet can traverse tens of OTN nodes without accumulating too much jitter or wander.

6.3.5.2 One-way Synchronous Ethernet Links

Most of the Ethernet interfaces are such that the direction of synchronization can be changed quickly as required by the synchronization protection switching described in Figure 6.1. However, there are some exceptions. In the 1-Gbit/s and 10-Gbit/s copper interfaces (1000BASE-T and 10GBASE-T) the two ends of the links will have the same frequency.

One end will always be selected as the master and the other as a slave. The direction of the synchronization can be managed by setting the auto-negotiation parameters of the interfaces. However, changing the direction would require setting the parameters in a different way and resetting the interface. The traffic might be interrupted for many seconds when several such interfaces in a chain would be reset. Therefore, the directions are set semi-permanently. Note, the one-way issue does not concern the corresponding optical interfaces. However, PON (passive optical networks) links are also one-way in terms of synchronization because the head-end is always synchronization master for running the PON transport protocol. This is not a problem, however, since PONs are used in the edge of the network oriented in the right direction concerning the synchronization hierarchy.

6.3.6 Chaining Different Synchronization Technologies

When chaining different synchronization technologies, e.g. packet timing and physical layer timing, one must take into account that each island consumes a proportion of the total wander budget. When considering chaining different synchronization technologies, the reader is encouraged to study the different deployment cases in G.8261.

6.3.7 Summary of ITU Recommendations Related to Frequency Synchronization in Packet Networks

6.3.7.1 G.8260, Definitions and Terminology for Synchronization in Packet Networks

Considers mostly packet timing since G.810, ‘Definitions and terminology for synchronization networks’, contains most of the definitions related to physical layer frequency synchronization applicable also for Synchronous Ethernet. G.8260 includes an appendix about packet timing metrics.

6.3.7.2 G.8261/Y.1361, Timing and Synchronization Aspects in Packet Networks

‘The handbook of timing in packet networks’. The recommendation describes timing in circuit emulation service (CES), Synchronous Ethernet, and packet timing including network limits for different deployment cases. Further, it defines test cases for packet timing.

6.3.7.3 G.8261.1, Packet Delay Variation Network Limits Applicable to Packet Based Methods (Frequency Synchronization)

This specifies the network limits on packet delay variation depending on reference model. The first version bases the limits on Hypothetical reference model 1 (HRM-1), a network of 10 nodes between the master and slave. The links are otherwise 1 Gbit/s fiber optic links except three links that are 10 Gbit/s fiber optic links.

6.3.7.4 G.8262/Y.1362, Timing Characteristics of Synchronous Ethernet Equipment Slave Clock (EEC)

The recommendation is very similar to G.813, Timing characteristics of SDH equipment slave clocks (SEC). The intent of synchronous Ethernet is to interoperate with existing synchronization networks based on G.813.

6.3.7.5 G.8263, Timing Characteristics of Packet Based Equipment Clocks (PEC) and Packet Based Service Clocks (PSC)

The recommendation describes PEC-M, Packet master clock, PEC-S, Packet slave clock, and PEC-B, Combined packet slave clock and packet master clock.

6.3.7.6 G.8264/Y.1364, Distribution of Timing Information Through Packet Networks

The recommendation describes the Ethernet Synchronization Messaging Channel and reference source selection mechanism.

6.3.7.7 G.8265/Y1365, Architecture and Requirements for Packet-based Frequency Delivery

Recommendation describes the general architecture of frequency distribution using packet-based methods. This is applicable to PTP and NTP.

6.3.7.8 G.8265.1/Y.1365.1, Precision Time Protocol Telecom Profile for Frequency Synchronization

ITU PTP profile definitions for frequency delivery

6.3.7.9 Synchronous Ethernet Over Optical Networks and Testing

G.709/Y.1331, Interfaces for the Optical Transport Network (OTN) about mapping clients to OTN (optical transport network) and the corresponding jitter and wander specifications G.8251, The control of jitter and wander within the optical transport network (OTN), have been developed to allow passing synchronous Ethernet across OTN with minimal jitter and wander accumulation. O.174 Jitter and wander measuring equipment for digital systems which are based on synchronous Ethernet technology completes the specifications related to Synchronous Ethernet.

6.3.8 TICTOC

The TICTOC (Timing over IP Connection and Transfer of Clock) working group in IETF has been running since 2006. The working group is concerned with highly accurate time and frequency distribution over native IP and MPLS-enabled IP packet switched networks. The charter overlaps with the agenda of ITU-T Question 13 (Network synchronization and time distribution performance) of study group 15. It has taken some time to find topics that naturally fall into the realm of an IETF working group and the early documents have expired. Currently TICTOC has two active working group documents ‘Transporting PTP messages (1588) over MPLS Networks’ and ‘Precision Time Protocol Version 2 (PTPv2) Management Information Base (MIB)’.

The first one defines the method for transporting PTP messages over an MPLS network. The method allows for the identification of PTP messages at the port level to allow for port level processing of these PDUs in MPLS equipment. Setting up corresponding pseudowires and label switched paths are described in related Internet-Drafts. In the case of frequency transport

the method does not usually need to be used, because adequate synchronization quality can be achieved by carrying PTP over the same pseudowires as data. In the case of time/phase transport or very strict frequency synchronization needs, the method may be useful for ensuring symmetric paths and filtering PDV on the way. However, in the case every node supports the protocol, then the forthcoming ITU PTP time profile, see Section 6.7.2, is probably a better choice. The MIB document defines managed objects used for managing PTP devices. The managed objects can then be used with network management protocols.

6.4 Synchronization Metrics for TDM and Synchronous Ethernet

A clock based on the physical layer signal is depicted in Figure 6.7. A comparator compares the clock input and clock output. In the case of phase comparator, the output will, in the long run, output the same number of clock cycles as goes in. The clock is thus called a phase-locked loop (PLL). Another method is to compare frequencies of the input and output. Because phase error accumulation is not an issue for base stations, as long as the frequency error remains within 16 ppb, phase lock is not necessary.

As the TDM specifications allow large short-term frequency variations, see Figure 6.9, a low-pass filter is needed to do averaging of the input phase or frequency. The most commonly used reference, 2 Mbit/s traffic signal needs about 1000-s averaging. Thus, the filter time constant should be of the same order of magnitude.

6.4.1 Stability Metric MTIE

Before going into the metric, let's quickly remind ourselves about the relationship between frequency and phase errors. Either or both terms will be used in this chapter depending on the context. Simply, the accumulated phase error over a time period is the average frequency error times the length of the time period. Frequency error is calculated as the difference between the measured and reference frequency, divided by the reference frequency. Since the nominator and denominator both have the unit Hz, the ratio between them has no unit, while the unit of phase error is the second. The acronyms ppm (parts per million) and ppb (parts per billion) are typically used to give the magnitude of the frequency error.

Sometimes it is more natural to use phase and sometimes frequency error even though there is a direct relationship. For example, it is more natural to measure very short-term noise as phase error since the phenomenon is basically the time inaccuracy when a rising or falling edge actually occurs. Long-term variations are more frequency error oriented since an oscillator can tick too fast or too slow for long periods and the accumulated phase error can be up to thousands of clock cycles.

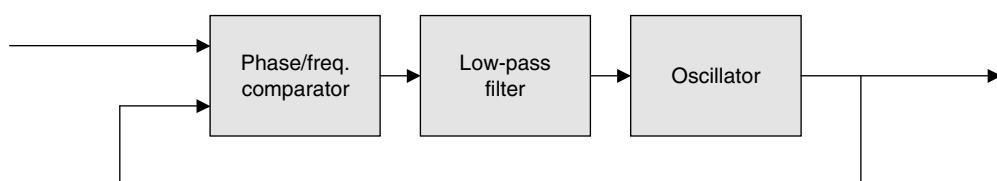


Figure 6.7 Function model of a TDM based clock.

Clock sources have short- and medium-term frequency/phase noise. On the other hand, as telecommunication signal traverses across a network, several mechanisms will cause short- and medium-term fluctuation to the frequency/phase compared to the original signal. As long as the accumulated phase error is not too large, the fluctuations can be tolerated even if it means significantly larger short-term frequency error than allowed in average. Therefore, a metric is needed, for specifying different amounts of fluctuations depending on the observation interval. In telecom, the noise phenomena have been divided into two classes, jitter and wander. Jitter is phase/frequency noise with periodicity below 0.1 s and wander above 0.1 s.

The TDM (time division multiplexing) synchronization metrics are described in ITU-T G.810. The most important metric is MTIE (maximum time interval error), which measures wander. It describes the time drift of a clock, compared with the reference. The x-axis of an MTIE graph is τ , observation interval width, measured in seconds. The y-axis $MTIE(\tau)$ is the maximum time drift that has occurred within the observation window. Thus, MTIE is a phase metric and correspondingly the unit is second. For example, for finding out the maximum time drift within a 10-s observation window, the 10-s window is slid over the whole measurement data of, for example 24 hours. The time difference between the maximum and minimum time error within the observation window is recorded at each position the window takes over the 24-hour data. The maximum recorded time difference represents MTIE at that particular value of τ (observation interval length). MTIE is estimated from discrete measurement samples as

$$MTIE(n\tau_0) \cong \max_{1 \leq k \leq N-n} \left(\max_{k \leq i \leq k+n} x(i) - \min_{k \leq i \leq k+n} x(i) \right), \quad n = 1, 2, \dots, N-1 \quad (6.5)$$

where $n\tau_0 = \tau$ is the observation interval (or window), $x(i)$ is the time error, $n+1$ is the number of samples in the observation interval, τ_0 is the sampling interval, and N is the total

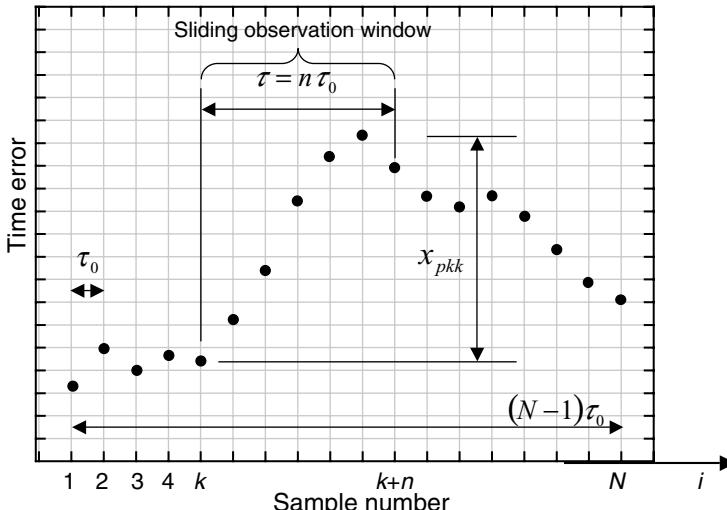


Figure 6.8 The variable x_{pkk} is the peak-to-peak x_i within the k -th location of the observation window. $MTIE(\tau)$ is the maximum x_{pkk} found from the $N-n$ locations the observation window takes when it is slid over the data.

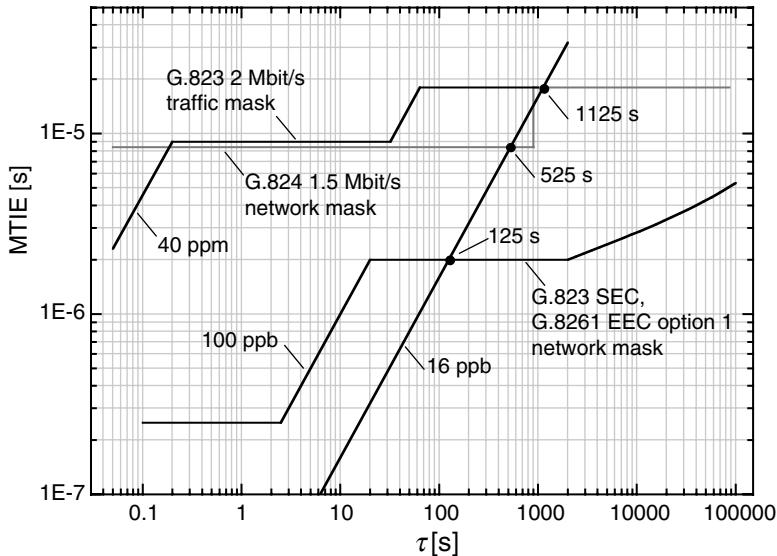


Figure 6.9 Wander masks for various TDM signals.

number of samples in the measurement. Figure 6.8 shows the sliding observation window of length $\tau = n\tau_0$ that is slid over the N samples of data.

As the name indicates, MTIE determines *maximum* drifts that have occurred during a measurement. Therefore, it is particularly handy for indicating limit values in the form of masks. Figure 6.9 shows a few of the most commonly used masks. Namely, 2.048 Mbit/s traffic mask, 1.544 Mbit/s network mask and SEC (SDH equipment clock) mask. The masks indicate maximum allowed time drifts compared to the reference. When looking carefully at the specifications, the metric for 2.048-Mbit/s wander is actually MRTIE (maximum relative time interval error). However, if the 2.048-Mbit/s signal is synchronized to a PRC instead of being free-run, then MRTIE = MTIE.

6.4.2 Relationship between TDM Wander Specification and Base Station Clock Accuracy

The wander masks can be compared with the 16-ppb frequency stability requirement of cellular base stations. The wander limits determine the applicability of a TDM signal for base station synchronization. The limits are quite forgiving concerning short-term phase variation. For example, G.823 2M traffic mask allows 40-ppm frequency error at the observation interval of 0.1 s. As the requirement of base station clock is 16 ppb even for short observation windows, 2-Mbit/s traffic interface wander would be three orders of magnitude too high. Since the TDM signals used for base station synchronization are phase-locked to PRC, the long-term average error is as low as 0.01 ppb. Thus, the base station requirement can be satisfied by averaging the reference for a long enough time.

The most commonly used reference is 2.048 Mbit/s PDH traffic signal. The wander is specified in Section 5.2.1 of G.823. From Figure 6.9 one can deduce that by averaging this signal for 1125 seconds, the 16-ppb requirement can be achieved. By being slightly picky, one

can see that the G.823 mask, at the level of 18 μs , ends at 1000 s, allowing a long-term drift of 18 ppb. However, phase-locked clocks will not drift very long to the same direction – so 16 ppb can be safely reached by averaging about 1000 s. By using 1.544 Mbit/s signal as a reference one could reach 16-ppb accuracy by averaging 525 seconds. Finally, by using a reference meeting SEC mask, adequate frequency stability could be achieved by averaging a mere 125 s.

6.4.3 TDEV

TDEV (Time Deviation) is another metric used in telecom. It describes time stability of a clock as a function of averaging time. The metric has been derived from Modified Allan Deviation (MDEV), which in turn describes frequency stability of a clock as a function of averaging time. MDEV is used, for example, to compare primary frequency references against each other. TDEV incorporates sliding windows similar to MTIE, and TDEV curve is also drawn as a function of the window width. However, instead of seeking maximum within the window, TDEV measures differences between three consecutive sliding averaging windows. The estimator formula is

$$\begin{aligned} \text{TDEV}(n\tau_0) &\cong \sqrt{\frac{1}{6n^2(N-3n+1)} \sum_{j=1}^{N-3n+1} \left[\sum_{i=j}^{n+j-1} (x_{i+2n} - 2x_{i+n} + x_i) \right]^2}, \\ n &= 1, 2, \dots, \text{integer part } \left(\frac{N}{3} \right) \end{aligned} \quad (6.6)$$

where $n\tau_0 = \tau$ is the averaging window size, $x(i)$ is the time error, n is the number of samples in the averaging window interval, τ_0 is the sampling interval, and N is the total number of samples in the measurement. Figure 6.10 shows the sliding observation windows of length $\tau = n\tau_0$ that

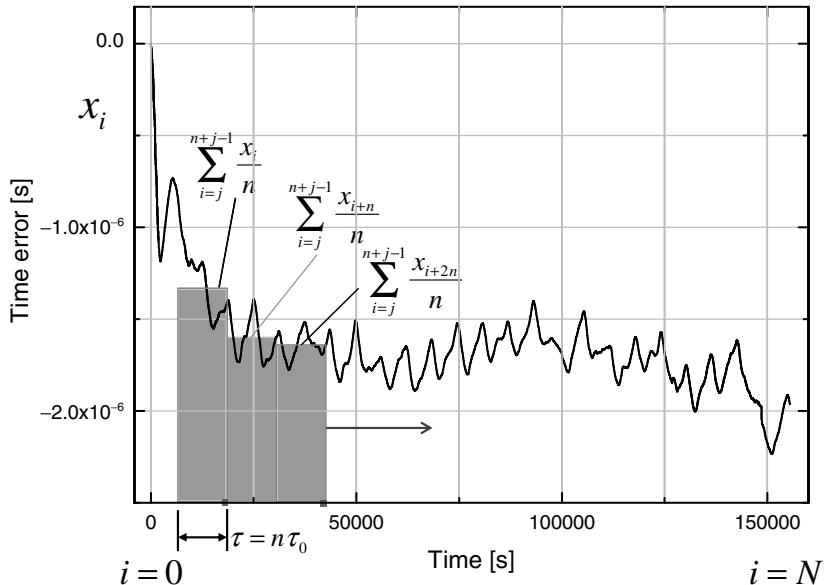


Figure 6.10 Description of the principle of TDEV function.

are slid over the N samples of data. There are two averaging processes. First, the contents of each time window indicated by the grey columns are averaged. This occurs inside the square brackets in the formula. Second, the outer sum averages the squared difference calculations as the averaging windows slide across the data.

$$\text{By noticing the relationship } x_{i+2n} - 2x_{i+n} + x_i = (x_{i+2n} - x_{i+n}) - (x_{i+n} - x_i) \quad (6.7)$$

where x_i , x_{i+n} , and x_{i+2n} represent the left, middle, and right window, respectively, one can see that the metric detects how the time difference between consecutive averaging windows vary. This corresponds to measuring how much the frequency error changes over the time $n\tau_0$. Thus, a clock, with a constant frequency error will yield $\text{TDEV} = 0$. On the other hand, a clock with high frequency/phase noise will yield a high TDEV. The relationship between TDEV and MDEV is simple:

$$\text{TDEV}(n\tau_0) = \frac{n\tau_0}{\sqrt{3}} \text{MDEV}(n\tau_0) \quad (6.8)$$

Figure 6.11 depicts TDEV of the clock time error shown in Figure 6.10. One can see how the phase noise level varies as a function of sliding window size. The curve shows that at time intervals below 10 s the average noise level is 0.2 ns. At larger time intervals the average noise level increases up to about 100 ns. When comparing this value to time error variation in Figure 6.10, one can see that some time error changes in Figure 6.10 at corresponding time scales are up to ten times larger than indicated by TDEV. This is due to the averaging property of TDEV.

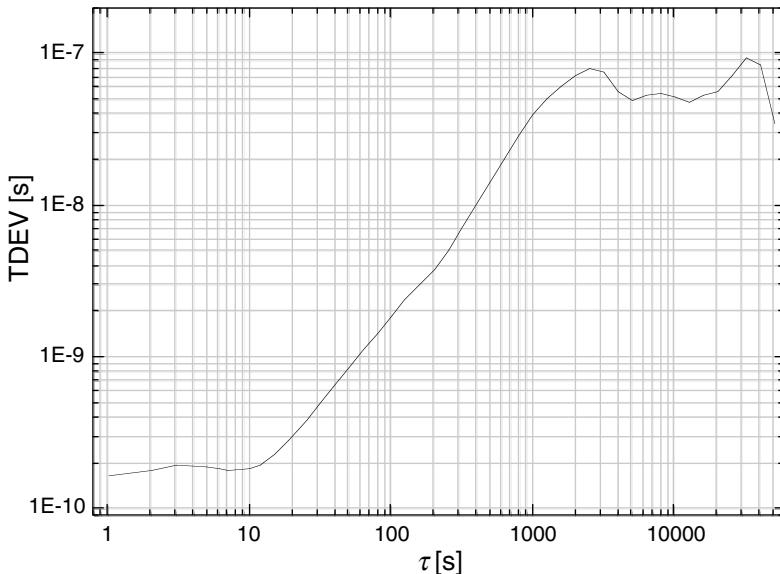


Figure 6.11 TDEV curve of a clock.

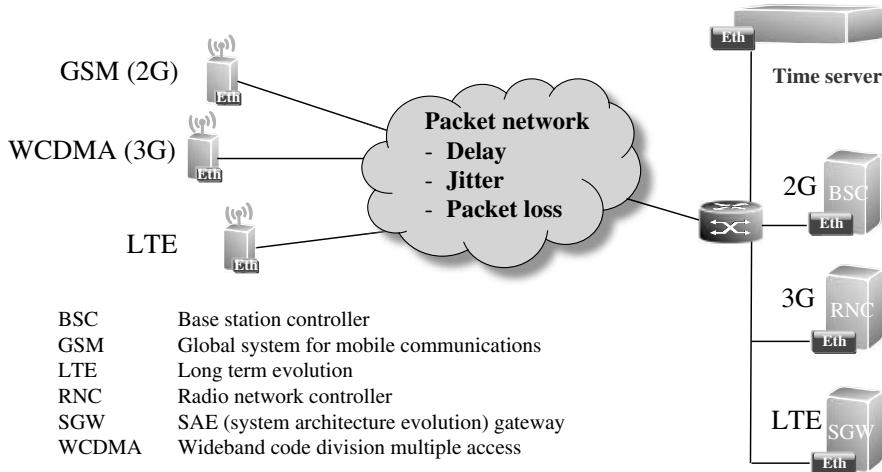


Figure 6.12 Packet timing architecture.

6.5 Packet Synchronization Fundamentals and Metrics

In the case of ACR, see Section 6.5.1, the interworking function mapping TDM frames into packets takes care that the pseudowire packets are sent out using a precise packet rate. In the case of NTP and PTP the mobile network operator runs time servers in central locations. A single master typically serves several hundred base stations providing an individual timing packet stream for each. The transport network can be treated as a cloud, which is unaware of the timing streams, see Figure 6.12. The clock recovery algorithm in the base stations is in a key role.

In TDM networks the phase drifts are in the microsecond range, which makes synchronization of the base station relatively straightforward. In packet timing, the delay variation of the packets is the ‘phase variation’ of packet networks. The delay variation in packet based transport networks is up to several milliseconds, two to three orders of magnitude higher. First, the principle of packet synchronization is explained, from where after the scheme to overcome the accuracy problem is discussed.

6.5.1 The Principles of Packet Timing for Frequency Synchronization

For transporting time the master clock sends timing packets to the slave clock. At least two packets are needed to convey frequency, see Figure 6.13. Either the period of the timing packets is known beforehand, as in adaptive clock recovery (ACR) or the timing packets are time stamped, as in, for example, PTP. If two timing packets are sent exactly 1000 s apart, the slave clock should advance exactly 1000 s between receiving the two messages. If it advances only 999.99999 s, then the slave clock has been 10 ppb too slow and the frequency should be increased by that amount.

Packets experience different amounts of delay, for example because of queuing, see Figure 6.14. If the first packet has been delayed by 2 ms and the second one 4 ms, the

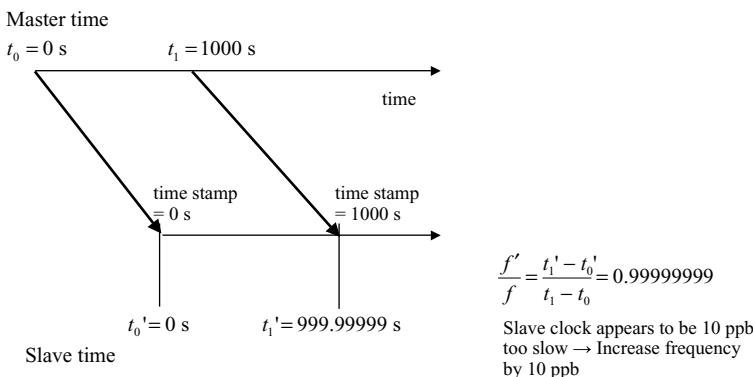


Figure 6.13 Principle of packet timing for conveying frequency.

delay difference is 2 ms. Consequently, the time interval between receiving the packets is 1000.002 s. If the slave clock is 10 ppb too slow as in the previous case, the slave clock has advanced by 1000.00199 s. The frequency ratio calculation yields a frequency error of +1990 ppb instead of the correct value -10 ppb. Clearly, as such, the accuracy of the frequency error determination would be far from enough and therefore some more tools are needed.

Let's examine the delay distribution of the packets to find out whether an improvement could be achieved by selecting a certain fraction of the timing packets. For determining delay distributions as a function of load, a network of five Ethernet switches was built, see 'Measured network' in Figure 6.15. The traffic streams were arranged so that they resemble those in a mobile transport network where, due to the tree structure of the network, traffic streams to base stations are branched off at the intermediate nodes. Since a five node chain is rather short compared to average chain length, a longer chain was approximated by summing up two copies of the 60000 data point measurement to produce delay values twice as large as on average. The order of the data points of the second copy was reversed first in order to remove

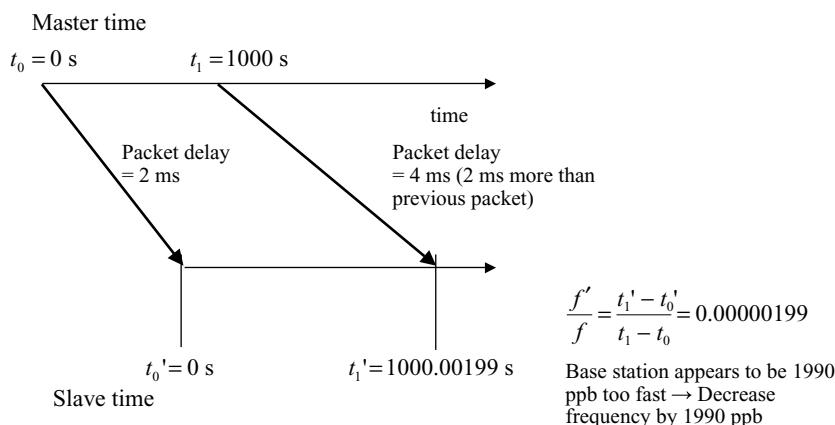


Figure 6.14 Effect of packet delay variation to packet timing.

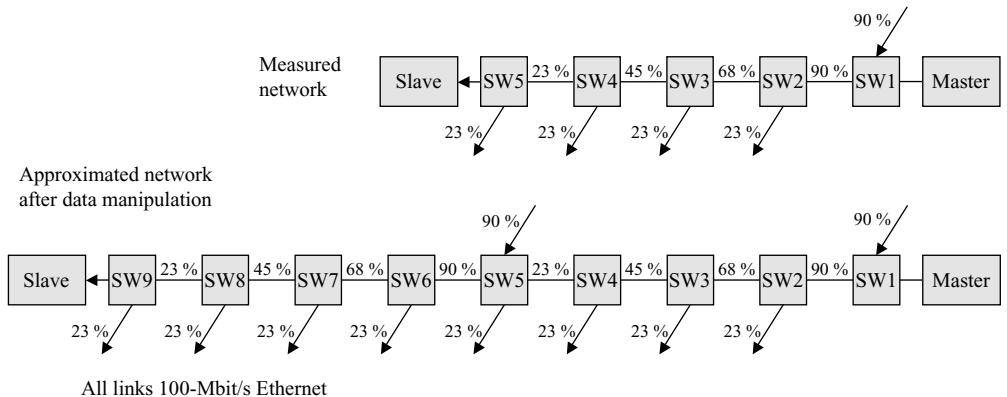


Figure 6.15 Measured and approximated network. The loads represent the ‘high load’ situation. In the ‘medium load’ situation the loads are correspondingly 50%, 38%, 25% and 13%.

correlation between the two. The approximation is not perfect but, nevertheless, probably leads to a sufficiently accurate outcome.

Figure 6.16 depicts delay distributions of the three load situations. If any packet were to be used while the load varied between the extremes, the delay uncertainty would be 2.8 ms (difference between the delay of the fastest packet of the lowest load and slowest packet of the highest load). As previously concluded, the resulting accuracy is not acceptable. The change of average delay going from low to high load is better, 0.7 ms, but still yields too inaccurate a frequency error approximations. The minimum delay varies only 0.005 ms. This variation

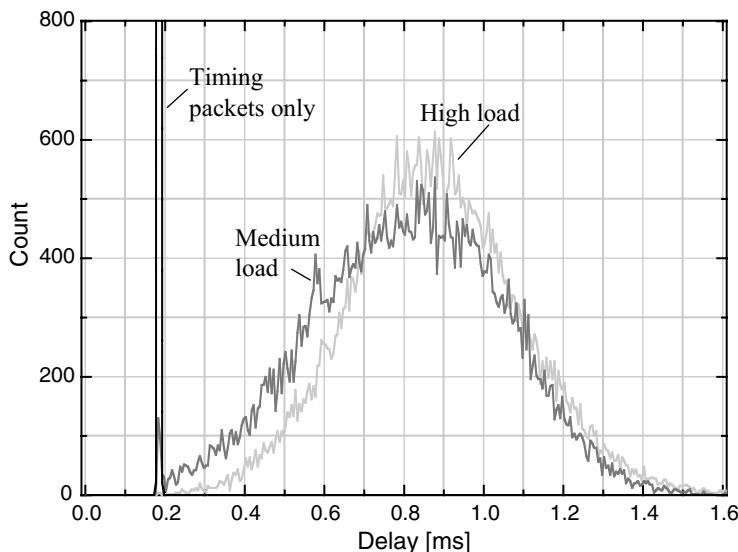


Figure 6.16 Minimum, average, and maximum delays: Low load – 0.180 ms, 0.185 ms, and 0.190 ms. Medium load - 0.179 ms, 0.810 ms, and 1.73 ms. High load – 0.184 ms, 0.888 ms, and 3.01 ms.

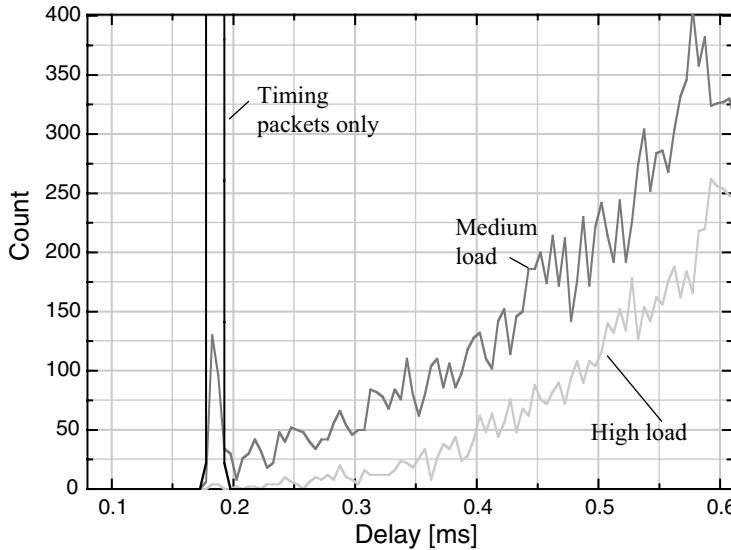


Figure 6.17 Small-delay tail of the measurement.

would be small enough. However, as the minimum delay packet in the low-load and high-load cases is the single fastest packet in each measurement, a more careful analysis is required.

Figure 6.17 shows that while all 60000 packets in the timing-packets-only situation are confined within a 0.01-ms range, in the medium load situation only about 200 packets remain. In the high-load situation, a mere 10 packets are left.

Figure 6.18 shows how the number of packets accumulates as a function of delay. The delay within 1%, 0.2%, and 0.03%-percentiles varies as a function of load 0.22 ms, 0.11 ms, and 0.045 ms, respectively. If the delays of the packets in the percentiles are averaged, then the percentiles differ from the minimum delay by 0.17 ms, 0.08 ms, and 0.02 ms, respectively. This shows that averaging decreases the error caused by delay variation. In the measured network, the minimum delay is very stable. In this case selecting a smaller and smaller fraction of packets improves performance. However, some transport technologies, like VDSL, have delay noise even without any load. In this case, instead of dropping the percentile to a very small one, it is better to use a slightly larger percentile to obtain more packets for averaging. Initially, a suitable compromise value could be 1%. For obtaining sufficient number of fast packets despite the small selection percentage, large packet rates between 16 pps 128 pps are typically used.

It is shown in Figure 6.18 that by using only the fastest packets and by averaging the information, more accurate estimations of frequency error can be obtained. The load-based delay variation of 0.17 ms after the packet selection and averaging would still yield 170-ppb error in the unlucky event that the increase occurs steadily over a long time so that the earlier presented 1000-s averaging of the clock algorithm is not long enough. Therefore, increasing the averaging time to about 10000 s would be needed to approach the 16-ppb level. Note that long chains of 100-Mbit/s links are mostly history in telecom. Most of the links have the capacity of 1 Gbit/s or 10 Gbit/s. Having said that, mobile networks in particular may have up

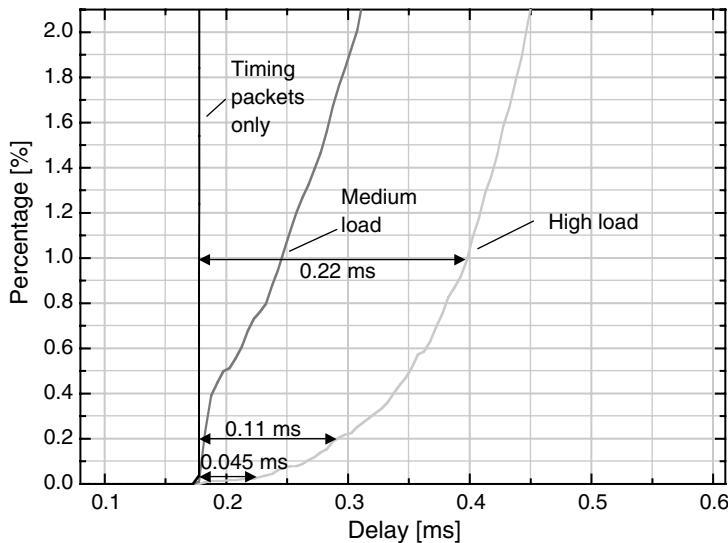


Figure 6.18 Accumulation of packets as a function of delay.

to about ten hops of microwave radio links with lesser capacity in a row. However, in such a long chain most of the hops would still be clearly higher than 100 Mbit/s in bandwidth.

The function model of a packet clock is shown in Figure 6.19. Compared with the function model of TDM based clock, see Figure 6.7, the packet based clock has two additional blocks, Local time scale and Packet selector. The oscillator output drives the pace of the local time. The packet selector uses the local time information for packet selection. The packets with the largest time stamp values compared with local time are the fastest and are selected. In addition, the comparator block has changed. Instead of the phase or frequency comparator there is a time scale comparator. As in the TDM based clock, the output of the comparator block may actually be a frequency error signal instead of a time error signal.

The averaging time needed in packet synchronization can be long. In this time temperature may change as much as 10° C causing drift in the oscillator frequency too quick for the algorithm to respond in time. Also the packet selection is affected. Therefore, the averaging time must be tailored on the basis of the possible oscillator responses to the environmental conditions.

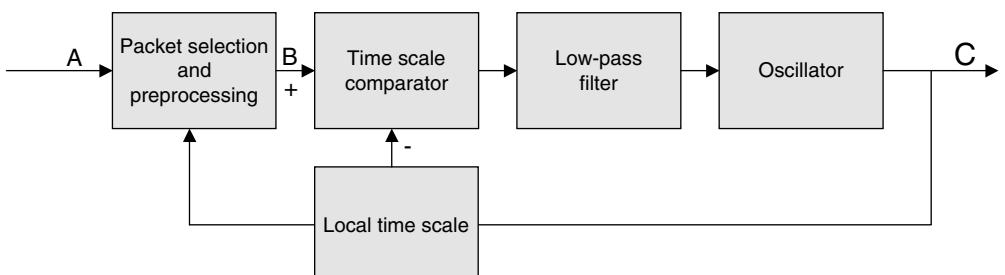


Figure 6.19 Function model of a packet based clock. At point A the timing packets enter the clock, at point B the packet selection and pre-processing have been carried out. Finally, C is the output of the clock.

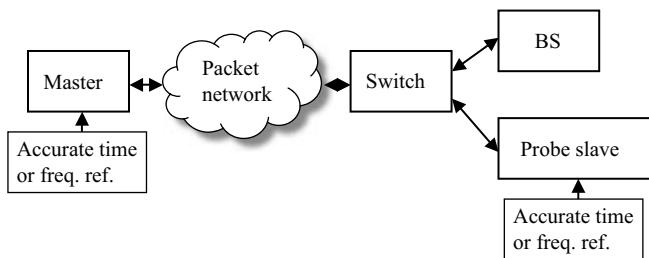


Figure 6.20 Packet delay measurement.

6.5.2 Packet Delay Metrics for Frequency Synchronization

As shown before, the ‘raw material’ for TDM metrics is the time error of the clock. For a packet clock, the time error is caused by the delay variation of the timing packets. Thus, the delays of the packets can be used directly as the raw data for corresponding packet timing metrics. The delays can be measured, for example using the setup in Figure 6.20. Due to the accurate time or frequency references at both ends all time stamps $t_1 \dots t_4$ of Figure 6.5 are accurate and known by the probe slave. If the reference is a frequency reference, then the absolute delays cannot be determined. However, this does not matter in the case of frequency synchronization.

As mentioned earlier, packet selection is crucial for packet based frequency synchronization. Therefore, metrics for packet based frequency synchronization always involve packet selection. There are two options, either integrate packet selection in the metric itself or pre-select packets and create a new data set on which the actual metric bases, see Figure 6.21.

The second selection model resembles more than that of a packet clock. Therefore it has usually been considered regarding PDV (packet delay variation) tolerance specifications. Due to limited space only this model is discussed below.

The network load is usually the main contributor for PDV. The load has a strong 24-h pattern. Therefore, typically delay measurements last for at least 24 hours. The packet time error sequence will consist of about one hundred thousand to a few million samples. For packet selection, the data is then divided into time windows of equal length. Then a certain percentage of fastest packets is selected from this group. The selected delays are then averaged to create a single new delay value from each window. This is the most commonly used preprocessing

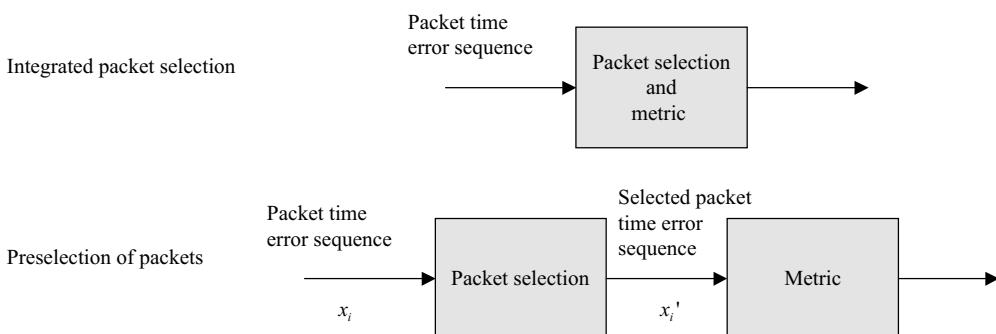


Figure 6.21 Metric models.

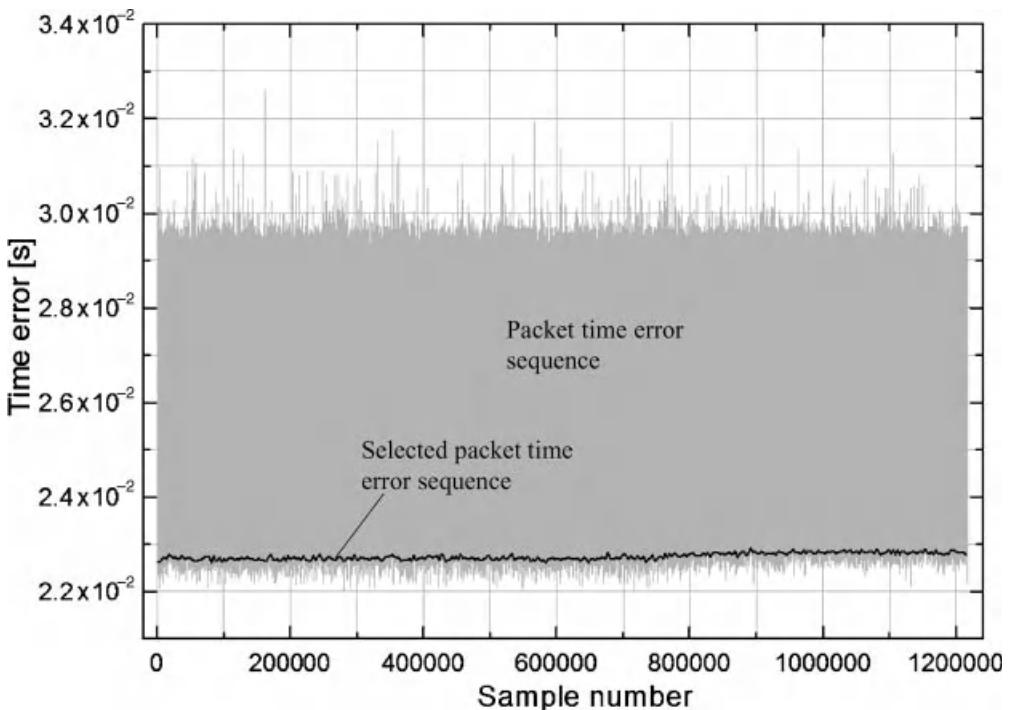


Figure 6.22 Time error sequence and selected packet time error sequence.

method. There are other schemes, as well, but these are not discussed. Figure 6.22 shows that selected packet time error sequence has significantly less delay noise than the original sequence.

6.5.2.1 TDEV

After the preprocessing, for example TDEV or MDEV can be calculated. The TDEV of the selected packet time error sequence will coincide quite well with TDEV calculated from a packet clock at τ values that are above the averaging capability of the clock algorithm, see Figure 6.23. Although TDEV calculated from preprocessed delay values estimates clock output TDEV very well, the metric has not been considered for tolerance specifications of packet clocks. This is because TDEV averages the perturbations over the whole measurement period and therefore it cannot be used to describe maximum allowed perturbations.

6.5.2.2 MATIE and MAFE

As the curves in Figure 6.23 almost combine, it suggests that the averaging inside the sliding windows (grey columns in Figure 6.10) resembles the averaging occurring inside a packet clock. Thus, one can imagine that the phase difference between two neighbouring averaging windows estimates the phase change of a clock using corresponding averaging in the tuning

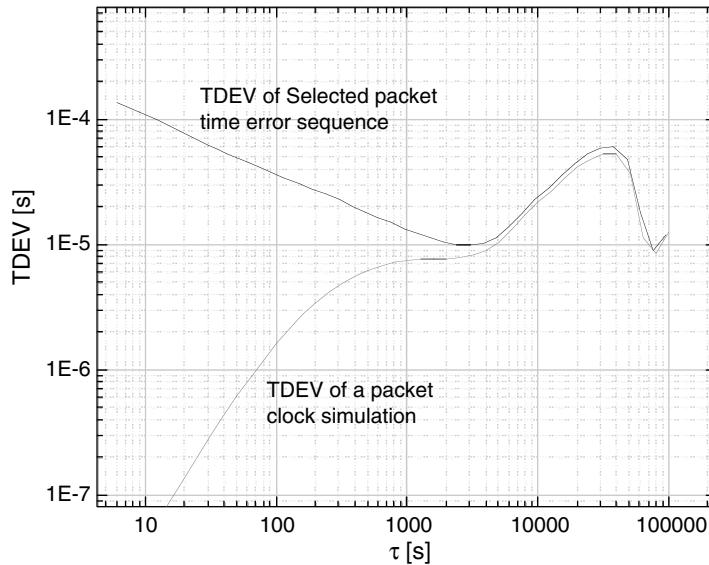


Figure 6.23 TDEV of packet time errors and of a packet clock.

algorithm, see Figure 6.24. By sliding the two windows and determining the maximum difference (instead of averaging as in the outer sum of TDEV), one obtains an estimate of the maximum phase change or maximum frequency error.

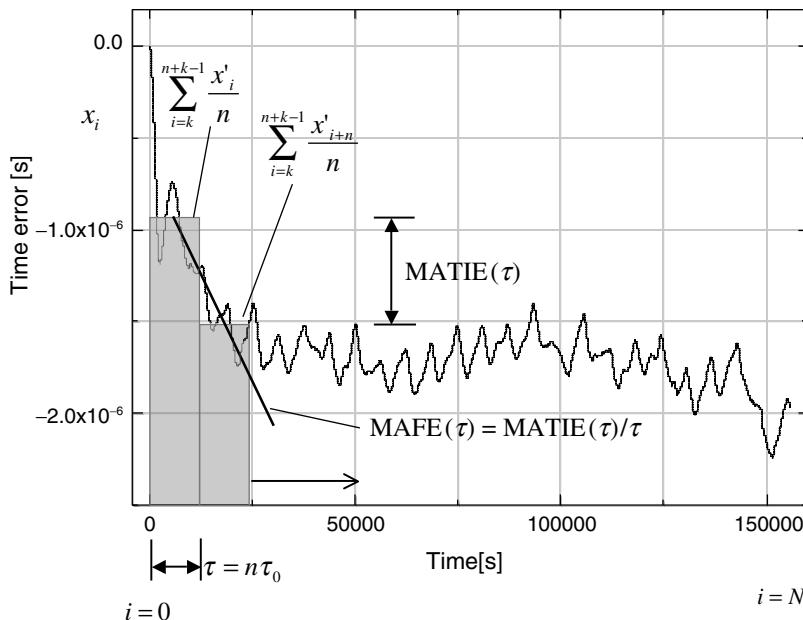


Figure 6.24 Principle of MATIE and MAFE calculations.

Equation 6.9 is used to calculate the maximum phase change, or maximum average time interval error (MATIE) from preprocessed delay sequence.

$$\text{MATIE}(n\tau_0) \cong \max_{1 \leq k \leq N-2n+1} \frac{1}{n} \left| \sum_{i=k}^{n+k-1} (x'_{i+n} - x'_i) \right|, \text{ for } n = 1, 2, \dots, \text{ integer part } (N/2) \quad (6.9)$$

When comparing to the MTIE and TDEV formulas one notices that MATIE is a mixture of both. The maximum average frequency error is calculated by Equation 6.10.

$$\text{MAFE}(n\tau_0) \cong \frac{\max_{1 \leq k \leq N-2n+1} \frac{1}{n} \left| \sum_{i=k}^{n+k-1} (x'_{i+n} - x'_i) \right|}{n\tau_0}, \text{ for } n = 1, 2, \dots, \text{ integer part } (N/2) \quad (6.10)$$

Figure 6.25 shows MAFE calculations from similar packet delay sequences as in Figure 6.22. As in the case of TDEV, the curves calculated from selected packet time error sequence and packet clock coincide with each other at τ values above the filtering capability of the clock. Note that the frequency stability of the clock is not adequate for a base station internal clock. It is just an example for showing the correlation. Clock simulations at different filtering bandwidths were tested on the selected packet delay sequence. It was observed that the τ_{estimate} derived from the maximum frequency errors of each clock has a linear, but not 1:1 relationship with the time constant of the PLL filter. Thus, the maximum frequency error of a clock could, in principle, be estimated from the time constant of the clock and the MAFE curve of the selected packet time error sequence.

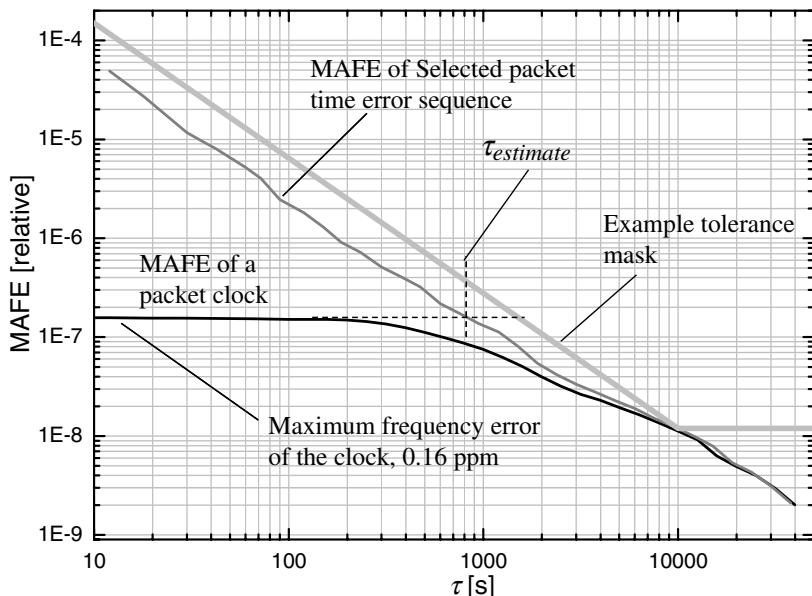


Figure 6.25 MAFE of selected packet time error sequence and a packet clock.

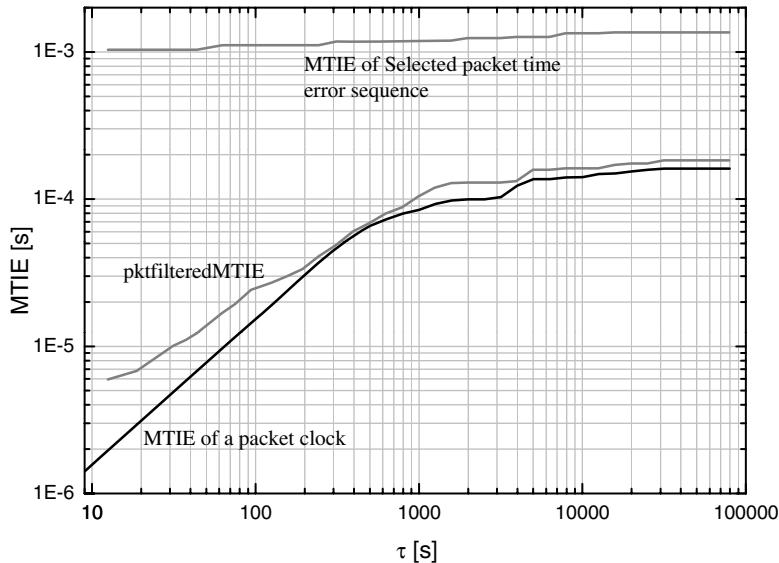


Figure 6.26 MTIE curves.

The light grey mask is an example of a possible tolerance mask of a clock that is able to achieve 16-ppb accuracy at this delay noise level. Approximately 10000-s averaging capability is needed. The corner point in the mask is therefore at 10000 s. The level is at 12 ppb for giving margin for delay jumps and oscillator drift. The slope can specify the limit to noise at the time scales affecting the algorithm.

6.5.2.3 pktfilteredMTIE

In TDM networks the most commonly used metric is MTIE. Therefore, MTIE would be the metric of choice if it could be used for tolerance specification. Figure 6.26 depicts various MTIE curves. Clearly it can be seen that MTIE applied directly into the selected packet time error sequence does not correlate with the corresponding calculation applied on the clock output.

However, if a sliding averaging window is first applied to the selected packet time error sequence, the situation changes dramatically. For accomplishing this, one more block, ‘Bandwidth filtering’, is added to the metric model, see Figure 6.27. The curve pktfilteredMTIE in Figure 6.26 depicts an MTIE calculation applied to data that has been processed with a 1500-s averaging filter. This matches quite well with the MTIE of the clock output.

The formula of bandwidth filtering

$$y_i = \frac{1}{n} \sum_{j=i}^{n+i-1} (x'_j), \quad i = 1, 2, \dots, N-n+1, \quad (6.11)$$

is more or less a direct copy of the sliding averaging windows in TDEV, MATIE and MAFE formulas. The variables \$x'_j\$ and \$y_i\$ are the Selected and Filtered packet time error sequences,

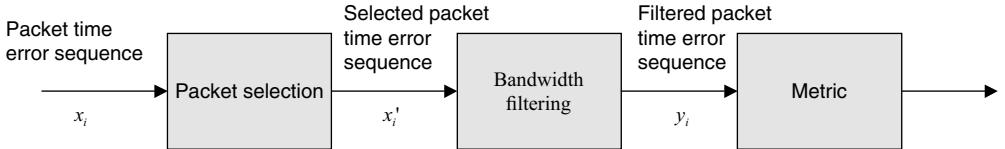


Figure 6.27 Metric model with bandwidth filtering.

respectively. The other variables are familiar from the TDEV formula, Equation 6.6. The notation used in Eq. 6.11 differs slightly from the one used in G.8260, but the calculated values remain the same. To conclude, pktfilteredMTIE is very useful for testing against MTIE masks that extend to long observation intervals, such as G.824 1.5 Mbit/s network mask and G.823 SEC mask, see Figure 6.9.

6.5.2.4 Floor Delay Packet Population

If all selected packets fit within a certain distance of the fastest packet in the measurement, one can define the maximal averaging time that is needed to remain within e.g. 16 ppb. For example in Figure 6.28, a slope from 0 to 150 μ s delay may occur in 9375 s or slower for remaining below 16-ppb error, if it is not filtered. If the swing occurs in a smaller time, the delay needs to be filtered to remain below 16-ppb slope. The most difficult patterns for filters are long slopes that are slightly shorter than 9375 but having the maximal, 150- μ s amplitude. By using about 9500-s averaging window, it can be ensured that no imaginable shape of floor delay pattern will ever cause a larger frequency deviation than 16 ppb.

A network limit can be specified based on defining such a delay window. The method was adopted to G.8261.1 concerning network limits for HRM-1, see Ch. 6.3.7.3. The network limit requires that from every 200-s selection window, at least 1% of packets have to be within the 150- μ s range from the fastest packet of the measurement for fulfilling the PDV limit.

The method is simple to comprehend and therefore straightforward to use as a metric. However, it is a conservative one since the worst case delay floor pattern will never occur in practice. Packet delay always has short-term variation that is easier to filter out. As the floor delay population method does not discriminate between the short-term and longer-term variation, delay scenarios that might be acceptable for a clock with such a long averaging are not necessarily acceptable by the metric. The large 150- μ s value despite the easy reference model received some criticism and the value might be revised.

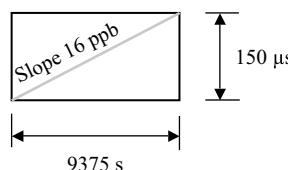


Figure 6.28 Floor delay packet population.

6.5.3 Two-way Messaging

The packet delay metrics have been discussed for simplicity's sake as if one-way-only timing traffic is used. The same formulas apply whether forward or reverse packets are used. In cases where timing packets in both directions are used, then both directions need to be considered. First, the selected packet preprocessing is carried out independently in both directions. Then, the difference between each pair of preprocessed forward and reverse delays is carried out and divided by two.

$$x'_i = \frac{x'_{forward_i} - x'_{reverse_i}}{2} \quad (6.12)$$

In the formula the weight of the two directions is the same. The formula can be further developed to accommodate asymmetric weighing. Even dynamic weighing has been proposed.

6.5.4 Delay Jumps

If the route between endpoints changes, then the delay also changes. Typically, delay jumps are from a few tens of microseconds to several milliseconds. Good packet clocks can accept from one to a few hundred microseconds of minimum delay variation depending on the spectral distribution of the delay pattern. Thus, a jump of several milliseconds certainly needs to be treated in a special way. Even a jump of tens of microseconds is unwelcome because it contains low-frequency components that are difficult to filter and correspondingly would eat a significant portion of the available delay variation budget. Thus, packet clocks need to have non-linear features for diminishing the effect of both large and relatively small delay jumps. Luckily, the number of delay jumps in a day is low. Thus, the algorithm can rely on low repetition rate of the events.

Delay jumps are especially cumbersome for clocks that are expected to satisfy MTIE limits extending to observation intervals of tens of thousands of seconds like G.823 SEC and G.824 1.5-Mbit/s network limits. SEC limit is just a few microseconds over a whole day, see Figure 6.9. One possibility is to use a very stable oscillator for detecting delay jumps of just a couple of microseconds. Another possibility opens if it can be assured that delay jumps are symmetric, i.e. both directions of the packet timing flow experience identical jumps. Two-way messaging is needed and the algorithm has to weigh the two directions equally. In this case no scheme for delay jump detection is needed.

6.5.5 Testing Packet Timing Slaves

Figure 6.20 showed a test setup for measuring packet delays in networks. Figure 6.29 describes a setup for testing a packet clock. A PDV file created from a delay measurement can be uploaded to an impairment emulator. Being able to create known and repeatable conditions is essential in packet clock testing. Therefore, impairment emulators are crucial for determining the quality of different clocks.

The emulator vendors have created delay files for their customers especially for the ITU test cases.

ITU test cases

G.8261, Appendix VI, describes various test cases utilizing a reference network consisting of ten Ethernet switches with Gbit/s interfaces. The network is loaded in various ways to create packet delay variation on the way of timing packets. The switches are loaded up to 80% load,

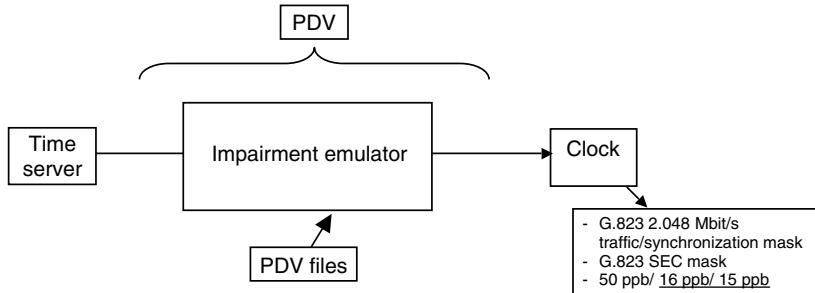


Figure 6.29 Packet clock test setup.

which is relatively high considering that all switches encounter the load at the same time. In real life the average load is lower, but, on the other hand, the chains of nodes can be longer.

The most challenging two-way synchronization test cases (TC) are TC13 and TC14. TC13 has 1 h periods of 80% and 20% load in forward direction and corresponding but half-hour shifted periods of 50% and 10% load in the reverse direction. Figure 6.30 shows the MAFE curves (using packet rate of 16-pps, 1% preselection from 60-s windows) calculated separately from forward, reverse, and two-way direction. The delay files of vendor B are more challenging in the fwd and 2-way case and the reverse delay file of vendor A is more challenging than the corresponding one of vendor B.

In both cases a two-way clock remains within 16-ppb frequency error by averaging for less than 500 s. This is a rather short time, actually shorter than is needed to reliably extract 16-ppb accuracy from 2-Mbit/s PDH signal. Therefore, also more difficult test cases are needed.

6.6 Rules of Thumb for Packet Timing Network Implementation

At the time of writing, the delay variation effects in complex networks have not been studied thoroughly enough for estimating delays at the fine-grain accuracies required for timing

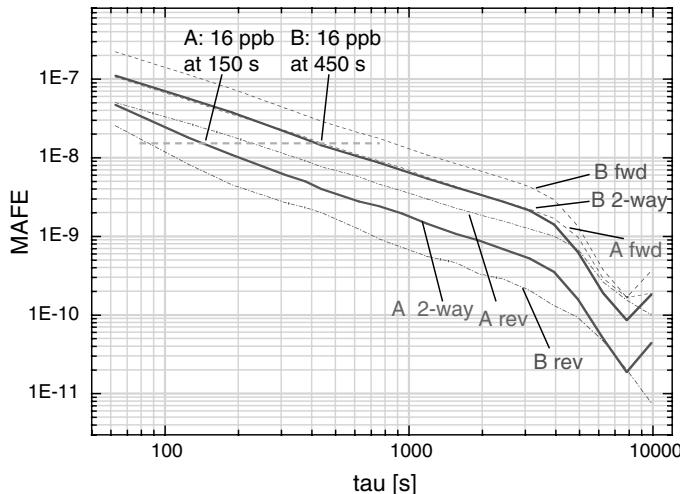


Figure 6.30 MAFE curves of TC 13. A: Vendor A, B: Vendor B.

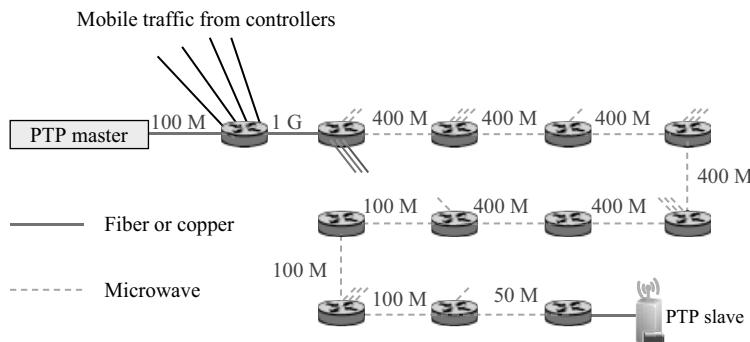


Figure 6.31 Example of a path in a commercial MWR backhaul deployment.

purposes. A network can be assessed by carrying out delay measurements over multiple connections at different load conditions. Unfortunately, just after building the network the loading is still much less than it will be after some time. However, for example MAPE and pktfilteredMTIE produce quantitative results where the margin between the limit and delay variation can be observed. If at a later measurement the margin has become clearly smaller, then some links might need upgrading.

Experience so far indicates that if the packet timing slaves are good enough, there is no need to decentralize time servers to the edges of the transport network. The path in Figure 6.31 is part of a commercial deployment. There are two copper/fiber Ethernet links followed by ten packet microwave radio (MWR) links.

Although the implementation above has worked well, it cannot be taken for granted that all such implementations will work. An example set of rules for 16-ppb clock accuracy target and G.823 2-Mbit/s traffic mask is given below. The values match with a packet clock that has a 10000-s MAPE corner point shown in Figure 6.25.

- Maximum one way delay should be < 100 ms.
- Jitter < 5 ms.
- Packet loss < 2%.
- Clock packet stream should have the highest priority or at least the same priority as the real-time traffic and receive expedited forwarding QoS.
- High-priority traffic share of BW should be ~60 % or less.
- Maximum number of hops: 20.
- Maximum number microwave hops: 10. In this case the total number of hops should be less than 15.
- The average load of the links along the path should not be persistently above ~50% if the path is long.
- The number of delay jumps should be limited to a few per day

For targeting clock stability requirements including long observation intervals, such as G.824 1.5-Mbit/s mask or G.823 SEC mask, clearly stricter rules apply. The ITU test case setup of ten hop chain with 1-Gbit/s links should be acceptable.

The long MWR chain described above utilizes FDD (Frequency Division Duplex) technology where both directions of the links are active continuously enabling good packet

timing performance. However, there are also some TDD (Time Division Duplex) radios on the market. In this case each direction is waiting for half of the time causing potentially problems in packet timing. Especially, more than one TDD link in a row could be risky for packet timing.

SHDSL (Single-pair High-speed Digital Subscriber Line) and VDSL (Very-high-bit-rate Digital Subscriber Line) have survived as mobile backhaul links at the same time as PDH is becoming obsolete. Neither DSL technology has notable issues in traversing packet timing except for the lower bit rate compared with point-to-point optical interfaces.

Passive Optical Networks (PONs) operate continuously in downstream but employ TDM (Time Division Multiplexing) technique upstream. Therefore, the delay floor in upstream direction is quite noisy and downstream direction is usually smoother.

6.7 Time Synchronization

As can be seen in Table 6.2, there are several cellular systems requiring time synchronization. The most widely deployed time synchronized system is CDMA used by several hundred million subscribers around the world. It is expected that the TDD version of LTE will become popular. If this happens, the number of time synchronized base stations will increase significantly in the coming years. At the moment time synchronization is provided mainly using GPS. It is expected that in the future PTP will take over due to the advantages obtained by network timing.

6.7.1 GNSS Systems

Time synchronized cellular systems have relied until these days solely on GPS (Global Positioning System) that is operated by the US government. However, another GNSS (global navigation satellite system), the Russian GLONASS (Globalnaya Navigatsionnaya Sputnikovaya Sistema or Global Navigation Satellite System) became ready for full global operation in October 2011. European Galileo is planned for 2014-19. In China the regional Beidou navigation system is planned to be expanded to a global system by 2020. In India a positioning system is also under development. The GNSS receiver vendors are currently adding at least the capability to receive GLONASS signal.

The accurate positioning is based on accurate time synchronization. The time can be retrieved at a better accuracy than ± 100 ns, which clearly leaves behind the requirements of cellular systems. For being able to autonomously obtain accurate frequency, a receiver needs to receive the signal from at least four GPS satellites. Therefore, the GPS antenna requires being able to see the majority of the sky, which is somewhat costly to arrange especially for base stations operating inside buildings or in street canyons. Vulnerability against jamming, and the small risk of the commercial signal being switched off due to a political situation have also been mentioned. Holdover periods up to 24 hours have been required by some operators increasing the cost of the local oscillator. The cost of the antenna installation is expected to drop since ever more sensitive receivers are entering the market. Another way to improve the situation is A-GPS (Assisted GPS). With the use of orbit information, the receiver is able to lock into weak signals more easily. After learning the location of the receiver, accurate time could be maintained even when receiving signal from a single satellite.

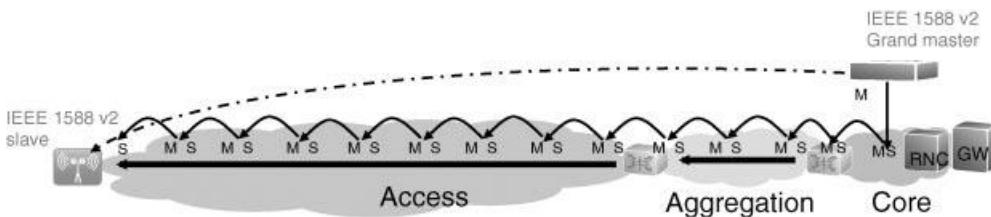


Figure 6.32 PTP synchronization chain.

6.7.2 PTP for Time Synchronization

As said earlier, PTP was originally designed for accurate time synchronization. The main asset to achieve the microsecond-level accuracy required by cellular systems is the boundary clock specification. If all intermediate nodes are equipped with boundary clocks, see Figure 6.32, the timing packets will not experience queuing at any node.

The PTP standard specifies several alternative ways to create time synchronization chains in telecom networks. For example, the protocol mapping could be PTP over Ethernet or PTP over UDP over IPv4 (or IPv6), the delay requests could be carried out using either end-to-end or peer delay mechanism. Therefore ITU is working on a PTP time synchronization profile. The following recommendations have been considered. Only G.8271 has been finished by February 2012 and the collection is subject to change.

- G.8271 Time and phase synchronization aspects in packet networks
- G.8271.1 Network requirements for time/phase
- G.8272 PRTC (Primary reference time clock)
- G.8273 Packet Time/phase clocks: Framework and clock basics
- G.8273.1 Grandmaster
- G.8273.2 Boundary clock
- G.8275 Packet architecture for time/phase
- G.8275.1 PTP profile for time/phase

6.8 Conclusions

Switching from TDM to packet networks has required significant research, development, and standardization activity within the area of synchronization. A second burst of activity has been caused by the definition of new time synchronized cellular technologies that are expected to be widely adopted.

In the case of frequency synchronization there are two choices. In the case of no on-path support there is a choice between Synchronous Ethernet and packet based technologies such as PTP. Synchronous Ethernet can provide high-quality synchronization regardless of network load. One of the drawbacks is the need for full on-path support, which might not be available for a long time from some leased line operators. Another problem is associated with 1-Gbit/s and 10-Gbit/s copper Ethernet links that cannot change the direction of synchronization for protection switching purposes in a reasonable time. PTP can, in principle, be used from day

one in all packet networks, because it does not need on-path support. There are some risks though because large load variations combined with a high number of hops may lead to performance limits being exceeded. In terms of deployments over the first two years PTP has been very reliable though.

Time synchronization has been for a long time a requirement in CDMA networks. LTE and the increasing number of small cells will explode the number of time synchronized base stations in several years, increasing the need for network based synchronization. The only practical network based alternative in this case is PTP. However, the use of PTP will be quite different to the frequency synchronization case and therefore frequency and time profiles should not be mixed. The work concerning PTP time profile for telecom has still not advanced very much by the writing of this book.

References

- [1] 3GPP TR 45.050, ‘Background for Radio Frequency (RF) requirements’.
- [2] 3GPP TR 25.951, ‘FDD Base Station (BS) classification’.
- [3] 3GPP TS 45.010, Technical Specification Group GSM/EDGE Radio Access Network; Radio subsystem synchronization
- [4] 3GPP TS 25.104, ‘Technical Specification Group Radio Access Network; Base Station (BS) radio transmission and reception (FDD)’.
- [5] 3GPP2 C.S0002-D, ‘Physical Layer Standard for cdma2000 Spread Spectrum Systems’.
- [6] IEEE 802.16-2009, ‘Air Interface for Broadband Wireless Access Systems’.
- [7] ETSI TR 101 190, ‘Implementation guidelines for DVB terrestrial services; Transmission aspects’.
- [8] 3GPP TS 36.133 ‘Evolved Universal Terrestrial Radio Access (E-UTRA); Requirements for support of radio resource management’.
- [9] Federal Communications Commission FCC 07-166, ‘Report and order’, 11-2007.
- [10] RFC 4553, ‘Structure-Agnostic Time Division Multiplexing (TDM) over Packet (SAToP)’, IETF, June 2006.
- [11] RFC 5086, ‘Structure-Aware Time Division Multiplexed (TDM) Circuit Emulation Service over Packet Switched Network (CESoPSN)’, IETF, December 2007.
- [12] RFC 5905, ‘Network Time Protocol Version 4: Protocol and Algorithms Specification’, IETF, June 2010.
- [13] IEEE 1588-2008, ‘IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems’, IEEE March 2008.
- [14] Jean-Loup Ferrant, Geoffrey M. Garner, Michael Mayer, Juergen Rahn, Silvana Rodrigues, and Stefano Ruffini, ‘OTN Timing Aspects’, IEEE Communications Magazine, September 2010.

7

Resilience

Esa Metsälä

7.1 Introduction

Network resilience is e.g. in [1] used as ‘the maintenance of both the connectivity and the quality of service (QoS) in terms of packet loss and delay during network failures.’ This introduces the key aspects: a) maintaining service when failures occur, and, in addition b) with the defined QoS.

One aspect to start with is to look into each of the backhaul protocol layers and see what they provide in the case of link and node failures. In this chapter the focus is on the protocol layers above the physical layer: Native Ethernet (Section 7.2), Carrier grade Ethernet (Section 7.3), IP (Section 7.4) and MPLS (Section 7.5).

Resilience in the access tier is discussed in Section 7.6. Resilience of the radio network - core network interface is a topic for Section 7.7.

7.1.1 Restoration and Protection

Network recovery can be divided into restoration and protection [2], [3]. Restoration relies on activation of a new path, when a failure occurs, while protection switching uses a path that is pre-configured. Both mechanisms increase resilience of the network and both alternatives are deployed with packet networks.

Figure 7.1 shows the two approaches. In the left-hand side, protection switching occurs on a preconfigured back-up path when the failure is detected. In the right-hand side, a failure causes a routing protocol to reconverge and consequently find a new active path (dashed arrow). The new path is selected with the help of a routing protocol (best path selection).

Sonet/SDH supports extensive and well-known protection switching methods, typically with recovery in less or in the order of 50 ms. This has set the performance standard for transport networks, even though 50 ms is not a hard limit for many (if not most) of the

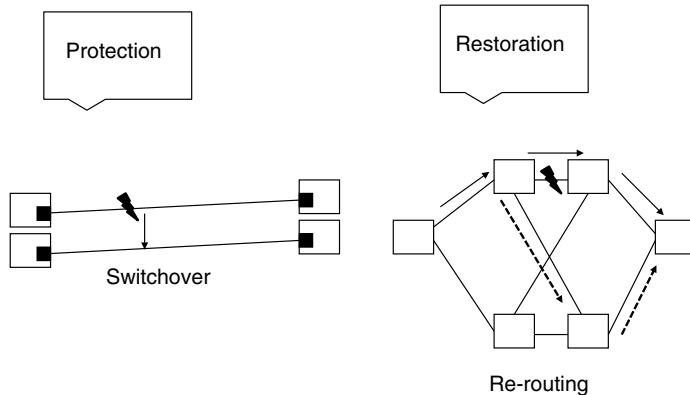


Figure 7.1 Protection and restoration examples.

applications. Also in mobile backhaul that is the case: calls and contexts are lost only after hundreds of milliseconds or even seconds, depending on the mobile network implementation and parameterization. Still a fast recovery, such as 50 ms, is a benefit as it increases availability.

Sonet/SDH type of fast protection switching is also supported at the Ethernet layer, with the use of ITU-T G.8031 and G.8032. Spanning tree is feasible as a local area (LAN) feature at the Ethernet layer but not as a wide area network recovery solution.

Routing protocols enable restoration by exchanging information about reachability to destination networks. When links or nodes fail, a new best path is calculated/selected, and traffic flow is restored via this new route. Typically, re-routing takes more time than protection switching. IGPs (Interior Gateway Protocol) support recovery times of subsecond to several seconds, depending on a number of factors.

MPLS TE Fast Reroute (FRR) and similarly IP FRR can bring the recovery time to values comparable to that of Sonet/SDH. FRR relies on a pre-calculated back-up path. MPLS-TP similarly supports 50 ms recovery, again depending on detection time and other topics.

Many of the concepts familiar from Sonet/SDH are reused in networking technologies that target a ‘transport network behavior’: extensive OAM, fast recovery, and a connection oriented (potentially deterministic) nature. An example of this is MPLS-TP.

7.1.2 Recovery

In RFC3469 [3], Sharma and Hellstrand present recovery cycle times for MPLS (Figure 7.2). The model suits a discussion of recovery in general, also a discussion of recovery in the mobile backhaul transport layers.

After time T1, network impairment leads to the fault being detected. Time T2 elapses before the failure is propagated further. Propagation to another network device takes T3, after which recovery starts, taking time T4. After T5, traffic is completely recovered.

It is clear that different recovery mechanisms take a different amount of time, depending on the type of failure, detection, network topology, and so on. Recovery as an operation is protocol and technology – dependent.

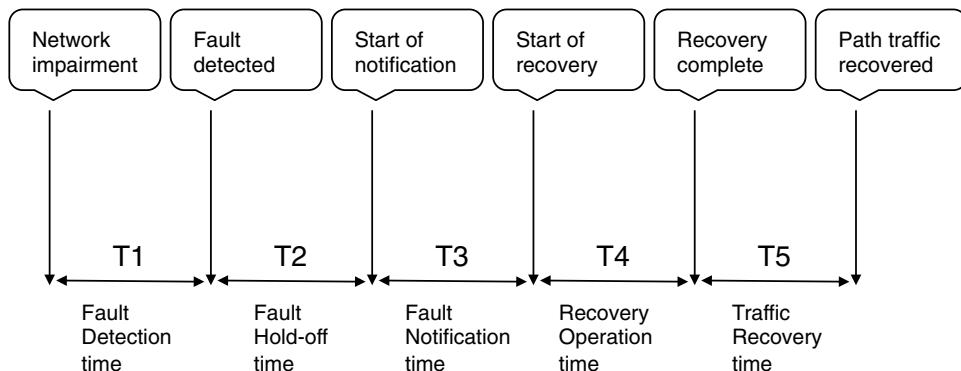


Figure 7.2 Recovery cycle times ([3], Sharma, Hellstrand).

7.1.3 Availability

Availability is commonly expressed in terms of nines: e.g. 99.99% ('4 nines') availability means the system is unavailable 0.01% of the time. Calculated over a year, and assuming 365 days* 24 hrs, e.g. 99.99% (4 nines) availability turns into a value of 52.56 minutes of service unavailability during a year. Other values are shown in Figure 7.3.

High availability is defined as 5 nines or more. 5 nines translates to a calculated probabilistic service downtime of 5.26 minutes per year.

Availability is defined with two variables; mean time between failures (MTBF) and a mean time to repair (MTTR) [4], [5].

$$\text{Availability} = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}$$

In order to make a meaningful calculation, MTBF and MTTR should include all relevant factors; not only HW faults, but also SW, and any configuration errors etc. Unavailability is caused both by planned outages, and unplanned outages.

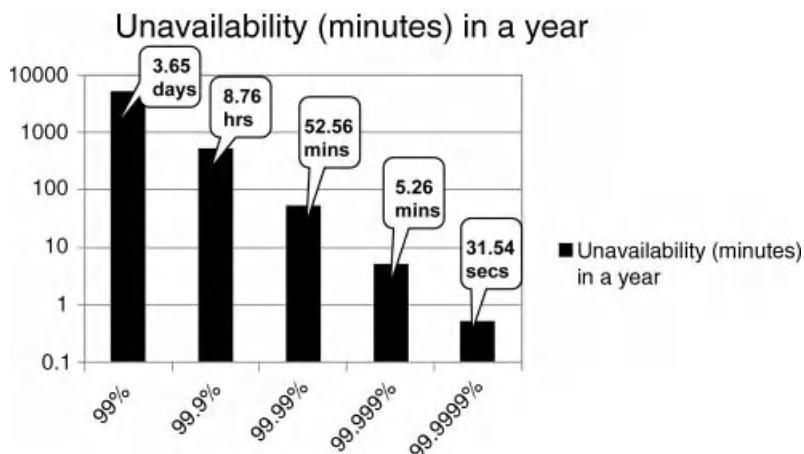


Figure 7.3 Unavailability during a year (note logarithmic y axis scale).

Table 7.1 Availability example of some transport network nodes and links. Based on [2] and [1] + information of microwave link availability.

Component	MTBF (hours)	MTTR (hrs)	Component availability	Availability as ‘nines’
IP interface card	10^5	2	0,999980	‘4 nines +’
IP router	10^6	2	0,999998	‘5 nines +’
Sonet/SDH ADM	10^6	4	0,999996	‘5 nines +’
WDM OXC/OADM	10^6	6	0,999994	‘5 nines +’
Fibre (1 km)	$7.0 * 10^6$	24	0,999997	‘5 nines +’
Fibre (10 km)	$0.70 * 10^6$	24	0,999966	‘4 nines +’
Fibre (300 km)	$0.23 * 10^5$	24	0,998974	‘3 nines’
Microwave link (availability target) ^a			0,9999 ^a	‘4 nines’

^a Microwave link (consisting of two radios) is designed so that the availability is reached. To meet the target, antenna sizes, hop lengths, modulation, and other factors are considered.

7.1.4 MTBF and MTTR

Availability was calculated with the help of two variables, MTBF and MTTR. What would the availability become, assuming tentative values for the two?

In [2] and [1] some values are given as examples. For fibre optic cable, a cable cut (CC) metric definition is first needed. This is the amount of cable that experiences one cable cut in a year. [5]

$$\text{MTBF} = (\text{1 year}/\text{length of the cable})^* \text{ CC}.$$

For the CC, values from 450 km to 800 km are considered in [2] and [1]. Consider a CC of 800 km as an example. With this value and with a fibre cable of 10 km (metro area), calculation results in an MTBF of 0.7×10^6 hrs. Additional assumptions for node failures in terms of MTBF and repair times (MTTR) are listed in Table 7.1.

Table 7.1 is an example. Clearly, actual values of the network nodes are vendor specific, the value of CC is network dependent, and MTTR depends on the operations. Even if real values differ in detail, a conclusion is that network links introduce a significant contribution to the unavailability. Fibre links of some tens of kilometres are more susceptible to failures than network nodes are. Additionally a longer repair time (MTTR) is assumed for the fibre cabling.

In practice, arranging redundant links is often costly and a more time-consuming task than designing a redundant router configuration.

The column component availability should be interpreted as a contribution of the named single component to the availability of a system. It does not take into account any other components of a system. In isolation it serves to compare the availability between components.

7.1.5 Increasing Availability

Given the estimate as in Table 7.1, what can we do to increase availability?

Reading from the availabilities, network links should be considered first. However with mobile network, nodes that serve a large number of other nodes, are also critical, even though

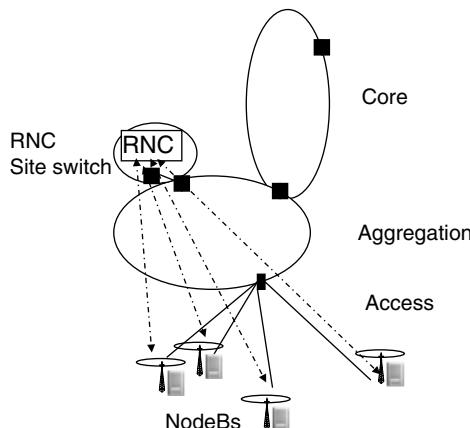


Figure 7.4 3G RAN topology.

the component level availability would be high. So one has to take dependency into account and focus on those nodes and links that are the most critical. Consider the topology of 3G RAN as illustrated in Figure 7.4.

Logical topology is shown as dashed lines: each NodeB needs connectivity to the RNC. Clearly, RNC is a single point of failure, as well as are RNC site devices, and aggregation nodes that connect the RNC to the mobile backhaul. Failure on any of these causes downtime for a large amount of NodeBs.

Similarly, aggregation nodes, routers and switches, that collect traffic from many sites (e.g. 20–50 base stations), are single-point-of-failures, unless duplicated. Even with high availability of the node as a component, a failure case causes downtime for a number of BTSs. The impact of a failure in the aggregation domain is larger than a failure in an individual BTS access link, be it then a failure of a link or a node – unless redundant components are deployed.

In LTE, the radio network consists of only a single node, eNodeB. Removal of a central radio network controller element (a BSC or RNC) leads to avoidance of one single-point-of-failure. This is one of the benefits of the flat architecture of LTE.

The physical transport topology with LTE, especially in the aggregation and core tier of the backhaul, is typically similar to the one in Figure 7.4 for 3G RAN. Traffic from multiple eNodeBs is aggregated before the core nodes (SGW and MME). The aggregation nodes are as critical as the 2G and 3G controllers are. In LTE, security gateways often exist, serving a large amount of eNodeBs.

Arranging for a redundant path is a powerful tool in increasing availability in the case of failing links. Assume that we can make the path via the 10-km fibre link in Table 7.1 redundant, with a 50 ms recovery time. This is achieved e.g. by Sonet/SDH protection, or MPLS FRR. Naturally the link still needs to be repaired, but this can now be done without the repair time impacting availability. It can also be carried out without a special urgency, leading to a lower operational cost.

Now the MTTR is 50 ms instead of 24 hrs. This makes a significant improvement to the availability, due to the much shorter MTTR. A precise calculation needs to take into account the additional individual links and nodes of the repair path and their failure rates.

7.1.6 Network Failures

There is little information publicly available about failure types and their causes in a real operational network. Markopoulou et al. [6] have published information of failures in the Sprint IP backbone network, and characterized these failures. Even though a nationwide large backbone network is in many aspects not comparable to a mobile backhaul, it is useful to review the main findings. In [7], IP Network downtimes (router failures) are analyzed by Kuusela et al.

The Sprint backbone uses IP restoration, and on the underlying layer (optical fibre), there is no protection deployed. The IP layer is highly meshed. IS-IS protocol is used, and restoration occurs through IS-IS calculating alternate routes. It is possible that the backup paths become congested, in which case packet loss may occur. Temporary routing loops are also possible. During normal operation, QoS objectives are reported to be met due to almost no queueing delay and negligible jitter.

20% of all failures occur during scheduled maintenance. Of the unplanned failures (80% of all failures), about 30% concern multiple links (a shared link failure), while 70% only affect a single link. Shared link failures are further categorized as router-related, optical-related, and unspecified.

Frequent short failures have a larger adverse effect on IP connectivity than a single long lasting failure, due to the network convergence needed after a failure. With a well working restoration, after the network has converged, there is no impact to connectivity, as opposed to the frequent recalculation of routes due to link state often changing. Links could also be divided into a high failure and a low failure category, with high failure links (2.5% of all links) being responsible for over 1/3 of all of the unplanned failures.

During a maintenance window, long failures were experienced (minutes to tens of minutes or hours), due to human intervention, router reboots etc. Maintenance windows are scheduled during the least busy hours, which is common practice in the mobile network as well.

For mobile backhaul, several key differences exist compared to the case presented. First, topology in the access is typically hub-and-spoke and not mesh. With hub-and-spoke, often no backup physical links are available for the spokes, at least before the aggregation tier. Second, the case presented results of IP restoration using IS-IS. Mobile backhaul is more heterogenous with a number of alternatives for protection and restoration.

7.1.7 Human Errors

Both node and link failures have been considered. It is easy to grasp that a node fails (due to a power outage, HW or SW failure or similar) or that a cable is accidentally cut during infrastructure works. What is not so easy to model, is that network planning, operation and maintenance, node configurations, SW upgrades, etc, are tasks that people carry out. During these tasks, errors happen. In many cases, network downtime is caused by this type of incidents.

Investing in highly available, redundant network nodes and protecting links increases availability up to a point. Investment in people, competencies, processes, tools and operating practices may be required as well.

One topic of importance is to consider what is the level of complexity added, when network components are made resilient. As the complexity grows, also the likelihood for failures in configuration and operation increases. Recovery mechanisms can be complex, especially if

multiple protocol layers are involved, or if the operation of the mechanism itself is not commonly understood and documented.

7.2 Native Ethernet and Resilience

Native Ethernet concepts, MAC address learning, flooding of unknown unicast frames, and spanning tree based resilience are intended for the local area. Correspondingly, Carrier Ethernet and resilience in Metro and Wide area is addressed separately. Why would we then discuss native Ethernet further?

First, BTS or controller sites may deploy redundant L2 switches. Within this site LAN, native Ethernet functionality may be needed. One reason is cost: cost of an Ethernet port in a L2 bridge is typically lower than the cost of an Ethernet port in a router. And redundant switches are needed for high availability.

Similarly, as e.g. E-LAN service (multipoint) emulates a L2 bridge, it is useful to review how L2 bridging works. Rapid spanning tree protocol (RSTP) is also included in the MEF definitions as a possibility for supporting resilience.

7.2.1 *Ethernet Bridging*

At the Ethernet layer (L2), native Ethernet bridging relies on MAC address learning, flooding of unknown unicast traffic, and spanning tree protocol for ensuring a loop free topology. With this approach, a single active topology exists, and load sharing is not possible. In a redundant L2 topology, restoration occurs by spanning tree calculating a new loop-free L2 topology. The result is a tree topology (hub-and-spoke): Frames from all stations (leaves) pass through the hub (the root of the tree).

With L2 bridging, there has to be only a single L2 forwarding path at a time. If two Ethernet bridges were to be connected in a way that the L2 frame finds its way back to the originating bridge (via a different port), a loop is formed. Unknown unicast frames are flooded out on all ports, as well as broadcast frames. If a loop exists, L2 frames circulate endlessly until the configuration is modified.

The reason why multiple paths may exist in the first place, is resilience. An alternate link may carry the traffic if the primary link fails. Similarly, if a bridge fails, uplink connection might be supported via another bridge. In these cases, one or more ports need to be blocked in order to avoid a loop. The control protocol standardized by IEEE for this purpose is spanning tree [8]. If a link fails, active L2 topology is changed and a blocked port changes state to forwarding, to mitigate the failure.

When the size of the L2 broadcast domain (VLAN) grows, unknown unicast and broadcast frame flooding becomes more visible, as every station in the VLAN receives frames from the other stations. Part of this traffic is essential control plane traffic. An example is ARP, which is needed in the initial phase to create the IP address to MAC address binding. Broadcast traffic exists, even though unknown unicast traffic would be rare. The number of stations in a broadcast domain typically is limited.

Spanning tree has many variants. Furthermore, some commonly used versions are not standardized by IEEE but are proprietary. Also, a number of enhancements were first introduced into the protocol as vendor specific functionality [9]. Initial spanning tree protocol is defined in IEEE802.1d. Rapid Spanning Tree (RSTP), IEEE 802.1w [10], improves the

performance by a faster selection of the root bridge, and by faster state changes for bridge ports that are connected to hosts. Multiple Spanning Tree (MSTP), IEEE802.1s [11], supports up to 64 spanning tree instances. Each instance can have a different root bridge, allowing VLAN based load sharing.

7.2.2 Spanning Tree Operation

While each of the variants has its own characteristics, the basic behaviour is similar. Spanning tree creates a logical tree topology by blocking ports out of the active forwarding configuration. It is effectively a distance-vector protocol. No bridge has a complete view of the network.

BPDUs (Bridge Protocol Data Unit) are control frames bridges exchange. The types of BPDUs include configuration BPDUs and topology change notification BPDUs. A structure of a configuration BPDU is shown in Figure 7.5.

Octet	
1	Protocol identifier
2	Protocol version identifier
3	
4	BPDU type
5	Flags
6	
7	
8	
9	
10	Root identifier
11	
12	
13	
14	
15	Root path cost
16	
17	
18	
19	
20	
21	Bridge identifier
22	
23	
24	
25	
26	Port identifier
27	
28	Message age
29	
30	Max age
31	
32	Hello time
33	
34	Forward delay
35	

Figure 7.5 Configuration BPDU [8].

Spanning tree relies on the principle that BPDUs are received constantly. Otherwise a failure is assumed, and spanning tree aims at correcting the topology accordingly. If for some reason the network and the link is still operational, but the BPDUs were lost, a loop may be created when a blocked port moves to a forwarding state.

One bridge is selected as the root bridge. Bridge ID defines which bridge becomes the root bridge. Bridge ID contains a bridge priority subfield and a bridge MAC address. The bridge with the lowest priority becomes the root bridge. Priority field is 16 bits with a configurable value (default value being 32768). If priority fields of bridges are equal, the bridge with the lowest MAC address becomes the root.

Selection of the root bridge is important, as all traffic exits the broadcast domain (VLAN) via the root bridge. For every domain there has to be a single root bridge. (With MSTP, one can configure different root bridges for different instances. Each instance contains a selection of the VLANs.)

Each other bridge (non-root bridge) will select its lowest cost path to the root. Lowest cost path is calculated, based on information received from the other bridges. Every bridge along the way adds its own cost to the root before announcing it further. The port which has the lowest cost to the root bridge becomes an active port (root port). Any other ports towards the root are in a blocking state. With each LAN segment, there is a single designated port that is in forwarding state.

Costs are based on the capacity of the links. The default cost values are shown in Figure 7.6.

RSTP defines new port roles: Alternate and backup ports. Root ports and Designated ports are forwarding, while other ports (Alternate, backup and disabled ports) are Discarding.

IEEE802.1d-1998 uses port states of Blocking, Listening, Learning, Forwarding and Disabled. With RSTP, the states Disabled, Blocking and Listening map to Discarding state, so with RSTP there are three states: Discarding, Learning and Forwarding [8]. An example topology is shown in Figure 7.7.

Transitions between states are guided by timers, Max age and Forward delay.

In the listening state, the bridge listens to BPDU messages from the other bridges. At the listening state, spanning tree has selected the port to be in the active topology (either a Designated or a Root port), but the port is temporarily in a Discarding state.

Bandwidth	Cost
< 100k	200 000 000
1M	20 000 000
10M	2 000 000
100M	200 000
1G	20 000
10G	2 000
100G	200
1T	20
10T	2

Figure 7.6 Cost of the links based on bandwidth [8].

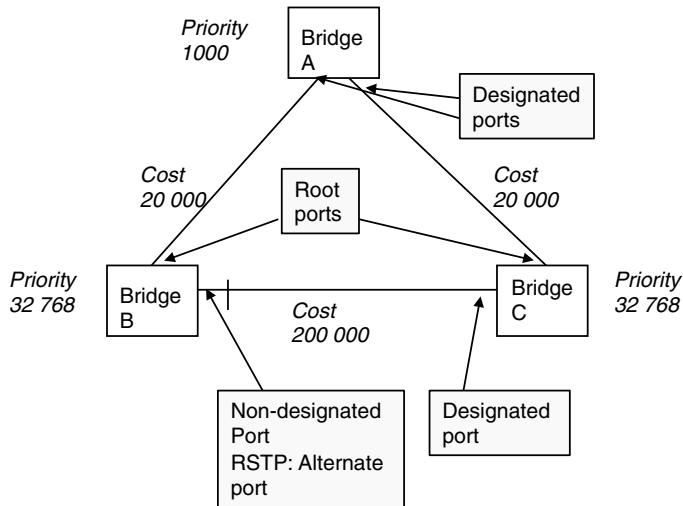


Figure 7.7 Example STP topology.

In the learning state, bridge starts populating its MAC address table with the source MAC addresses it observes. This reduces the amount of flooding when the bridge starts forwarding frames in the forwarding state.

With the default values assumed for the timers, convergence takes $\text{Max Age} + 2 \times \text{Forward delay}$ or $20 \text{ seconds} + 2 \times 15 \text{ seconds} = 50 \text{ seconds}$. See Figure 7.8.

With RSTP for bridge ports that are connected to hosts, listening and learning states can be omitted. This allows the bridge port to move faster to the forwarding state. Similarly, the operation is faster due to the support of a proposal flag and an agreement flag in the BPDUs.

With RSTP, also blocked ports send BPDUs as keep-alives. Missing 3 BPDUs with a default Hello time of 2 s, gives a 6 s detection time. Failure detection time is reduced, since instead of Max Age (20 s), missing BPDUs can be interpreted as a failure of a link or bridge. Depending on the case, further optimization is achievable.

With MSTP, one achieves a VLAN based load sharing function. Traffic is separated into different VLANs and these VLANs are then configured to MSTP instances. Each instance can have a different root bridge, and thus a different topology. Within any single VLAN, all traffic follows the same path.

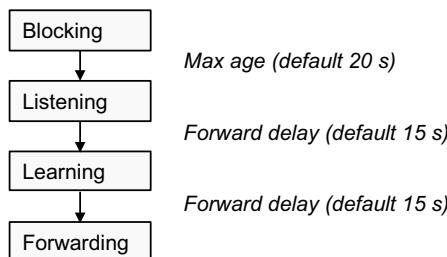


Figure 7.8 State transitions [8].

7.3 Carrier Grade Ethernet

7.3.1 Carrier Ethernet

Reliability is one of the attributes of the term Carrier Ethernet as defined by MEF [12]. How is reliability of Carrier Ethernet for the mobile backhaul then achieved?

The MEF service is delivered by a service provider. The provider network is carrier grade, and built with his technology of choice for resilience. The user sees the service through UNI characteristics. See Figure 7.9.

Provider's network is isolated from the customer networks and the resilience deployed by the provider is not visible to the customer. Similarly, the resilience within the customer site is separate from that of the provider network.

Standardization bodies have addressed the topic of carrier grade Ethernet with solutions that differ from the native Ethernet concepts. From the service user viewpoint, availability is simply one characteristic of the service. How it is implemented is a subject for the service provider.

Availability of Metro Ethernet services may be supported by multiple technologies in the provider network. In the case of Sonet/SDH, protection switching at the Sonet/SDH layer can be used. As with TDM in general, a drawback is the static and rigid capacity allocation. Virtual concatenation feature of NG-SDH makes the network more flexible, however the root of the issue remains.

Often Ethernet link aggregation helps achieving carrier grade resilience. Since the MAC layer does not see individual links of the aggregated connection, it is possible to support redundant ports, without involving the MAC layer.

ITU-T has defined two protection switching standards for the Ethernet: G.8031 and G.8032. These are intended for providing resilience in Ethernet networks.

IETF L2 VPN working group has produced a set of specifications that define an implementation of the L2 service with an IP/MPLS network. VPLS defines a multipoint service (E-LAN) while VPWS defines a point-to-point (E-Line) Ethernet service. Recovery is based on IP and MPLS.

7.3.2 MEF Services

For the user of MEF services, availability and QoS are characteristics of the service, and as such, they are more important than information of how protection or restoration is actually implemented. QoS and availability of an offered service can be defined in a service level specification, which is part of a service level agreement. The attributes of the service at the UNI are relevant.

For the service provider implementation of resilience is a key issue. As MEF compliant Ethernet services may be delivered over different technologies, resilience is tied to the technology used in the provider network. It also depends on the type of service, E-Line,



Figure 7.9 Provider network and customer network.

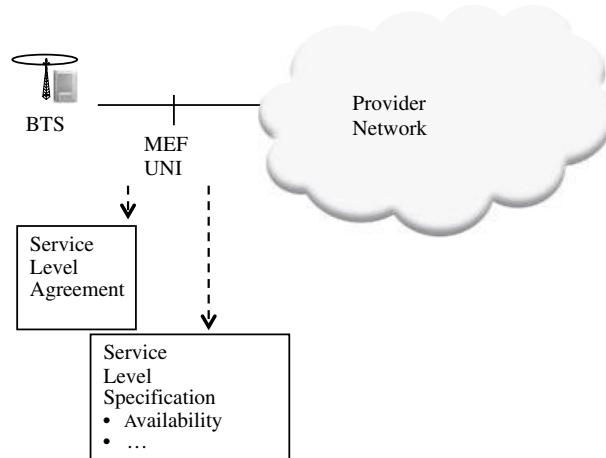


Figure 7.10 Availability attribute in a MEF SLS.

E-LAN, or E-Tree. Example technologies are Ethernet over Sonet/SDH, Ethernet over MPLS, Ethernet over DSL and a point-to-point Ethernet. For the Ethernet layer, link aggregation and RSTP are mentioned in MEF [13].

As MEF looks at the situation mainly from a service viewpoint, availability is an attribute that can be negotiated with the customer and the service provider. This attribute is in the scope of a Service Level Specification (SLS) [14] as shown in Figure 7.10.

The value of the attribute is then subject for negotiation. It is defined in MEF as specific to the Class of Service (CoS) label. This allows separate availability for each CoS type.

Having the availability attribute defined in the SLS is the main tool for the user of the service. Primarily, if the BTS or other customer equipment needs a higher availability for the MEF service, it is a subject of negotiation with the provider.

7.3.3 Ethernet OAM

Ethernet OAM (Operations, Administrations and Maintenance) provides for failure detection rather than for recovery itself. With MEF services, it is a tool that can be used to detect failures and monitor connectivity. Ethernet OAM can be used as a trigger for protecting switching. In IEEE, Spanning Tree Protocol is intended for the recovery.

Ethernet OAM was introduced in Chapter 4. Figure 7.11 shows an example application. An Ethernet service (e.g. E-Line) provided by a service provider is used between a BTS and an IP router.

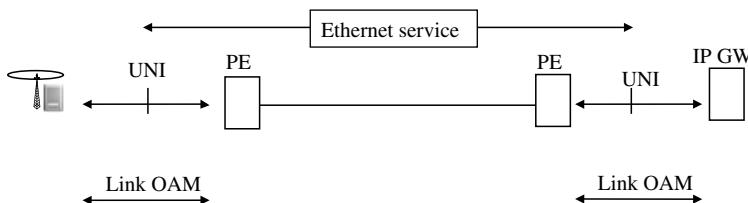


Figure 7.11 Link OAM between BTS and PE.

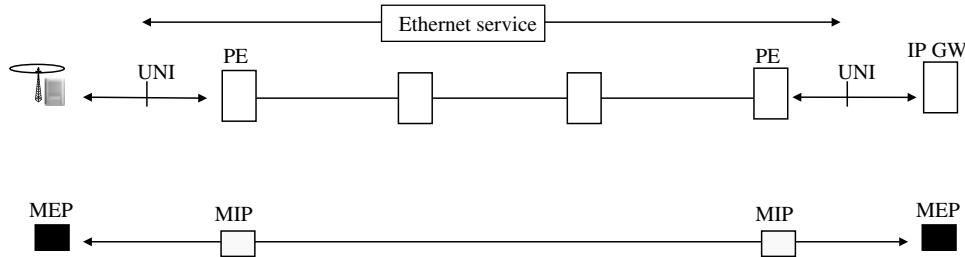


Figure 7.12 CFM for a mobile backhaul Ethernet service.

Ethernet link OAM [15] monitors the first mile from the BTS to the PE device. Link OAM is capable for single-hop application only, so the Link OAM is terminated at the PE device.

An example of using CFM [16] in the same case, is shown in Figure 7.12. The mobile operator configures a MEP (Maintenance End Point) to the BTS and to the peer entity, and can now use connectivity monitoring. The mobile operator sees the PE as a MIP (Maintenance Intermediate Point). Provider network however is not accessible to the customer.

For connectivity monitoring, connectivity check messages (CCMs) are sent regularly. For fast failure detection, the message interval can be as short as 3.3ms. Values are configurable. A failure is detected, if three CCM messages are lost. An alarm is raised as a consequence, additionally, Ethernet automatic protection switching application may use Ethernet OAM as a trigger. The minimum value mentioned (3.3 ms message interval) is aggressive. If the links are not constantly up but may be flapping, a false trigger is easily caused.

While CCM is operated constantly as a keep-a-live at Ethernet level, Ethernet loopback is comparable to ping: a MEP or MIP is asked to reply. Ethernet link trace corresponds to a trace route application. CFM protocols relate to a service (VLAN) and to a maintenance domain.

In cases where the failure cannot be recovered immediately e.g. by protection switching Ethernet level ping and traceroute help in troubleshooting. This shortens the time needed to find the faulty link or bridge, and consequently increases availability.

Ethernet Connectivity and fault management (CFM) supports nested maintenance domains. Maintenance domains allow customer and provider domains to be kept separate. A service provider may use connectivity and fault management within his maintenance domain, to monitor the Ethernet layer connectivity within his network – provided that the service is implemented using Ethernet technology. If the service is implemented using MPLS (or NG-SDH), service provider uses Sonet/SDH or MPLS layer OAM mechanisms instead of the Ethernet OAM.

7.4 IP Layer

7.4.1 VRRP

VRRP (Virtual Router Redundancy Protocol) allows two (or more) separate routers to present themselves as a single virtual router to the hosts. One router is always active (Master Virtual Router) while the others are passive in a VRRP group. IP packets from hosts are sent to the default gateway address which is the virtual IP address of the VRRP group. VRRP protocol [17] (RFC5798) supports both IPv4 and IPv6.

The VRRP group consists of a master virtual router and one or more backup virtual routers. VRRP messages are sent within the routers in the group. With IPv4, the master router sends VRRP advertisements using an IP multicast address 224.0.0.18 with a protocol number of 112. For IPv6, the assigned multicast address is FF02:0:0:0:0:12, with 112 as the IANA assigned number for the IPv6 Next header field.

During a failover of the master router, a backup router with the highest priority is elected as a new master. When the original master recovers, it then again assumes the master role.

VRRP master router responds to ARP requests (IPv4) and replies with a virtual MAC address. This allows the same IP and MAC address to be kept, irrespective of which of the routers assumes the master virtual router role.

With VRRP a single router is active while others are passive. For load balancing over all routers, multiple VRRP groups may be created, with each group having a different VRRP master.

7.4.2 Load Sharing

One of the benefits of IP is the capability of having multiple active links simultaneously in a load sharing configuration. By this way, network resources are consumed more efficiently compared to an active-passive operation, where some of the links are idle, while others may be congested.

Load sharing also reduces network downtime in case of failures, since recovery is faster. If one of the links fails, the other links can continue to be used, as they are already in the active forwarding table.

As forwarding is unidirectional, the return path may be different. So with load sharing, return traffic may use a different link. This is an issue for all stateful devices. A stateful firewall is an example.

Equal cost multipath is supported e.g. by OSPF and IS-IS routing protocols. When multiple paths to the same destination exist with the same entire cost, a load balancing algorithm distributes traffic to the links. Load sharing is discussed in RFC 2991 [18] and RFC 2991 [19].

Two types of load balancing exist: Per-packet and per-flow load balancing. Per-packet load balancing forwards each packet individually to the outgoing links, e.g. in a round-robin fashion. The downside of this approach is that packets may arrive out of order due to different transmission delay on the links. As well, path MTU may be different. Per-flow load balancing is thus often preferred.

Per-flow load balancing uses the IP packet header fields to identify a flow, e.g. source and destination addresses and protocol type (3-tuple). Also layer-4 ports may be used (5-tuple). Algorithms are not specified. Recall from Chapter 3, that for GTP-U, the UDP destination port number is 2152, while the source port is locally allocated by the sending node. So, e.g. on the S1-U (eNodeB – SGW), load sharing depends on the implementation.

7.4.3 Routing Protocols

IP control plane means the use of any routing protocol: OSPF, IS-IS, RIP, BGP - and so on. IP control plane, with a topology of redundant links, makes IP forwarding resilient to failures in links and nodes in the network. When IP control plane is used for MPLS as well, MPLS benefits from the resilience routing provides.

Details of operation of routing protocols differ. For recovery, the failure first needs to be detected. This may occur via a physical or link layer indication, via the use of a routing protocol, or via a dedicated detection protocol (e.g. BFD). Second, other network nodes have to be informed. Third, a new topology (best paths) needs to be calculated (shortest path first algorithm in case of OSPF and IS-IS). Finally, the new topology needs to be taken into active forwarding.

Distance-vector routing protocols (e.g. RIP) send updates periodically. Link-state routing protocols (e.g. OSPF and IS-IS) respond to changes (e.g. losing a link) which then trigger updates. Depending on the application, link state protocols are often more attractive, due to the improved characteristics.

Distance-vector protocols do not maintain a complete topology of the network. They also in some cases suffer from a count-to-infinity issue: a route is advertised back and forth between two routers, with an increasing metric. Convergence after a failure is not complete, until the metric (e.g. hop count) reaches its maximum. Only then both of the routers conclude that the destination is not reachable. Count-to-infinity is addressed by split-horizon (not advertising the same route back) and a poison reverse (advertising the route back with the maximum metric). See [20] for routing protocols.

Due to the faster response, link-state protocols are preferable for fast recovery. OSPF or IS-IS is also needed for potential traffic engineering applications. The main drawback of link state protocols is that they in general are more complex than distance-vector protocols. In the next section, OSPF is discussed in further detail.

7.4.4 OSPF

OSPF (Open Shortest Path First) supports a two-level hierarchical area structure with a backbone area, and non-backbone areas. Full topology is only maintained within an area, with less information available from other areas. After a change, Link State Advertisements (LSAs) are flooded within the area, and routers update their Link State Databases (LSDBs). Consequently, routers run Dijkstra's Shortest Path First (SPF) -algorithm placing themselves as the root, select the shortest paths to the destinations, and store the results into an OSPF routing table. OSPF consists of essentially three subprotocols: Hello, Exchange and Flo.

OSPF is defined for IPv4 as RFC2328 (OSPF v2) [21] and for IPv6 in RFC 5340 (OSPF v3) [22]. RFC5838 [23] adds support of non-IPv6 address families to the OSPF v3. An example is the IPv4 address family.

The section is based on OSPF v2, however, OSPF v3 in many cases uses the same concepts and procedures. One difference is that OSPF v3 operates per link. Link local addressing is used. Note that also IS-IS, although not discussed further here, is in many aspects comparable to OSPF.

OSPF essentially uses three subprotocols: Hello, Exchange and Flooding. Hello protocol monitors link availability, and elects Designated and Backup Designated routers when needed (broadcast/non-broadcast multiaccess networks). Exchange protocol brings up the adjacencies and synchronizes databases. Finally, flooding is used to update the database after the initial synchronization. OSPF uses cost as the metric, and this can be configured administratively for each network link. The cost can reflect bandwidth or delay or it may simply be configured to match the preference as seen by the network administrator. OSPF has been enhanced for a Traffic engineering application (OSPF-TE), supporting additional parameters like bandwidth of the link in the OSPF-TE metrics.

Within OSPF essentially three subprotocols exist: Hello, Exchange and Flooding. Hello protocol establishes and maintains neighbour relationships, elects Designated and Backup Designated routers when needed (broadcast/non-broadcast multiaccess networks), and

monitors link availability. Exchange protocol brings up the adjacencies and synchronizes databases. Finally, flooding is used to update the database after the initial synchronization.

OSPF operates within an Autonomous System (AS) – within an administrative domain.

One benefit of link-state protocols, such as OSPF, is the capability to react to changes in the network, e.g. link failures. Link state information is propagated in Link State Advertisements (LSAs). LSAs are flooded to all routers in an area, and the flooding is triggered by changes in the network.

As a consequence, all routers have a complete view of the topology, including the state of the links, of an area. OSPF routers collect this information into a LSDB from the LSAs. Link state database maintains a complete information of the state of links in an area. Shortest path first – algorithm is run to select the shortest paths to each destination network. The results are stored in the OSPF routing table.

OSPF LSAs carry network and subnet mask information, and thus allow variable length subnetting. Longest prefix match routing is used.

Large networks are divided into areas. Routers at the border of an area assume a special role and a separate LSDB is maintained for each area a router connects to. LSAs of Type 1 and Type 2 are not flooded into other areas. Instead, Type 3 Summary LSAs advertise networks from other areas. Inter-area routing takes place via the OSPF backbone (area 0).

AS-external routes are informed by external LSAs.

LSAs are triggered when there are changes in the network, e.g. a router's interface goes down. LSAs are also refreshed periodically after LSRefreshTime (30 minutes in the specification), even if there are no changes in the contents of the LSA. LSAs age out after MaxAge (60 minutes), if not refreshed.

LSAs are carried as OSPF packets. The OSPF packets are:

- Hello packets.
- Database description.
- Link state request.
- Link state update.
- Link state ack.

All of the OSPF packets may carry a list of LSAs, except the Hello packet, which is meant for discovering and maintaining neighbour relationships. Hello protocol includes HelloInterval and RouterDeadInterval parameters, which need to be agreed by the routers. HelloInterval means the interval Hello messages are sent. RouterDeadInterval means the time routers wait until declaring a neighbour dead, after not receiving a Hello packet. RouterDeadInterval is a multiple of HelloInterval. The parameters values are important, if Hello packets are used for failure detection.

Database description summarizes database contents and is used when neighbour adjacency is initialized. Link state request packet asks for LSAs that are outdated. Full adjacency is reached, when databases have been synchronized.

Link state update implements the flooding of LSAs. A single link state update packet may include several LSAs. Link state ack packet acknowledges an LSA. Flooding of LSAs is reliable, because of the acknowledgement. The header of the LSA is included in the information field of the Link state ack message.

LSA format depends on the type of LSA. The LSA types include (list is not complete):

- Type 1 Router LSAs – Describe the state of router's interfaces, including their cost.
- Type 2 Network LSAs – Describe routers attached to the network.

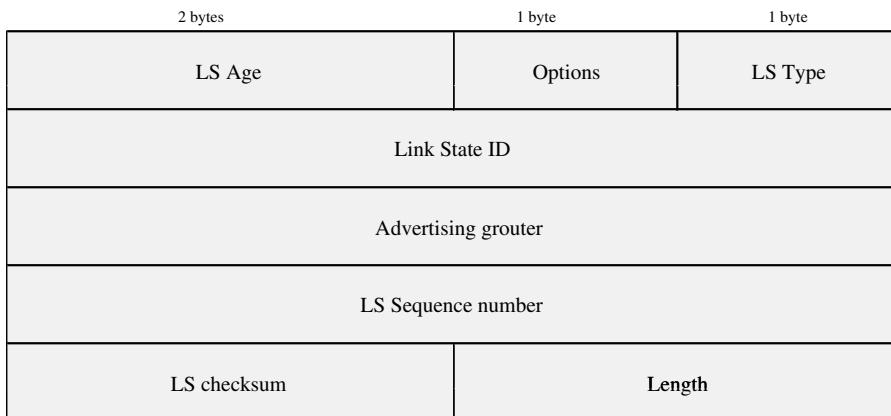


Figure 7.13 LSA header.

- Type 3 and 4 Summary LSAs – Describing routes to other areas. Type 3 LSA describes routes to networks and Type 4 routes to ASBRs (Autonomous System Boundary Routers).
- Type 5 – AS-External LSAs, from ASBRs. Autonomous System Boundary Routers create Type 5 and Type 7 external LSAs.

The LSA header (20 bytes) is shown in Figure 7.13.

The header fields are:

- LS Age – time in seconds since LSA originated.
- Options – indicates optional OSPF capabilities (see RFC for further detail).
- LS type (see previous paragraph; Type 1 Router, Type 2 Network LSA and so on)
- Link State ID – depends on the LS type.
- Advertising router – Router ID.
- LS Sequence number – used for detecting old or duplicate LSAs.
- LS Checksum – a checksum that is calculated over the whole LSA content and header, excluding the LS Age field.
- Length – Length including the header.

OSPF packets run directly on top of IP with a protocol number 89. OSPF Hello packets are sent to the multicast address of AllSPFRouters (224.0.0.5). AllIDRouters (224.0.0.6) address refers to designated routers and backup designated routers. To ensure that these packets are not forwarded, the TTL (Time-to-live) field is set to 1.

OSPF packets are control traffic and thus should be given priority in the Diff Serv architecture by marking them as internetwork control.

OSPF packet header (24 bytes) is shown in Figure 7.14.

The OSPF header fields are:

- Version means the version of the OSPF protocol, which in RFC 2328 is 2. (For IPv6, Version 3 is specified)
- OSPF packet types (see previous page).

1 byte	1 byte	2 bytes
Version	Type	Packet Length
Router ID		
Area ID		
Checksum		Authentication type
Authentication		
Authentication		

Figure 7.14 OSPF header.

- Length: Length of the OSPF packet, including the header.
- Area ID: Identification of the OSPF area.
- Checksum is calculated over the OSPF packet header, excluding the authentication fields, and content.
- AuType identifies the authentication type, the subsequent field is reserved for the authentication: 0 means no authentication. 1 means simple authentication, e.g. a password. 2 means cryptographic authentication. RFC2328 defines the use of MD5 (Message Digest).

RFC 5709 [24] defines additional authentication options:

- HMAC-SHA-1.
- HMAC-SHA-256.
- HMAC-SHA-384.
- HMAC-SHA-512.

Implementation of RFC5709 assumes HMAC-SHA-256 to be mandatory.

Time required for recovery depends on a number of factors: Detection, flooding of the LSAs, calculation of the SPF, updating the routing table, and recovery in forwarding. For a fast recovery, LSAs should be flooded immediately. However a guard time may be needed before an LSA is flooded.

Two types of timers are defined: Single shot timers, and interval timers. Single-shot timers are used after an event is detected. When the timer expires, the event is processed. Benefit of not processing the event immediately, is to allow some guard time in case of flapping links. Reacting immediately to each separate event would cause a new flood of LSAs each time the link changes state. Drawback of the timer is naturally the longer time needed for the network to converge. In the original (dated) RFC the timer granularity is 1 sec.

When the routers receive the LSA they update the information in their LSDB. After the update, the SPF algorithm is calculated. Processing time depends on the size of the network, on the amount of routers, networks and links. To speed up convergence, information about networks that are not used should not be flooded. However basically the whole area receives the same information. With stub areas external LSA flooding may be limited. An area can be configured as a stub area if there is only a single exit point. Many variants of stub areas exist (not all standardized), allowing also further reduction of LSAs.

Interval timers define the sending interval for packets that need to be sent at regular intervals like OSPF Hellos. For this timer, the granularity is similarly 1 second. In practice, smaller values are supported. For fast detection, BFD may be used instead of optimizing the Hello interval. Relying on OSPF hellos, the detection is on seconds range, depending on the parameter values. With BFD, sub-second detection can be achieved.

For a calculation of the SPF algorithm, it is beneficial to have the LSDB only with routes that are really needed. If the LSDB holds a large amount of information, also changes are more frequent, SPF calculation is needed often, taking more time. Information from other OSPF areas can be reduced using summarized routes at the ABRs (Area Border Routers). In general, information from other areas is condensed by the ABR, and full topology is not given in the Type 3 Summary LSAs. For stub areas, as mentioned, external LSAs are not allowed, however other information is still received. LSDB size is as well impacted by the selection of the network type (point-to-point, broadcast, non-broadcast multi-access NBMA, point-to-multipoint or virtual link). The amount of information within the LSA varies accordingly.

7.4.5 BFD

Bidirectional Forwarding Detection (BFD, [25]) is specified as a light-weight mechanism for detection of failures over any L2 or L1 media. It monitors the liveness of the path between IP forwarding elements. Routing protocols relying on Hello packets tend to be slow. For faster detection, BFD is well suited. Routing protocols can then use BFD for a faster detection of failures over the path.

BFD is also essentially a Hello protocol. Messages are transmitted periodically between the two end-points. If an end-point is not receiving these messages from the other end, it assumes a failure somewhere in the forwarding path. BFD control packets, and BFD echo packets are defined.

A few different operating modes are supported. In asynchronous mode, systems transmit packets periodically. If packets are not received in a certain time, a failure is assumed. In demand mode, it is assumed that some way of monitoring connectivity exists, and control packets can only be sent on demand. In that case BFD messages are exchanged after which the system goes back to the quiet mode. Additionally, Echo mode allows echoing BFD messages back from the other system. If a number of these messages are not received, a failure is assumed.

BFD supports four different states. Admin down means administratively down. The other states are Down, Init, and Up.

BFD is specified for IPv4 and IPv6 over single-hop (RFC 5881) and for multi-hop (RFC 5883) [26], [27]. Single-hop BFD monitors the path to the next hop router. Multi-hop BFD monitors the path over several hops. BFD is encapsulated over UDP. IP source and destination addresses and UDP ports identify the BFD session. Single-hop BFD uses UDP destination port 3784 for BFD control packets and port 3785 for BFD echo packets.

BFD supports authentication. The options are simple password, MD5 or SHA-1 authentication.

7.4.6 Further Topics

The mentioned OSPF RouterDeadInterval is typically e.g. four times the HelloInterval. BFD relies on a similar type of lightweight Hellos. When the link fails (a cable is accidentally pulled off or the site suffers a power outage), no messages can be sent and the failure is detected by missing Hellos.

With BFD, a more aggressive detection is possible and this is beneficial since it shortens the recovery time. Sometimes a link may be flapping; i.e. the state is not permanently down, but is oscillating before (maybe) stabilizing into either state. Having state changes propagated during short intervals through the network causes LSA flooding (in the case of OSPF) and consequently, SPF calculations. Clearly, the detection should not be too aggressive. The value depends on the network.

For BGP route dampening is defined in RFC2439 [28]. A figure of merit per route is incremented every time the route is withdrawn. Based on the figure, changes may be suppressed for those routes that are considered unstable. The underlying assumption is that routes that have been flapping, have a high probability of doing so in the future.

On the other hand, in stable networks RFC 4136 (OSPF Refresh and Flooding Reduction in Stable Topologies) [29] suggests that LSA flooding e.g. every 30 minutes is not necessarily needed, if the network is stable. This reduces OSPF control traffic. The implementation relies on a Do Not Age -Bit in an Indication LSAs as defined in RFC1793 [30].

Another issue is packet loss. What happens if Hello packets are lost e.g. due to congestion? Lost Hello packets trigger declaring neighbour down after RouterDeadInterval. In [31], it is proposed to tune the RouterDeadInterval according to Hello packet drop rate. This reduces false detection, if Hellos are lost due to congestion. However, if OSPF is marked as Network control (See QoS chapter of this book) and treated accordingly, Hello packets should not first suffer from congestion.

A further topic is that even though Hello packets (as network control) would go through also in the case of network congestion, service to the user can be poor. In the mobile backhaul case, the user is the radio network layer application.

To detect QoS impairments (increased delay and packet loss), performance measurement is needed. RFC4656 defines a One-Way active measurement protocol (OWAMP) [32], and RFC5357 [33] a two-way active measurement protocol (TWAMP). The protocol supports an active measurement that can be used to monitor QoS. Packet delay (Round-trip-time) and packet loss can be measured. One component is the TWAMP sender, and another one is the responder. In the mobile backhaul, TWAMP can be used e.g. to estimate the QoS between a base station and a controller/GW.

Figure 7.15 depicts a TWAMP application for the mobile backhaul. A TWAMP sender (eNodeB in the figure) sends the initial UDP/IP packet with a transmit sequence number and a timestamp. A TWAMP responder (e.g. a router/gateway) replies and adds his sender sequence number and timestamp. Alternatively, TWAMP functionality can be supported by a cell site router.

TWAMP supports measurements per QoS class, as indicated by the DSCPs. This is important as well. If network is congested, background traffic may suffer delays even though voice and control traffic is carried without impairments.

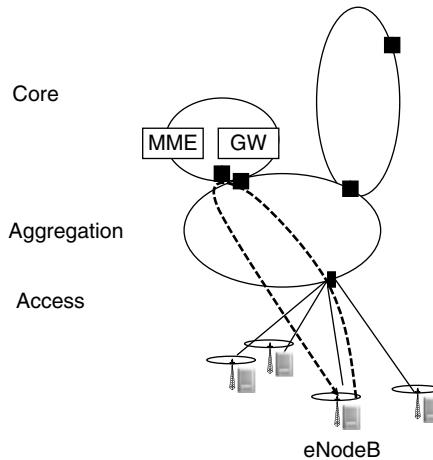


Figure 7.15 Example TWAMP application.

It may also be used to monitor SLA. Violations of SLA are otherwise difficult to detect. IP SLAs are discussed e.g. by Benlarbi [34] and Kilpi [35].

7.4.7 Loop Free Alternates

With IP routing protocols, a single best path is typically used. A network change causes then a recalculation/selection of the new best path, which, depending on the size of the network, takes an amount of time.

Recovery time can be reduced, if a ‘second best path’ is calculated in advance, so that when the best path fails, the router immediately knows which next hop should be used instead. Clearly, this second best path has to be loop-free.

RFC5286 [36] defines an IP Fast Reroute (IP FRR) – comparable to the operation of MPLS FRR. The goal is in protecting IP unicast traffic over MPLS/LDP networks. Back-up loop-free next hops are calculated in advance, and are used to mitigate a failure. Convergence of the routing protocol still takes place in the background. During convergence, traffic is forwarded towards the back-up next hop. This reduces packet loss during convergence.

In addition to the MPLS FRR, a non-standardized routing protocol Enhanced Interior Gateway Protocol (EIGRP) includes a definition of a feasible successor [37].

7.5 MPLS Resilience

In the following section resilience topics related to MPLS are discussed. Further reading can be found e.g. by De Ghein [38], Minei and Lucek [39] and Guichard, Le Faucheur and Vasseur [40].

7.5.1 Label Allocation

With IP forwarding, MPLS Labels are assigned by LDP [41] to the destination networks. Information of the destination networks is obtained via an IGP.

For the label distribution two operating modes exist: downstream on demand and unsolicited downstream. In both cases labels are allocated locally and LDP delivers the locally assigned label to the upstream node. In downstream on demand mode the label is allocated only as a response to a request. In unsolicited downstream the downstream LSR allocates the labels to all destination networks it knows of without a specific request.

When labels are allocated to each destination network and all upstream neighbours are informed of the labels, how does the upstream LSR decide what to do with multiple labels to the same destination network? With conservative label retention labels that have not been requested, will not be kept. In liberal retention mode all labels that have been received are kept.

In RFC3037 [42], a motivation is given that downstream on demand with conservative label retention should be used when labels are a scarce resource, and should be conserved. ATM and cell mode –MPLS is given as an example. In frame mode MPLS this is not the case. Standard allows all combinations.

Liberal retention mode has the benefit of having labels already allocated, in case of failure towards the next LSR. If the destination network can be reached via another link or node, the label is already available and can be used, without having to first exchange label information with LDP.

In Figure 7.16, labels for the destination network 10.0.0.0/26 have been allocated locally by each LSR. LSR Delta has given a label value of 68 to the destination network 10.0.0.0/26, and this value is announced by LDP to both upstream neighbour LSRs, Beta and Gamma. The example uses per-platform label space: same label is advertised on all interfaces by LSR Delta.

LSR Beta and LSR Gamma similarly locally select a label and advertise this to the LSR Alpha. A value of 223 is allocated by Beta, and 33 by Gamma. LSR Alpha receives both of these labels. Shall Alpha now use Label 223 towards Beta or Label 33 towards Gamma ? This is selected by the IGP. In liberal retention mode, Alpha stores both label values, even though it would only use the one preferred by the IGP.

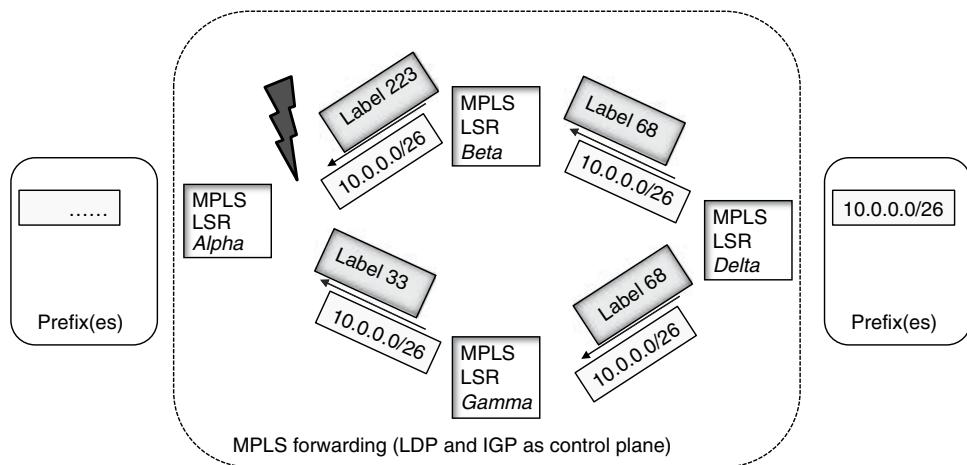


Figure 7.16 Liberal retention mode.

Assume that the IGP prefers Beta as the next hop. So LSR Alpha pushes a label with a value of 223, and forwards the labeled IP packet towards Beta. What happens if the link to Beta fails? Due to the change (link down), IGP convergence is needed. How fast this happens depends on the detection and on the IGP used.

When IGP has converged, LSR Alpha knows that the new best path to the destination network is via LSR Gamma. Due to the liberal label retention mode, the label is already available. Recovery is faster since LDP is not needed for obtaining a label binding. Alpha forwards traffic towards Gamma, with a label 33 now pushed instead of label 223.

If label 33 had not been available, LDP would be needed to obtain a label from Gamma before being able to use MPLS label based forwarding to Gamma.¹

7.5.2 LDP Sessions

As discussed in the previous section, liberal label retention is beneficial since it shortens the recovery time. User plane forwarding can continue immediately after IGP convergence.

Another issue is the operation of the LDP control plane. LDP session needs to be operational, as otherwise the user plane forwarding based on MPLS labels is disrupted as well.

LDP uses TCP, and each LDP session requires a TCP session. Multiple LDP sessions between LSRs may exist. When the LDP session is not with a directly connected LSR neighbour, it is a targeted LDP session.

LDP maintains a Hello adjacency with its neighbour(s). Neighbours may be directly connected or multiple hops away. When Discovery Hello messages are not received from a neighbour within a defined time, the LSR concludes that the peer has failed, or for another reason the peer not wishing to switch based on labels.

TCP session includes a Keepalive timer. Unless messages are received within a defined time, LSR considers the peer down, or that the session has failed, removes the LDP adjacency and terminates the LDP session. This is considered a fatal error.

7.5.3 IP MPLS VPN

In addition to what is provided for resilience within the ‘IP/MPLS cloud’, the PE-CE routing protocol and the routing peering with the provider should be considered.

In Figure 7.17, CE A1 is multihomed to PE x and PE z. IGP advertises customer prefixes reachable via CE A1 to both PE x and PE z. Assume that the BGP best path towards the customer networks behind CE A1, is via PE x. If the CE A1 – PE x link fails, BGP convergence is required to find the alternate path via PE z. BGP convergence may take time (tens of seconds and more) as it is a distance-vector (or, a path-vector) protocol and not optimized for a rapid recovery after a link state change. For a faster recovery in the IP MPLS VPN application optimizing BGP convergence is important. Solutions for this type of optimization exist. Additionally, an expired draft on BGP multipath exists ([draft-bhatia-ecmp-routes-in-bgp-02.txt](#)) [43].

¹ Assuming that the LSRs are capable of IP unicast forwarding, Alpha could have forwarded the IP packets to Gamma without any label. The IP packet would travel one hop without a label imposed, until the label is delivered from Gamma to Alpha.

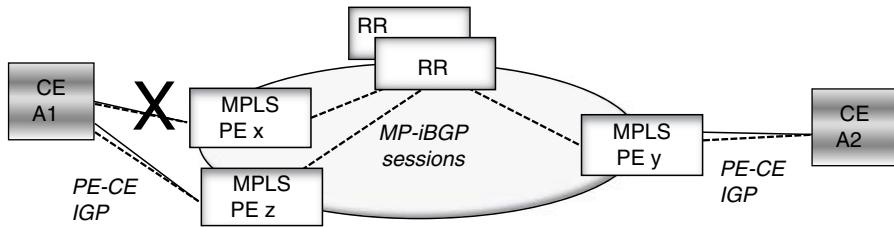


Figure 7.17 Multihoming to two PEs (control plane).

With OSPF as a PE-CE protocol, the redistribution of routes by BGP causes OSPF LSA types to change. Inter-area OSPF routes appear as external routes due to redistribution. OSPF however prefers intra-area routes. Any possible direct connectivity such as a back door link connecting two customer sites directly would be used by OSPF, even though in a normal case traffic would be intended to be routed through the MPLS network. The use of OSPF as a PE-CE protocol for the IP MPLS VPN application is covered in RFCs 4576 (Using a Link State Advertisement (LSA) Options Bit to Prevent Looping in BGP/MPLS IP Virtual Private Networks) [44] and 4577 (OSPF as the Provider/Customer Edge Protocol for BGP/MPLS IP Virtual Private Networks) [45].

In general, any routing protocol (or static routes) can be used between the CE and the PE (BGP, RIP, etc). The original RFC 4364 [46] assumes BGP is used.

In Figure 7.17 Route Reflectors (RRs) are used. Since the MP-iBGP sessions between the PEs rely on route reflection, RR redundancy is essential. Route reflectors are specified in RFC 4456 [47].

7.5.4 VPLS

In L2 VPN deployment, in addition to the resilience supported by the PSN tunnels, one may need to consider failures on the PE-CE attachment circuit (AC), and a failure of the PE. MEF2 states that different CE attachment redundancy mechanisms are not in the scope. It is however acknowledged, that this may be wished by some customers [13].

If the customer does not use L2 switching on the ports which attach to the PEs, technically a CE could connect to multiple PEs. The service provider however in general has no control of the customer network, yet it is essential for the operation of his own network that the multihomed CE configuration does not cause any adverse effects. Thus multihoming with VPLS needs attention.

Candidate solutions rely on having the multihomed CE with one active link to either PE (PE1 or PE2), the other link passive, and additionally a support for PW redundancy. See e.g. an approach described in [48], using multi-chassis link aggregation as the AC. Principle of the multi-chassis link aggregation is presented in Figure 7.18. While Ethernet link aggregation itself is standardized, the protocol between the two chassis is not, so it is implementation specific.

An internet draft ‘draft-ietf-l2vpn-vpls-multihoming-03.txt’ [49] proposes a BGP based solution, which is intended to be applicable also for LDP-based VPLS. Here again, the idea is that one AC and one PE is active at a time.

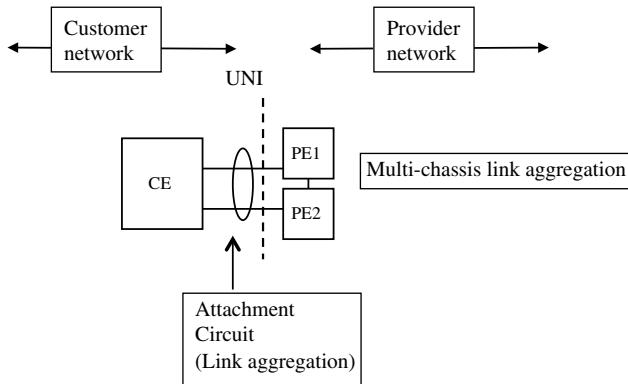


Figure 7.18 Multi-chassis link aggregation.

7.5.5 MPLS TE and Fast Reroute

MPLS TE was shortly introduced in Chapter 4. MPLS TE fast reroute (FRR) [50] allows a fast recovery, comparable to that of Sonet/SDH benchmark of 50 ms. MPLS FRR relies on local protection: in the case of a failure, traffic is locally directed via another path to a merge point. From the merge point onwards, the LSP uses its original path.

Local rerouting allows a fast recovery, since the failure of a link is often detected directly by the physical layer ‘signal down’ indication. (If the failure can not be detected directly, but e.g. via IGP or BFD, the recovery time grows.) The local LSR acts immediately without informing other nodes.

The protecting path needs to be set up in advance. FRR can be configured to protect against a link failure or a node failure. It can also support 1:1 or 1:N type of protection, depending on the configuration.

As an example, consider FRR link protection in the topology of Figure 7.19.

As LSR Alpha detects the link failure, e.g. no signal, it reroutes the labeled IP packet via the preconfigured back-up path. Label is swapped to 68 in the same way as would happen without the failure. Additionally, a new label, label 40, is pushed, and the packet is sent towards LSR Gamma. Gamma being the penultimate hop LSR, pops label 40, and forwards the packet to LSR Beta. Beta receives the packet with the same label as it would have in the normal operating condition. This is an example of a facility protection.

LSR Alpha is the point of local repair, as it detects the failure and reroutes the packet. LSR Beta is the merge point, as from Beta onwards, the packet follows its initial path.

Fast reroute takes place immediately. Simultaneously the MPLS-TE head end of the path is informed, and a new path computation occurs, with the information that the link Alpha-Beta is down. A new path avoiding the failed link (if available with the given constraints) will be set-up by RSVP-TE. In the meanwhile, packets use the locally repaired path. This shortens the time during which packets may be lost while the path computation and the new path establishment is ongoing.

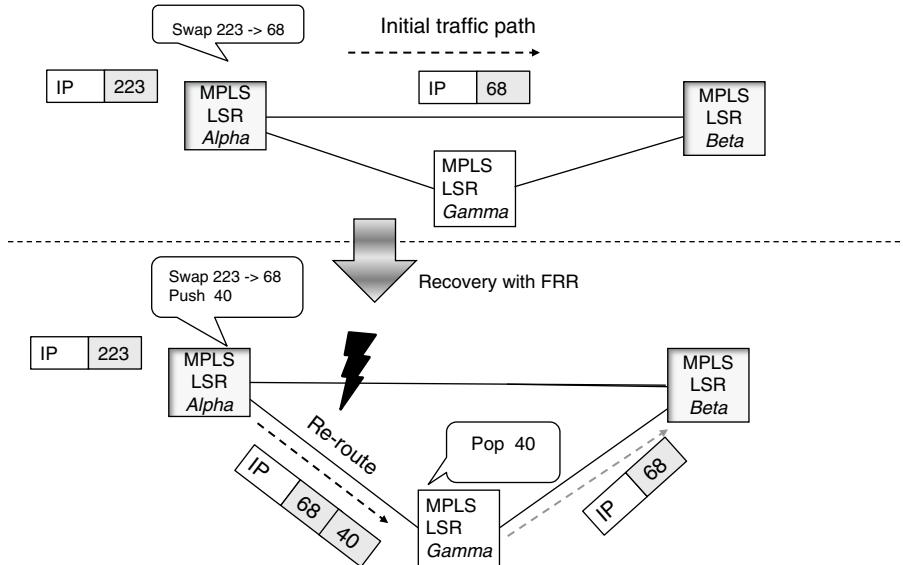


Figure 7.19 Fast re-route (Facility protection example).

7.5.6 MPLS OAM

RFC 4379 [51] defines an MPLS echo request, and an MPLS echo reply, that can be used for MPLS-ping and MPLS-traceroute applications. An MPLS echo request is an UDP packet, with TTL of 1 in the IP header.

The destination address in the IP header is an IPv4 address (any address) from the 127/8 network (internal host loopback addresses, RFC 1122). The source IP address is the address of the sender. The source port is selected by the sender, and the destination port is 3503. Use of this address ensures that the packet is not further forwarded. Also it allows the LSP packets to be identified.

The packet is then sent with the label stack of the FEC to be tested. LSP-ping is used to test the connectivity of the user plane. The LSP ping travels through the same path as the MPLS user plane (FEC). At the egress LSR, the packet is directed to the control plane, which checks that the packet belongs to the same FEC where it came from.

Additionally BFD is defined for fast failure detection on the data plane. RFC5884 defines BFD for LSPs [52] and RFC 5885 [53] defines BFD for the Pseudowire Virtual Circuit Connectivity Verification (VCCV).

7.5.7 MPLS-TP

MPLS-TP standardization is ongoing in IETF and there are many active drafts. Recently, RFC 6372 defines a Survivability Framework for MPLS-TP [54]. Protection capabilities are comparable to Sonet/SDH, so that both linear and ring protection are supported, based on fast detection. This makes MPLS-TP attractive also for mobile backhaul application. For

example, in the access/aggregation tier a Sonet/SDH type of fast protection switching with deterministic behaviour is achieved.

MPLS-TP includes an inband OAM channel, both at the LSP level and at the PW level. Connectivity control (CC) can be used to detect failures, and trigger protection switching. Connectivity verification (CV) can be used like LSP ping, to verify connectivity in a reactive mode.

RFC5586, MPLS Generic Associated Channel, generalizes the control channel of RFC5086, so that the same mechanism is applicable also to sections, LSPs and PWs [55]. RFC5085 defines a connectivity verification for pseudowires [56]. OAM control messages can then be exchanged using the new Generic Associated Channel (G-ACH). This can be identified with a new reserved label, Generalized Associated Label (GAL).

7.5.8 GMPLS Control Plane

As the name implies, Generalized Multiprotocol Label Switching (GMPLS) control plane allows for control of different types of underlying transport technologies [57], [58], [59]. This control plane essentially consists of RSVP-TE (Resource Reservation Protocol with Traffic Engineering extensions) and a link state routing protocol similarly with Traffic Engineering extensions (OSPF and IS-IS, [60], [61], [62]). With MPLS-TP, GMPLS control plane may be used but this is not mandated as MPLS-TP can as well rely on network management.

With GMPLS, switching capabilities can in addition to MPLS be:

- Layer-2 switch capable (L2SC).
- TDM capable.
- lambda-switch capable (LSC), or
- Fibre-switch capable (FSC).

This interface capability is one constraint. TE capability (and label information on that TE link) depend on the switching capability of the link.

Link protection type defines protection capabilities. These are presented in Table 7.2 from the lowest to the highest.

The information of link protection capability may be used in path computation, when setting up LSPs for MPLS-TE. Path selection algorithm then tries to find a path that meets at least

Table 7.2 Link protection capabilities [60].

Protection capabilities	Description
Extra traffic	Link is protecting another link or links. If any of the links it protects fail, LSPs will be lost.
Unprotected	No other link is protecting this link, and LSPs will be lost if the link fails
Shared	One or more Extra traffic type of links are protecting this link.
Dedicated 1:1	One disjoint link of type extra traffic protects this link
Dedicated 1 + 1	A dedicated disjoint link is protecting this link. Protecting links is not advertised in the LSDB and is not available for routing of LSPs
Enhanced	A protection scheme more reliable than dedicated 1 + 1, is used to protect this link

the minimum criteria. The information of link protection capabilities, is optional. If it is not received, it is unknown. For MPLS-TE, routing extensions are defined, into OSPF and into IS-IS.

7.6 Resilience in the BTS Access

In this section the access part of the backhaul is addressed – protocols and functionalities that allow building a resilient access tier.

7.6.1 BTS and BTS Site

As a network component, BTS itself is a single point of failure. Failure of a BTS, however, does not impact a large geographical area, and the failure of a single BTS is often at least partially compensated by radio coverage of the neighbouring BTSs – either by the same radio technology cells, or by another one (2G, 3G or LTE).

BTS sites are often shared between multiple BTSs. In these cases, transmission links may be wished to be shared as well. This is shown in Figure 7.20.

Sharing the physical link requires a cell site gateway for combining the traffic. Alternatively, the gateway function can be integrated into the BTS. For availability, the BTS integrated GW should be as independent as possible from the radio-related BTS functions.

When the backhaul service is common for multiple networks, such as 2G and LTE, the two systems become dependent on the same network components. The physical link, and the cell site GW, are shared. Synchronization is typically also obtained via the common gateway – via Synchronous Ethernet, or via packet based timing.

From availability point-of-view, the two radio networks should be independent, so that failure on one system does not impact the operation of the other. Traffic separation can be deployed to keep the traffic types from interfering. QoS has to be implemented on the gateway to guarantee priority for the critical traffic types (voice, real-time and control) in case of congestion.

Even if the site is shared, transmission links may be separate for the two BTSs. An example is an LTE eNodeBs that is added to an existing site (2G or 3G). If the existing transmission link is TDM based (e.g. 2 x E1/T1s), there is no capacity available for the eNodeB. As an expansion of the TDM network at best buys some time, a new packet network may be deployed as an overlay – delivering a high capacity Ethernet port to the site for the eNodeB to use. See Figure 7.21.

In this case, the two networks (2G and LTE) remain independent of each other from a transmission viewpoint. Failure on one transmission system does not impact the other.

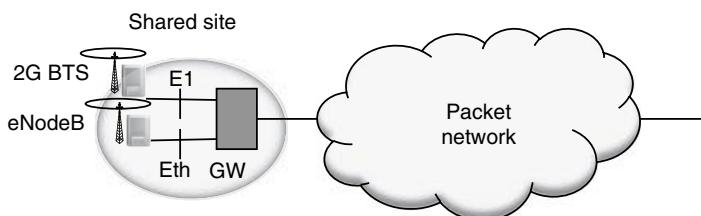


Figure 7.20 Shared transmission, 2G + LTE example.

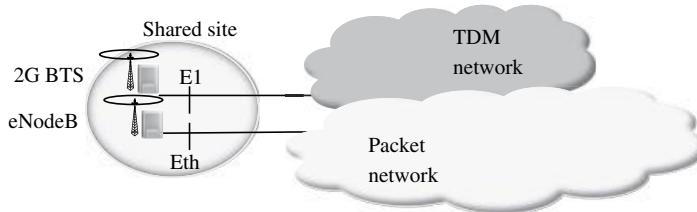


Figure 7.21 Overlay of a packet network for the eNodeB access.

Operating two parallel networks leads to an increased operational expenditure compared to operating a single packet backhaul network. Often the aim is in gradually merging also the 2G into a common packet network. A pre-requisite for this is that the availability and QoS for the 2G service can be guaranteed.

As a variant, Dual Iub presented in Chapter 3, may also direct voice service to the E1s/TDM network, while supporting the high peak rate HSPA traffic via the packet network. In this case all traffic types can be merged to a common backhaul in a suitable timeframe. Until that point, critical voice service can be supported on the traditional ‘telecom-grade’ TDM network.

7.6.2 BTS Access

For the mobile backhaul, the ‘first mile’ access to the BTS is typically not protected. If resilience is needed against a link failure, another link would have to be provisioned. To realize an improved availability the redundant link should not suffer from the same failure condition as the primary link does. Cost is an obstacle, as doubling the amount of access links increases both the access transmission capital expenditure (capex) and operational expenditure (opex) as well.

In many cases there are further issues to be addressed. Even if costs would be acceptable, implementation may not be possible – due to simply not having a feasible way of arranging a redundant link. Even if network infrastructure is available near-by in urban areas, arranging connection to the nearest point-of-presence (POP) is not straightforward. Laying out new cable is time consuming. Microwave radio link deployment may be blocked due to antennas that are not always allowed in city centres, due to aesthetic reasons.

New sites, and site space in general, are a scarce resource, especially in urban areas. It is difficult to fit transmission equipment to the sites. New small footprint BTS generations (without any cabinet) can be fitted into an unconventional environment, and the concept of a ‘site’ is changing. Designing the BTS small in dimensions and lightweight in order to utilize the new site types means also less options for transmission links. There simply is little extra space to be consumed.

7.6.3 IP Addressing

In the previous section difficulties in providing redundant first mile access were brought up. In many cases the base station itself and the first mile access remain single-point-of-failures.

When however redundant physical links are available, IP layer connectivity needs to be arranged between the peer mobile network entities. It is thus useful to review IP addressing. The following sections could be valid for any mobile network elements depending on their

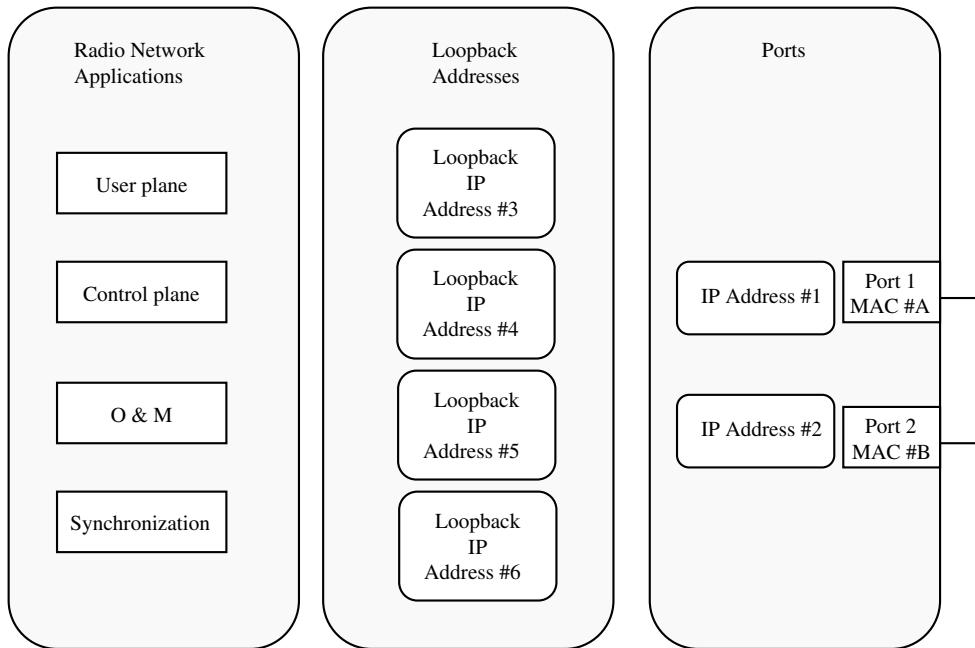


Figure 7.22 Loopback addresses.

implementation. 3GPP requires IP address or addresses that need to be used but does not define any particular addressing architecture. Mobile network elements may also include additional L2 switching or IP routing capabilities between the interface ports. This can be considered as an integration of additional transport functionality that is not covered by 3GPP. It however impacts addressing and also the design of the resilience in the backhaul.

For user plane, recall from Chapter 3 and Chapter 4, that the receiving side informs the peer entity of the IP address that is to be used for the bearer in the bearer set-up phase. For the user plane, this address can basically change dynamically for each new bearer set-up: it is transferred from the BTS to the RNC by NBAP and from the eNodeB to the MME (and from there further to the user plane element, GW) via S1-AP.

For control plane and management plane this is not the case. The IP address is configured; changing the address ‘on-the-fly’ breaks the IP layer connectivity. So the IP address has to remain the same. This leaves two main options: Either transfer the IP address from the failing port to the operating port, or use a loopback address that is reachable via any physical port.

An example in Figure 7.22 shows the interfaces (ports) which are assumed to have an IP address configured. Radio network applications, User plane, Control plane, O&M and Synchronization, are shown in the left hand side. These applications are tied to the virtual (Loopback) addresses. Loopback addresses may also be useful for termination of the IPsec tunnels (see Chapter 9).

In general, how addressing is implemented in the mobile elements: whether there are one or more IP addresses, and whether the addresses are tied to the ports, or should be loopback addresses, is not covered in 3GPP. So it is an implementation issue.

Alternatively, one could also have a single loopback address for all of the applications (user, control, O&M, and synchronization planes). In the case of a single IP address, port numbers and protocol type information is used to direct the IP packet to the correct termination point. As said, there is no mandate to use any loopback addresses in the first place; IP addresses can be tied to physical ports.

Loopback address allows radio network layer applications to continue working based on loopback addresses even if a port fails. If the next hops change, there is no need to involve the radio network layer.

Having the radio network protocols terminating on loopback addresses instead of interface addresses provides additionally a level of independence from the IP address configuration of the transport network. Changes in the transport network IP addressing do not need to be reflected in the loopback addresses. This may be beneficial especially if the transport network is managed by a separate entity.

When the IP addresses used by the radio network layer are simultaneously physical (port) IP addresses, the IP address needs to be transferred in case of protection switching (unit switchover).

7.6.4 Active-Passive Ports

In an active-passive operation one IP port is active, while the other one is not used and does not carry any traffic (See Figure 7.23). If the active port or the link to which the port connects to, fails, another link is used instead.

The cost to destination via Network 1 or Network 2 defines which link is used. Assume that Network 1 is preferred due to a better metric (lower cost). The other link (Network 2) is idle. Network 1 is used all the time unless a failure occurs on Network 1. This could be e.g. a physical link failure such as a cable cut or a failure of a physical port (Port 1 of eNodeB, or Port 9 of router). Failure is detected by a physical layer indication, OSPF Hellos, BFD, etc. While physical layer indication is typically the fastest method, it does not necessarily detect failures that occur on higher protocol layers. These are noticed by missing Hellos or BFD.

After detecting the event, eNodeB calculates a new lowest cost path towards the destination (SGW in the user plane) and starts using Network 2, assuming that Network 2 stays operational. In the other direction, after similar steps, Router starts using Network 2 in order to reach Loopback addresses of eNodeB.

In general routing protocol metrics (like hop count in RIP) define which one of the links is used. This assumes that a routing protocol is implemented to the eNodeB. If the configuration is based on static routes (no routing protocol), a preference parameter or an equivalent way is needed to identify which of the two routes is primarily used. The secondary route can be taken into active forwarding if the primary link fails. In this case as there is no routing protocol to detect the link failure, physical layer indication or a dedicated detection protocol (Ethernet OAM, BFD, etc) is needed.

The case as shown in Figure 7.23 is most relevant when the eNodeB and the router are on separate sites and it is possible to arrange a redundant physical port in the eNodeB and a redundant physical link between the sites. If the router is located on the base station site, the site internal link between the eNodeB and the router is not as error-prone as a WAN link and

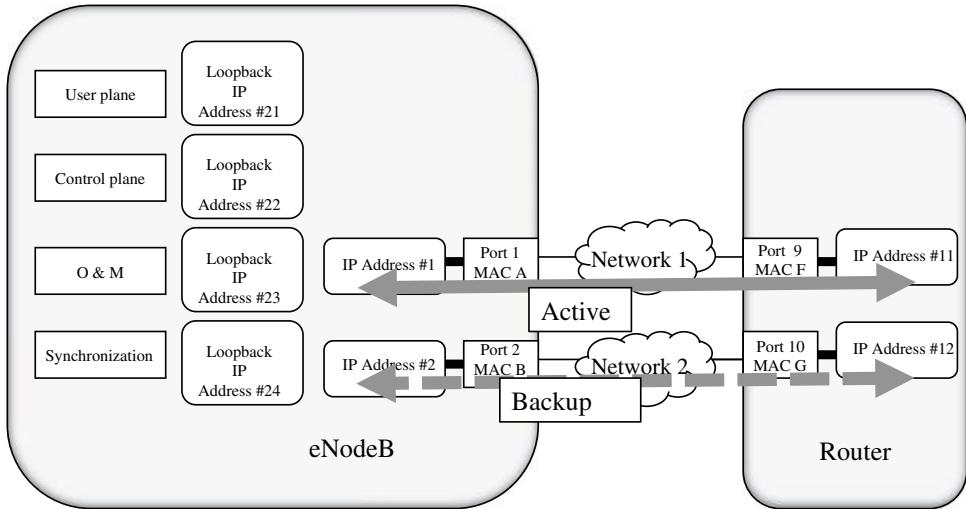


Figure 7.23 Active-passive configuration.

would not typically be made redundant. As well, assuming that the failure rate of a physical port is low the contribution to the eNodeB availability due to a port failure is not significant even with no redundancy.

The situation is different for a controller. Instead of an eNodeB, the principle shown in Figure 7.23 could be valid for a controller element as well. Some differences exist, however. Number of ports in a controller is usually much higher than in a base station and typically more IP addresses are needed. In general controllers need to tolerate failures on ports and units because a number of base stations depend on the operation of a controller (BSC or RNC). A failure of a controller port can be mitigated by some form of unit protection switching (switching the traffic from an active unit to a backup unit) or by load sharing (having multiple active units). In the case of unit protection switching on the controller, the used IP address (IP Address #1 of Port 1) could be transferred to the new active port, Port 2. These are very much platform and implementation specific topics. With a controller, often a site device exists, and resilience is a topic for the site solution design. Virtual Router Redundancy Protocol (VRRP) example is presented in section 7.7.2.

7.6.5 IP Load Sharing

The benefit of load sharing is that traffic load can be distributed to use network and node resources more equally. In the example shown in the previous section, Network 2 was not used at all and all traffic was carried by Network 1.

As a variant, the two ports and the two networks can be active at the same time with IP load sharing. Both ports forward traffic and a load sharing algorithm distributes IP flows to the ports. OSPF and IS-IS both support Equal Cost Multipath (ECMP). ECMP needs to be supported by the devices as shown in Figure 7.24. Additionally, for the ECMP algorithm, some variation in the input information is needed in order to distribute the flows.

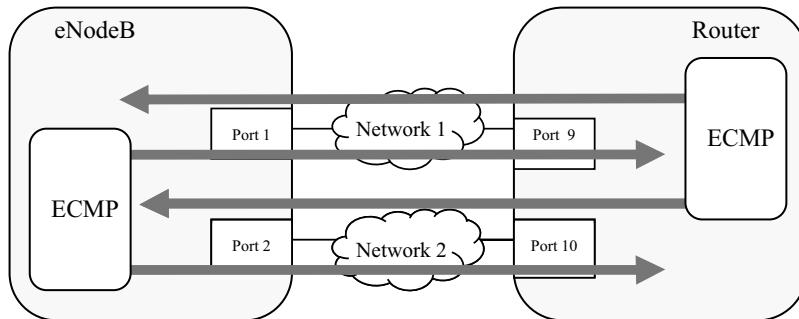


Figure 7.24 Load sharing.

In addition to the possibility to use both links in an active mode, another benefit is the possibility for a faster recovery after a link failure. Assume now that Network 1 fails in Figure 7.24.

Because two active paths exist, forwarding can continue using the remaining active path. (There in general could be more than the two active paths shown in Figure 7.24). Soon as the link failure is detected, load sharing algorithm ceases using the failed path and all traffic is distributed to the single link. Depending on the implementation, this is a faster process than running the Shortest Path First algorithm and installing the new next hop to the forwarding table. And since the other link is already used in forwarding, the MAC address (in case of Ethernet as L2) of the next hop is already known and there is no need for an ARP request.

7.6.6 Ethernet Link Aggregation

Ethernet link aggregation, when implemented to the base station or other mobile network element, also allows the use of multiple ports simultaneously. If any of the links are operational, the aggregate stays operational. Ethernet link aggregation allows a simple way for link/port redundancy, as the function is not visible to the Ethernet MAC layer, nor to the IP layer.

If one of the ports fail (Port 1 or Port2) in Figure 7.25, this does not change the MAC address. The link partner at the other end continues to use the same IP and MAC addresses. In the example two ports were grouped. There might be more ports aggregated.

As with IP based load sharing, with Ethernet link aggregation a distribution algorithm is needed so that load can be balanced over the individual links.

7.6.7 OSPF in the Access

In the mobile backhaul, OSPF can be used for rerouting traffic around failures. This is a common way of providing resilience in networking. OSPF in general supports IP connectivity between nodes and reduces configuration effort. OSPF is also one option as an IP control plane protocol for many MPLS based applications: As a CE-PE protocol for IP MPLS VPNs, as MPLS core IGP, or as OSPF-TE for MPLS-TE applications.

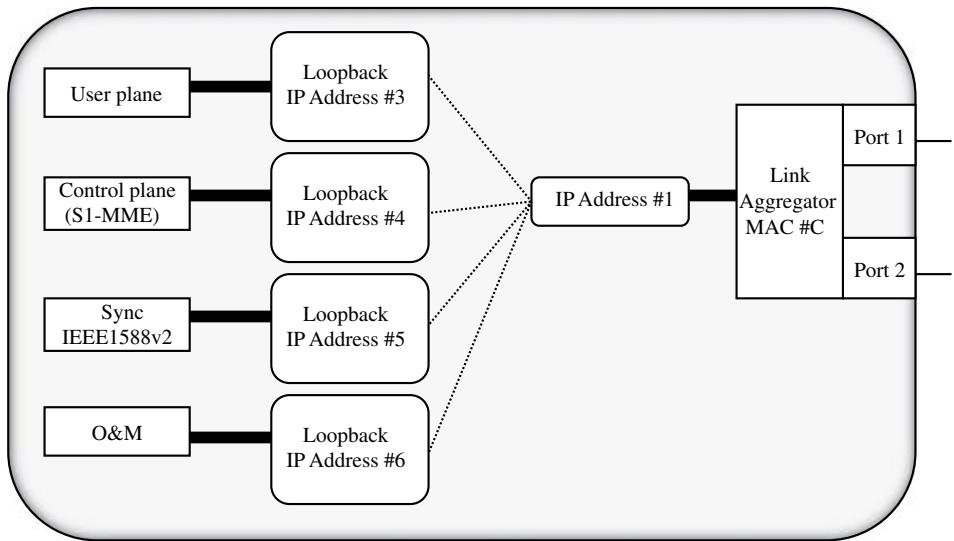


Figure 7.25 Addressing with Ethernet link aggregation.

An LTE access network with eNodeBs is shown in Figure 7.26. eNodeBs run OSPF as well as the routers in the aggregation network. Aggregation router (R1, R40 or R50) acts as a default gateway for the eNodeBs connected to it. OSPF can be used to obtain the default gateway address.

The connectivity to the core network (S1) exists via the aggregation and potentially via the backbone network. For direct X2 traffic, neighbour eNodeBs reach each other via R1, or

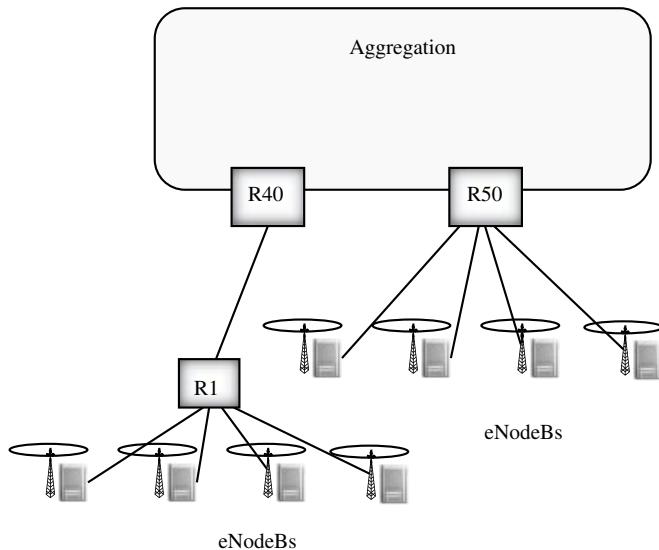


Figure 7.26 eNodeB access example with OSPF routing.

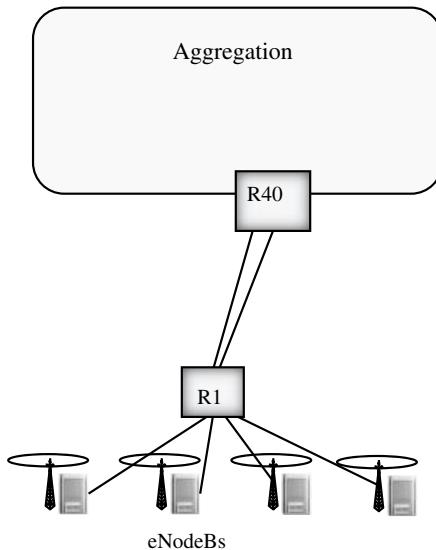


Figure 7.27 Redundant uplinks.

via R1-R40-R50. In the aggregation network, OSPF re-routes around link and node failures, assuming that redundant links are provisioned.

For resilience against a R1-R40 link failure, another link between R1 to R40 can be added. See Figure 7.27.

Assume that the destination network is reached with a lower entire path cost via the link on the left hand side. This link is used all the time. If the link fails, traffic is routed via the other link. The links may be implemented with any underlying technology. An example is a microwave radio access link, or an Ethernet service (E-Line).

Load sharing is another option. In this case costs to the destination need to be equal via both links. Load sharing algorithm distributes traffic, on a flow basis, to both paths.

Instead of the redundant link to the same router R40, one can implement the other link between R1 and R50, as shown in Figure 7.28. This adds resilience against failure of R40 – traffic may use R50 should R40 fail. Similarly, OSPF cost defines the path to be used.

As an alternative, higher availability of the router can be achieved by installing redundant modules into the router node.

The first mile access from each eNodeB to R1 is a single point-of-failure. To improve, a redundant link could be configured between R1 and the eNodeB. Again, OSPF between eNodeB and the R1, advertises both links with the configured metrics.

While this is technically feasible, cost of the redundant access lines is high. If the radio network includes specific BTS sites that need to be continuously available, redundant access lines could be deployed on these few sites only. As generally, here as well the need for protection is assessed based on the availability targets and on the estimated failure rates.

In the case of a direct X2 implementation, the application benefits from a direct spoke-to-spoke connectivity. In other cases in mobile backhaul, such a need (direct base station to base station connectivity) does not exist. A basic characteristic of OSPF is to maintain the full

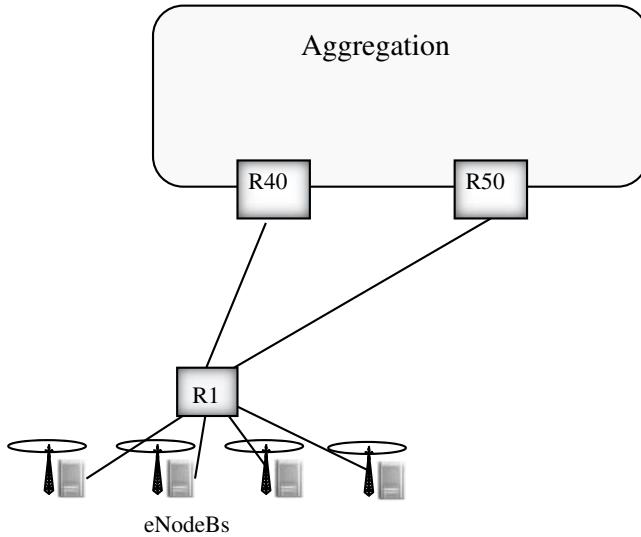


Figure 7.28 Redundant uplinks to redundant uplink nodes.

topology information of an area. With an OSPF stub area, external LSA flooding may be reduced. Stub areas still receive LSAs from the same area.

OSPF areas can be configured in order to limit the amount (and size) of LSAs. A group of eNodeBs are mapped into an OSPF area. The OSPF areas are connected via an OSPF backbone (area 0). Each eNodeB is still able to reach any other eNodeB, if needed. When the neighbour eNodeB belongs to another OSPF area, inter-area routing takes place via the OSPF backbone.

7.6.8 Static Routes

If no routing protocol is in use, static routes are configured on the eNodeB and on the router. With more than one static route, a parameter is needed to define, which of the static routes is primarily used. When the preferred/primary route fails, forwarding table is changed so that instead of the primary route, it uses the outgoing interface/next hop router corresponding to the secondary path. As there is no routing protocol, failure needs to be detected by other means.

In the case where the eNodeB and the Router are directly connected (e.g. via a direct cable connecting the ports), a signal down indication is received from the physical layer. When this is detected, both the eNodeB and the Router, can remove the corresponding entry from the active forwarding table. If a second route to the same destination exists, this can be taken into use.

Physical layer indication may be available e.g. from Sonet/SDH, TDM, or Ethernet layer. However, in case the failure occurs somewhere within the WAN, neither eNodeB or Router necessarily see the physical layer indication. The device in between might be a microwave radio, a Sonet/SDH node or any other intermediate device that masks the signal down indication.

With static routes there is no routing protocol to detect this type of failure, so a dedicated failure detection protocol is needed. At the Ethernet layer, Ethernet OAM supports

connectivity check at a service level (VLANs) as well as link OAM (port-based, one hop only). At the IP layer, BFD (Bidirectional Forwarding Detection) can be used as a hello protocol without a routing protocol.

7.6.9 First Hop Gateway Redundancy

BTS can also be seen as a host without a routing protocol. After all, there is typically no more than a single link towards the aggregation network.

As previously discussed, if multiple static entries are supported, the first hop gateway can be selected among these static entries, based on a preference parameter, and on information of the link availability (detected by physical layer, Ethernet OAM, or BFD, as examples).

A standard way of supporting first hop gateway redundancy for hosts on a LAN is with VRRP [17]. VRRP example is presented for a controller site LAN. Similarly it can be deployed for base stations within a site LAN.

7.6.10 Microwave Access Links

Microwave Radio (MWR) links are the most common BTS access technology currently. Even though fibre is becoming more generally available, the majority of sites do not have access to fibre.

Ring topology (Figure 7.29) has been widely used with microwave radio link transmission, since 2G.

Many variants exist for the support of ring protection in the MWR ring. For Ethernet based ring, ITU-T has standardized G.8032 [63]. There are also vendor specific implementations.

Capacity needed by the base station sites versus the capacity in the ring is a consideration. Long chains and/or large rings that have been used with TDM based MWR networks for 2G, have limitations for 3G HSPA or LTE. Especially for high capacity eNodeBs of LTE, capacity in the ring easily becomes insufficient. Even with a MWR link capacity of hundreds of Mbit/s, room for growth is limited. Upgrading capacity in the ring often involves a change in the topology, which is effort and time consuming as new links need to be provisioned. Capacity upgrades in the ring are difficult to carry out if all the microwave spectrum is already in use.

The availability of a single microwave radio hop is limited by the duration of exceptional rains (couple of minutes to tens of minutes per year, depending on the geographic location). During this time it is likely that all links from a site at least partially suffer from the same

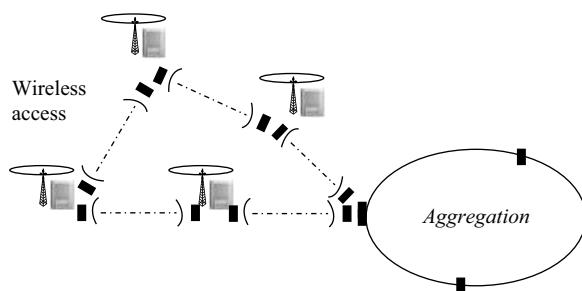


Figure 7.29 Microwave link based ring.

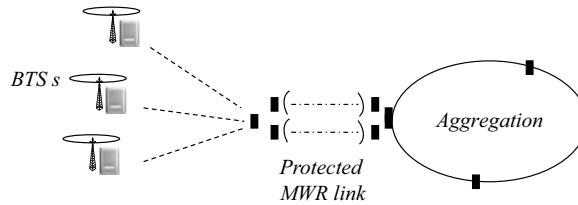


Figure 7.30 Protected MWR link.

condition. So full resilience against link failures is not achieved. Equipment failures on a remote site, as well as link failures on remote sites, can be mitigated.

As an alternative to the ring topology, MWR can be used to access the eNodeB with a few hops only before attaching to the aggregation network. With few hops to a resilient aggregation network, resulting availability may be acceptable without any protection on the access links. This of course depends on the availability target.

It is also possible to deploy a protected point-to-point link, making a single link redundant. Ethernet APS protection G.8031 is one option [64], Ethernet link aggregation also solves the case shown in Figure 7.30.

Again, if the impairments in the radio propagation are the main source of unavailability, having two point-to-point microwave radio links parallel to each other, does not add much resilience. In this sense, an improvement is achieved by connecting the MWR hub site to two different aggregation devices. See Figure 7.31.

Now the two MWR links do not fully share the same radio propagation conditions. If the wireless transmission is impaired due to local conditions in the vicinity of Edge Node A, MWR link 2 is not necessarily impacted. The configuration is also resilient against failures of the aggregation Edge Node, since either Node A or Node B can forward traffic.

Assuming the MWR hub is an IP device, and an IP aggregation network, redundant links and nodes can be used, with a routing protocol operating between the MWR Hub and the Edge Nodes. Routing protocol metrics define, which of the links is used, or whether both are used in a load sharing configuration.

7.6.11 Attachment to a MEF Service

In the basic case, Ethernet service availability is an agreement between the service provider and the customer.

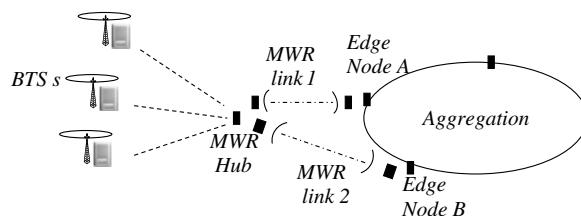


Figure 7.31 Multihoming a MWR hub node to two aggregation devices.

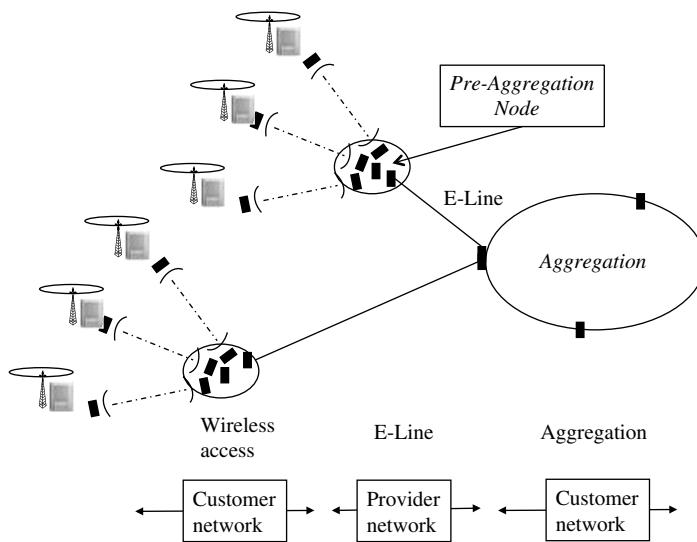


Figure 7.32 Pre-aggregation serviced by E-Line.

Additionally resilience may be needed against failures of the attachment circuit, and against provider edge nodes. Consider a pre-aggregation node (part of customer network), which is connected further via a MEF service (by a third party service provider) to a customer aggregation network, as in Figure 7.32.

At the pre-aggregation site, traffic from a few BTSs is collected into a single physical Ethernet port. An Ethernet service (E-Line, E-LAN, or E-Tree) provides the transport for the BTSs, mapped into VLANs. Availability of the MEF service is negotiated as previously.

Further failure cases are the failure of the customer network's hub node, the attachment circuit to the service, and the failure of the provider edge node. See Figure 7.33. Similarly

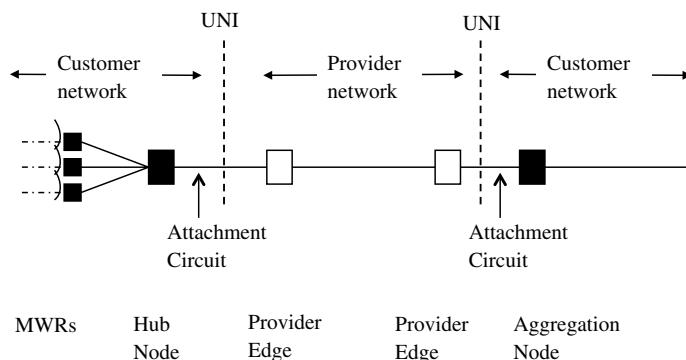


Figure 7.33 Connecting hub nodes via an Ethernet service to customer aggregation network.

the attachment circuits and nodes at the UNI at the other end of the service have to be considered.

The attachment circuit (consisting e.g. of VLAN tagged Ethernet frames) at the Hub node end is physically site internal cabling. The cable may still be disconnected e.g. during site visits. The physical interface ports on both sides of the UNI, at the hub node, and at the provider edge node, are single point-of-failures.

UNI 2.0 implementation agreement (MEF20) requires support of link aggregation for UNI type 2.2 [65]. Link aggregation can protect against a failure of the link in the attachment circuit, since the attachment circuit now consists of multiple ports and multiple links as shown in Figure 7.34. It also protects against a failure of an individual port on either side of the UNI. Link aggregation thus provides protection against a failure of the attachment circuit and against failures of ports connected to the attachment circuit. UNI type 2.2 can be used for both of the UNIs: At the hub node UNI (left hand side of the UNI), and at the aggregation node UNI (right hand side of the figure).

What is left then unprotected after the above are the nodes themselves: Hub node and aggregation node at the customer sites, and the provider edge nodes at the provider network. The provider edge node protection is more critical, as it typically would connect to several Hub nodes.

MEF recognizes the need to protect the attachment circuit and the CE connecting to the service, however MEF2 does not address dual attachment of the customer equipment (CE) to the service. This is left to the service implementation. Candidate solutions of VPLS multi-homing were discussed in section 7.5.4. In the case of Figure 7.34, it may be that the availability is acceptable without redundancy, of course again depending on the failure rates and on the availability targets. Additionally, network nodes often support making critical components redundant, thereby increasing resiliency against node failures.

Alternatively, two EPLANs could be considered over the UNI, and use routing in the customer network. This increases availability mainly if separate nodes would be used at both sides of the UNI.

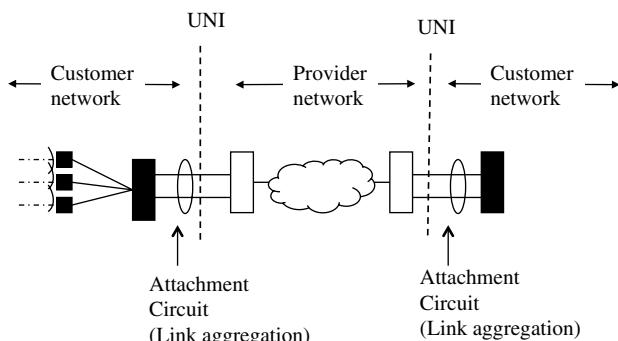


Figure 7.34 Link aggregation over the UNI.

7.7 Resilience in the Controllers and the Core Interface

Controllers and gateways are often connected to the aggregation tier. Resilience in the aggregation network is provided e.g. by the functionality of MPLS and IP as discussed in sections 7.4 and 7.5.

Lower layers also (Sonet/SDH and fibre/wavelengths) may support resilience. These were addressed in Chapter 5.

7.7.1 BSC and RNC and Their Site Solutions

2G and 3G/HSPA BTSs are attached to a single controller element (BSC or RNC) in the radio access network. Controllers present a single point of failure in the system architecture. So the controller elements need to be highly reliable. BTSs can be rehomed to another controller, but this is accomplished by configuration in the management plane, and is not intended for protection but rather for radio network expansions. To assure high availability, controllers are built on fault tolerant platforms.

A controller site solution needs to be as highly resilient as the controller itself. Otherwise the failure of a site device will lead to an event similar to that of the failure of the controller. Security gateways and higher tier aggregation nodes also serve a high amount of BTSs. These nodes, as well as the links connecting to the elements need to have a high availability.

The architecture of the mobile network elements is not covered in 3GPP. What kind of resilience, unit redundancy, protection switching, load sharing and routing these elements support is an implementation issue. The detailed element architecture, as already mentioned, however directly impacts the solutions related to network resilience.

In the telecom world, 1:1 type of protection is one alternative, matching well with Sonet/SDH protection architectures. If the processing units of a controller support switchover from one unit to another, MAC address of the termination point may change.

7.7.2 VRRP Example

For a site device, redundancy may be supported with VRRP as shown in Figure 7.35.

If one of the site switches fail, the other one takes over. The host, RNC in the figure, can continue using the previous IP and MAC addresses in its forwarding. If the MAC address in the RNC changes due to a unit switchover, the switch needs to learn the new MAC address.

Another type of resilience at the site is achieved with load sharing. Several units in the mobile network element are active and share the processing of user plane bearers, as an example. Load sharing is supported at the IP layer by equal cost multipath, or at the Ethernet layer by the use of Ethernet link aggregation. Ethernet link aggregation is not visible to the IP layer. With load sharing, an algorithm is needed to balance traffic over multiple units.

7.7.3 Signaling Resilience with SCTP Multihoming

Connectivity failures are critical for the signaling links. User plane bearer set-ups depend on signaling messages over NBAP/SCTP and RANAP/SCTP for 3G, and S1-AP/SCTP and X2-AP/SCTP for LTE.

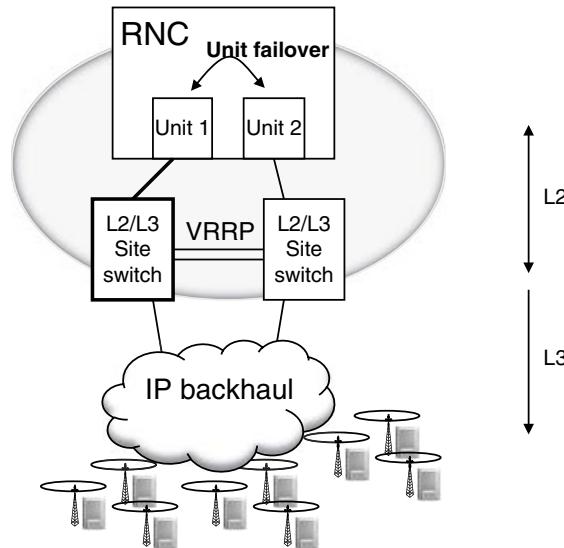


Figure 7.35 Unit protection and site solution.

Additionally, in 3G, RNC is responsible for managing the resources of the radio cells and other BTS resources via the NBAP/SCTP. A prolonged break (depending on the implementation it may be tens of seconds to minutes) on NBAP connectivity causes BTS resets and shutdowns/restarts.

SCTP has a built-in multi-homing support. This helps to keep the SCTP level connectivity up even if a single interface fails [66], [67]. If the IP interfaces are accessible only via the same physical link, there is no resilience against a failure of the physical link.

On the radio network – core network interface, multiple links are typically available. An example SCTP application on the Iu interface with multi-homing is presented in Figure 7.36.

Multihoming capability means an endpoint can have multiple IP addresses. This increases resilience against path or interface failures. A multi-homed endpoint has a single SCTP port, but multiple IP addresses. An example is shown in Figure 7.36 with the radio layer control plane signalling (RANAP) as the SCTP user.

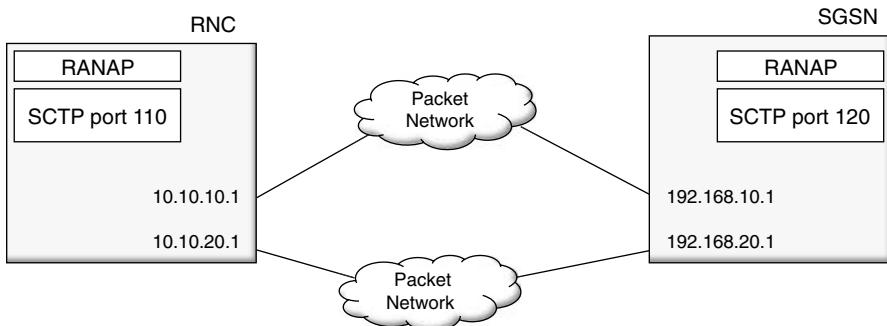


Figure 7.36 SCTP Multihoming example.

Both RNC and SGSN are configured for multihoming with two IP addresses on both ends. One of the transport addresses is selected as a primary transport address. The sender will transmit packets to the primary transport address of the receiver.

If a peer is multihomed, the SCTP association is kept active, if any of the destination address(es) of the peer is available.

Functionality of SCTP includes a detection of unreachable address and a detection of an unreachable endpoint (all destination addresses of the endpoint are not reachable). SCTP sender expects acknowledgements from the receiver. If there is no data to be sent, HEARTBEAT chunks are used to monitor peer reachability.

When the SCTP sender does not receive a response in a defined time, it considers that the destination address is potentially unreachable. Sender keeps track of missed responses, and if a defined number of consecutive responses are not received, the destination address is considered unreachable, and marked inactive. The protocol parameter for the amount of missed responses is Path.Max.Retrans.

7.7.4 Use of Multiple Core Network Nodes

Core network must be highly available. While the implementation of the elements is not covered in 3GPP, for core network, 3GPP has defined functionality that enables the radio network to connect to multiple core network elements. This is for 3G defined in TS23.236, ‘Intra Domain Connection of RAN Nodes to Multiple CN Nodes’ [68]. Sometimes terms like ‘Flexible Iu’ or ‘Multipoint Iu’ are used, which also describe the functionality well.

The idea is to remove the restriction of having only a single core element connecting to the radio access network. Instead, core network elements form a pool, and any element within the pool may be used to serve the user. Functionality is defined for 2G (A and Gb), 3G (Iu-cs and Iu-ps) and LTE (S1-U and S1-MME). In addition in supporting resilience, load sharing is achieved as well as an optimized signaling.

A concept of pool area is defined, and now a RAN node may be served by any of the core network elements (MSC pool for the CS domain, and SGSN pool for the PS domain) configured for that specific pool area. Pool areas may also overlap, as is shown for the RAN nodes in Area 2 and 6 in Figure 7.37. Separate pool areas exist for the CS and PS domains.

Network Resource Identifiers (NRIs) identify the CN intended, from the CN element pools. The RAN node (RNC in the case of 3G) provides a NAS (Non-Access Stratum) node selection function, which selects the specific MSC or SGSN to which the initial NAS signaling is routed. This selection function can also balance the load between the available CN nodes.

As such, the Flexible Iu operates at the radio network layers, and provides resiliency against failure of a core network node, since service is available via the other nodes in the pool. If the core network node (or all links to it) fails, signaling is routed to other CN nodes in the pool. Ongoing calls/sessions are lost when the serving CN fails, but new bearer set-ups may be directed to the operating core network node.

The mechanism is transparent to the mobile transport layers. Naturally links to multiple core network elements need to be configured. The underlying (Gb/A/Iu/S1) links at transport level may additionally be configured to support transport level redundancy, e.g. by using IP routing in the user plane, and SCTP multihoming in the control plane.

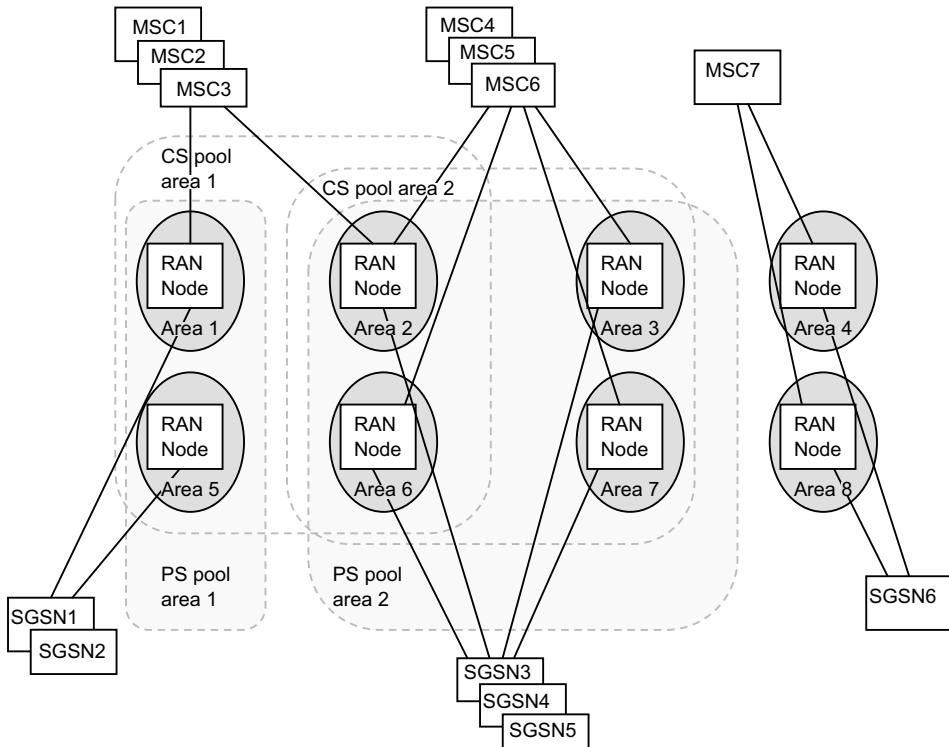


Figure 7.37 Connection to multiple core network nodes [68].

Basic operation on LTE S1 is comparable to the Iu, in that the core network node is selected as per each bearer. If one core network node becomes unavailable, another node is capable of supporting the new bearers.

For mobile backhaul a difference exists in LTE due to the flat E-UTRAN topology. Each eNodeB is logically connected to multiple core network nodes. Physical transport topology is likely to be different, since there typically is no alternate path directly from the eNodeB to each core network element. Instead, a separate path could exist from an aggregation node upwards – located somewhere where RNC would be in the case of 3G.

7.8 Summary

In the mobile backhaul, resilience to link and node failures is achieved by features that depend on the technology used. In the base station first mile access, redundant physical links are seldomly available, due to cost and difficulties in provisioning these links to the compact base station sites.

Microwave radio is the most commonly used access media for base stations. Microwave rings have traditionally provided resilience against path failures. With high capacity base station sites, capacity in the ring easily becomes an issue. In this case, the aggregation tier supports path and node resilience and individual base stations access the aggregation network with a single or few hops only.

With packet networking, IP routing can mitigate network failures by routing around a failed link or node. OSPF is a widely deployed link-state routing protocol. It is well suited for a fast recovery, as changes in link states trigger link state advertisements (LSAs) which cause a recalculation of the shortest path tree to the destination networks. This allows recovery from a link or node failure. Recovery time depends on the size of the network (in terms of links and nodes) and on the failure detection. With fast detection, subsecond recovery can be achieved. In addition to resilience that IP routing provides, MPLS supports a fast recovery by MPLS Fast re-route (FRR).

With MEF services, availability is an attribute of the service, defined in SLA and SLS (Service Level Specification). For the MEF service user, availability is increased by renegotiating the contract. The service itself may rely e.g. on MPLS and IP, or on Sonet/SDH as examples. Ethernet link aggregation can add resilience to the UNI.

Mobile network controller nodes, RNC and BSC, introduce a single point of failure. Several base stations depend on the operation of a controller. High availability of the controller site is a must, taking into account the controller node itself and any site devices (LAN switches and IP routers). In LTE, a single node exists in the E-UTRAN, which connects directly to the core network. At a system architecture level this removes one single-point-of failure from the LTE mobile system. Aggregation nodes and security gateways of LTE support a high amount of base stations similarly to controllers in 2G and 3G and high availability of these nodes is required.

References

- [1] Milbrandt, Martin, Menth, Hoehn, Risk Assessment of End-to-End Disconnection in IP Networks due to Network Failures. IPOM 2006, Springer-Verlag Berlin Heidelberg, 2006.
- [2] Vasseur, Pickavet, Demeester: Network Recovery. Elsevier, 2004.
- [3] IETF RFC 3549 Framework for Multi-Protocol Label Switching (MPLS)-based Recovery
- [4] Ross: Introduction to probability models. Academic Press, 2003.
- [5] ITU-T G.911 Parameters and calculation methodologies for reliability and availability of fibre optic systems, 04/97.
- [6] Markopoulou, Iannaccone, Bhattacharyya, Chuah, Ganjali, Diot: Characterization of Failures in an Operational IP Backbone Network. IEEE/ACM Transactions on Networking, Vol. 16, No. 4, 2008.
- [7] Kuusela, Norros: On/Off process modeling of IP Network Failures, IEEE/IFIP International Conference on Dependable Systems & Networks, 2010.
- [8] IEEE802.1D-2004.
- [9] Froom, Sivasubramanian, Frahim: Building Multilayer Switched Networks. Cisco Press, 2007.
- [10] IEEE 802.1w.
- [11] IEEE 802.1s.
- [12] AnOverviewoftheMEF.ppt, www.metroethernetforum.org, retrieved October 2011.
- [13] MEF 2, Requirements and Framework for Ethernet Service Protection,
- [14] MEF 22, Mobile Backhaul Implementation Agreement (2/09).
- [15] IEEE 802.1ah-2008.
- [16] IEEE 802.1ag-2007.
- [17] IETF RFC 5798 Virtual Router Redundancy Protocol (VRRP) Version 3 for IPv4 and IPv6.
- [18] IETF RFC 2991 Multipath Issues in Unicast and Multicast Next-Hop Selection.
- [19] IETF RFC 2992 Analysis of an Equal-Cost Multi-Path Algorithm.
- [20] Huitema: Routing in the Internet. 2nd Edition. Prentice-Hall, 1999.
- [21] IETF RFC 2328, STD 54, OSPF Version 2.
- [22] IETF RFC 5340 OSPF for IPv6.
- [23] IETF RFC 5838 Support of Address Families in OSPFv3.
- [24] IETF RFC 5709 OSPFv2 HMAC-SHA Cryptographic Authentication.

- [25] IETF RFC 5880 Bidirectional Forwarding Detection.
- [26] IETF RFC 5881 BFD for IPv4 and IPv6 (Single Hop).
- [27] IETF RFC 5883 BFD for Multihop Paths.
- [28] IETF RFC 2439 BGP Route Flap Damping.
- [29] IETF RFC 4136 OSPF Refresh and Flooding Reduction in Stable Topologies
- [30] IETF RFC 1793 Extending OSPF to Support Demand Circuits
- [31] Siddiqi, Nandy: Improving network convergence time and network stability of an OSPF-routed IP network. Networking 2005, IFIP International Federation for Information Processing, 2005.
- [32] IETF RFC 4656 A One-way Active Measurement Protocol (OWAMP).
- [33] IETF RFC 5357 A Two-Way Active Measurement Protocol (TWAMP).
- [34] Benlarbi: Estimating SLAs availability/Reliability in Multi-services IP network. ISAS 2006, LNCS 4328. Springer-Verlag, 2006.
- [35] Kilpi: IP-Availability and SLA. International Workshop on Traffic Management and Traffic Engineering for the Future Internet. Available at www.iplu.vtt.fi, retrieved October 2011.
- [36] IETF RFC 5286 Basic Specification for IP Fast Reroute: Loop-Free Alternates.
- [37] Teare, Paquet: Building scalable Cisco Internetworks. Third edition. Cisco Press, 2007.
- [38] De Ghein: MPLS Fundamentals. Cisco Press, 2006.
- [39] Minei, Lucek: MPLS Enabled Applications. Second Edition. Wiley, 2008.
- [40] Guichard, Le Faucheur, Vassuer: Definitive MPLS Network Designs. Cisco Press, 2005.
- [41] IETF RFC 5036, LDP Specification
- [42] IETF RFC 3037 LDP applicability
- [43] Halpern, Bhatia: Advertising Equal Cost Multipath routes in BGP. IETF draft. [draft-bhatia-ecmp-routes-in-bgp-02.txt](https://datatracker.ietf.org/doc/draft-bhatia-ecmp-routes-in-bgp-02.txt).
- [44] IETF RFC 4576 Using a Link State Advertisement (LSA) Options Bit to Prevent Looping in BGP/MPLS IP Virtual Private Networks
- [45] IETF RFC 4577 OSPF as the Provider/Customer Edge Protocol for BGP/MPLS IP Virtual Private Networks
- [46] IETF RFC 4364 BGP/MPLS IP Virtual Private Networks (VPNs)
- [47] IETF RFC 4456 BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP).
- [48] Bocci, Cowburn, Guillet: Network High Availability for Ethernet Services Using IP/MPLS Networks. IEEE Communications Magazine, March 2008.
- [49] Kothari, Kompella, Henderickx, Balus, Uttaro. BGP based Multi-homing in Virtual Private LAN Service. IETF draft '[draft-ietf-l2vpn-vpls-multihoming-03.txt](https://datatracker.ietf.org/doc/draft-ietf-l2vpn-vpls-multihoming-03.txt)'.
- [50] IETF RFC 4090 Fast Reroute Extensions to RSVP-TE for LSP Tunnels.
- [51] IETF RFC 4379 Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures.
- [52] IETF RFC 5884 Bidirectional Forwarding Detection (BFD) for MPLS Label Switched Paths (LSPs).
- [53] IETF RFC 5885 Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV).
- [54] IETF RFC 6372 MPLS Transport Profile (MPLS-TP) Survivability Framework
- [55] IETF RFC 5586 MPLS Generic Associated Channel
- [56] IETF RFC 5085 Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires.
- [57] IETF RFC 3471 Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description
- [58] IETF RFC 3473 Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions
- [59] IETF RFC 3945 Generalized Multi-Protocol Label Switching (GMPLS) Architecture
- [60] IETF RFC 4202 Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS).
- [61] IETF RFC 4203 OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS).
- [62] IETF RFC 5307 IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS).
- [63] ITU-T G.8032 Ethernet Ring Protection Switching (03/2010)
- [64] ITU-T G.8031 Ethernet Linear Protection Switching (06/2011)
- [65] MEF MEF20 UNI Type 2 Implementation Agreement.
- [66] IETF RFC 4960 Stream Control Transmission Protocol.
- [67] Stewart, Xie: Stream Control Transmission Protocol (SCTP), A Reference Guide. Addison-Wesley, 2002.
- [68] 3GPP TS23.236, Intra Domain Connection of RAN Nodes to Multiple CN Nodes, v10.2.1.

8

QoS

Thomas Deiß, Jouko Kapanen, Esa Metsälä and Csaba Vulkán

In this chapter, we concentrate on the QoS of the transport service. A starting point is in looking at the end user layer, the radio network layer, and the transport layer in Section 8.1.

End user services, based on TCP/IP and UDP/IP, are discussed in Section 8.2. These protocols are used to carry the end user applications over the mobile network. One well known aspect is that TCP reacts to packet loss as congestion. In the mobile network, packets may be lost due to the air interface. On the other hand, lost packets are retransmitted by the radio network layer – potentially recovering the packet before end user TCP reacts.

Section 8.3 discusses backhaul QoS marking. Traffic is classified and marked preferably at the sources. The traffic sources in the backhaul are the mobile network elements: base stations, controllers, gateways. QoS is marked to the Differentiated Services Code Point (DSCP). Further marking to MPLS and Ethernet (or yet other layers) is carried out based on the DCSPs.

Ingress and egress functions are the scope for Section 8.4. These are deployed at the IP layer, and also potentially at the lower layers. When the backhaul is provided as a service by a service provider, ingress policing by the service provider is an example of a topic that needs to be addressed at the user to network (UNI) interface.

Sections 8.5, 8.6 and 8.7 provide a radio network technology (2G, 3G, LTE) specific discussion, with an example backhaul QoS mapping for each of the radio technologies.

Finally, Section 8.8 presents a summary.

8.1 End User Service, Radio Network Layers and the Transport Layer Service

Recall Figure 4.1 from Chapter 4 stating that transport layer provides a service for the radio network layer. The topmost layer is the end user service that receives services from the radio

network and transport network layers. Each involved network layer and protocol has an impact on the quality of service experienced by the end user.

8.1.1 Transport Layer Service

The end user service (voice, internet access, etc.) is provided by the mobile network as a whole (recall Figure 4.1 from Chapter 4). Transport layer provides a transport service (end to end connectivity according to QoS, resilience, security and other requirements) to the radio network layer. Transport network is a shared resource that is used by the user plane, radio network plane, transport and radio network control plane, synchronization and management plane, each having specific service requirements. See Figure 8.1.

As for QoS of the mobile backhaul, in this chapter we are interested in providing the transport layer service with the required QoS.

8.1.2 End-to-End QoS

End users generally are interested only in the issues that are visible to them. For QoS of the service delivered by the mobile system, it is the end user perception of the received quality that matters. Service itself may be a voice conversation, a short message, mobile broadband access, or any service built on top of the TCP/IP protocol suite.

The end user perception is a subjective measure that is difficult to quantify as it depends on many factors. As an example, expectations can be lower in some cases, if the service is aggressively priced. However transport level quality of service is characterized by exact parameters: delay, loss, throughput, latency, etc.

Quality of service is defined in TS22.105 as the ‘collective effect of service performances which determine the degree of satisfaction of a user of a service.’ Factors affecting performance are service accessibility, service retention, service integrity, and yet other factors that are service-specific.

QoS is an end-to-end topic, encompassing all elements in the service delivery chain. QoS has to be supported by each element in the chain that delivers the service for the user. It also includes the external network (e.g. PSTN, or the internet) if involved, and the QoS of that network should be accounted for.

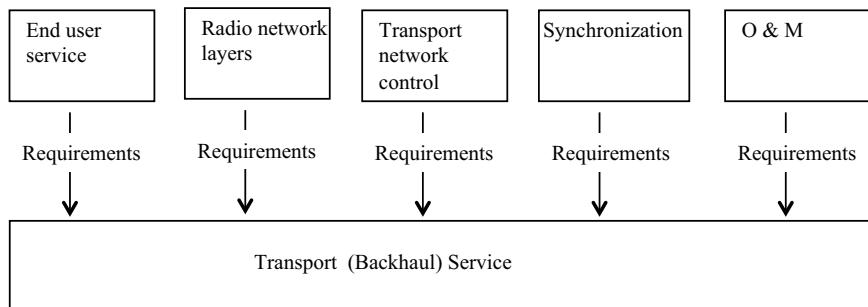


Figure 8.1 Sources of requirements for the transport service.

8.1.3 Need for Backhaul QoS

Do we need QoS in the mobile backhaul – and if yes, why?

QoS is needed – the service for the user needs to meet the quality level that is essential for the service itself. As an example, a voice call needs to have sufficiently low delay. Excessively delayed voice frames are discarded by the decoder at the receiving end. QoS for voice in the transport layer may be implemented with different QoS functionalities.

What might not be needed is the architecture providing the end-to-end and node level QoS mechanisms in case the network has no shortage in resources. Simply stated, if it can be assured that all traffic types are served according to their specific requirements at all times, there is less need for sophisticated QoS functionalities.

TDM networks are deterministic providing a fixed rate service. However, they lack flexibility and capability for statistical multiplexing and also low cost high capacity ports. High mobile broadband peak rates increase the peak-to-average ratio in the network. In the aggregation tier, this ratio can be exploited with packet networks due to time diversity of the aggregated traffic. The benefit is in less bandwidth needed on the transmission links, and less physical ports in all of the nodes.

Abundant bandwidth has traditionally not been available in the BTS first mile access. Instead, due to a high amount of BTSs, and a high amount of links needed, first mile access has been – and often continues to be – a bottleneck.

Deployment of fibre optics and DWDM/CWDM makes high bandwidth gradually available at least for an amount of BTS sites. Still the majority of today's sites are not connected to fibre. Locations of BTSs can typically not be selected according to the availability of fibre. In order to provide the required level of service, installation of new fibre is often required.

Even if enough bandwidth at first appears to be available, it may become inadequate due to the exponential growth of mobile traffic as it has already been experienced with HSPA. Mobile backhaul should have enough spare resources in order to accommodate the potentially increased future traffic; therefore at least new investments should be planned accordingly.

This is due to a long lead time for the new access lines – in most cases, it is not straightforward to increase capacity to the existing lines. In microwave radio access, new frequency bands, related licenses and potentially HW must be acquired at a network extension.

With a leased service, negotiating a higher bandwidth may as well take time, if a higher capacity service is not yet available. Expansion of capacity in the access tier often also causes a change to the network topology. This then impacts a larger area and causes at least a partial backhaul redesign. For reasons like these, bandwidth of the access lines often remains a scarce resource, at least until the upgrade phase has been completed.

Services in general differ with respect to their tolerance to latency, jitter and loss. Backhaul QoS features in the network nodes can be configured so that network resources (links and nodes) are used efficiently. At the same time, service level to the end user becomes more predictable and easier to manage. For these reasons, a level of QoS is typically deployed in the mobile backhaul.

Capacities are explored in Figure 8.2. The first mile access should be dimensioned so that peak radio rates are not limited by it. If the maximum air interface peak rate needs to be achievable in the system (even for a single user), this peak rate (as a minimum) needs to be supported on the first mile access as well. So, e.g. 14.4 Mbit/s HSPA peak rate requires 14.4 Mbit/s + overheads in the first mile access link. This allows in optimal radio conditions

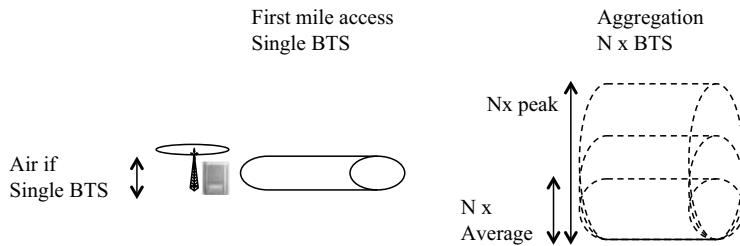


Figure 8.2 Capacities in the first mile and in the aggregation tier.

and with no other traffic load, a single user to reach the 14.4 Mbit/s rate. Average rate from the same site may be considerably less. In the end, dimensioning is guided by business decisions: target level of traffic and service.

When multiple BTSs are aggregated, it is realistic to assume that not every BTS transmits at its maximum rate ($N \times$ peak of a single BTS) simultaneously. This is due to user behaviour, mobility, daily routines, etc. The higher the amount of BTSs whose traffic is aggregated together, the lower the probability that $N \times$ peak rate is needed. Traffic amounts from the individual sites vary: downtown and suburban BTSs are not experiencing busy hour load at the same time. However, for two neighbouring BTSs at a downtown area this might not be true.

With a high peak-to-average ratio, the difference between $N \times$ average and $N \times$ peak becomes large. Correspondingly, the aggregation layer can produce a high gain, depending on the assumptions discussed.

Figure 8.2 also points out two candidate capacity bottlenecks: the first mile access and the aggregation. The first mile access may be of low capacity (couple of megabits/s) if it is supplied by traditional TDM lines. In these sites, QoS functionality is a must, at least if HSPA or LTE based data services are rolled out. The second area to study is the aggregation.

Instead of a single aggregation tier, BTS sites are often chained. Also small pre-aggregation sites may be created, to connect sites to the next level aggregation node. Chained and pre-aggregated sites belong to the access tier, and are then attached to a typically fibre based aggregation network. Any of these tiers may introduce congestion.

In addition to a self-deployed backhaul, such as the microwave radio access, part of the network may be implemented as a service from a third party provider. An example is the use of Ethernet service (E-Line shown in Figure 8.3). The capacity available depends on the CIR and EIR and other attributes defined for the service at the UNI. The service may as well introduce congestion that eventually leads to delay and loss.

The first and most important step in QoS provisioning is to serve the high priority traffic (control, voice) with the right level of QoS. This is not problematic as today data applications are dominating the user traffic mix, being of a lower priority than voice and control/signaling. With a suitable differentiation, e.g., a strict priority service, the intended level of QoS can be guaranteed without much of an added complexity. This of course requires that the transport network has at least enough bandwidth to carry the high priority traffic. Multiplexing gains can be considered for the high priority traffic, too.

At a minimum, two priority levels are needed for this: One for the critical/real-time traffic with a high priority/guaranteed throughput, and another one for anything else, with a best effort type of service. Guaranteeing control traffic and voice sustains availability of the basic service and access to the network even during congestion.

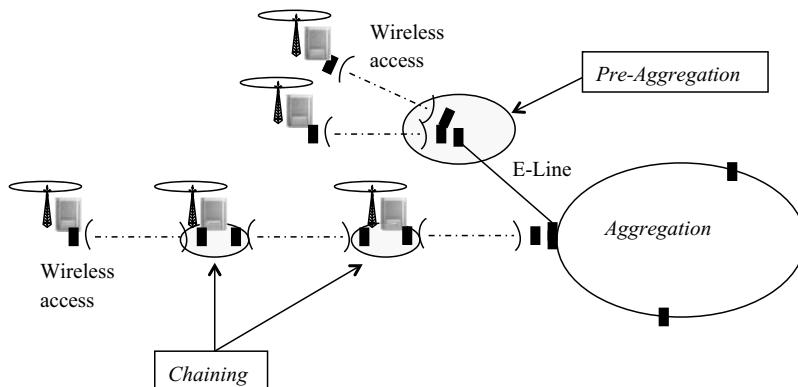


Figure 8.3 Aggregation.

In packet based networks, interface schedulers multiplex traffic forwarded to the same direction to the same physical links. If the total rate of the aggregated traffic exceeds the egress rate of the scheduler, packet queues will build up eventually leading to packet drops if the available buffering capacity is exceeded. Assuming critical/real-time traffic is guaranteed, further classes allow network nodes to prioritize elastic traffic according to a predefined QoS policy. In the case of congestion, real-time traffic and control channels are serviced with an uncompromised quality.

Elastic traffic is prioritized according to traffic type or subscription (user differentiation). With increasing granularity of differentiation, complexity increases as well. Sophisticated QoS features may not be available in all of the smaller platforms used in the pre-aggregation.

When voice and control traffic is prioritized, background/interactive connections will easily experience congestion. QoS functions can then improve the efficiency and fairness of the system. Without an investment into backhaul network capacity, QoS for the background users may remain poor. Having adequate capacity to start with is one of the most important QoS features. An overloaded network will not perform well, irrespective of the QoS features supported.

8.1.4 QoS Alignment with Radio and Backhaul

Resource shortages may occur on both the air interface and on the mobile backhaul. QoS features, such as scheduling, are however separate functions for the radio and transport resources. Air interface scheduling is a function of the RNC/BSC or BTS (HSPA NodeB and LTE eNodeB), and located in the radio network part of the implementation. Transport scheduling occurs e.g. at the network interface card of the BTS and other mobile network elements, and also in the network nodes of the mobile backhaul.

The radio interface requires special scheduling mechanisms that are able to achieve high resource utilization and are able to provide the level of service that each of the connections needs. This requires that at each scheduling decision the channel quality of each active connection is considered. Therefore the air interface scheduling is done on a connection/flow level. In contrast, transport schedulers are providing QoS on an aggregate level. Traffic is aggregated into a relatively small amount of QoS classes at the ingress of the transport

network. Therefore QoS is not guaranteed on a flow level at the transport layer. In order to enforce the end-to-end QoS within a mobile network, radio and transport QoS mechanisms should be harmonized.

Providing such harmonized QoS mechanisms is not trivial, taking into account a continuously changing traffic mix.

Alignment means that QoS mechanisms such as transport and radio schedulers implement consistent treatment of the traffic classes, in order to achieve the intended level of end-to-end QoS. A starting point is that the QoS mapping is consistent: packet flows can be identified by both radio and transport. They also need to be interpreted similarly. The actual algorithms may well be different, but the behaviour on both layers should support the same goal (e.g. a guaranteed throughput, low loss, etc).

Scheduler was used as a practical example. Admission control is another item; clearly if an admission control is to be applied for a bearer, it should be applied for both radio and transport resources, assuming that either one may be a bottleneck.

Radio network layer also includes functionality that directly impacts transport layer QoS. An example of this is 3G HSPA flow control and congestion control. These features operate at FP (Frame Protocol) layer, but they govern the traffic flow over the Iub. As algorithms are generally not specified by 3GPP, details are subject to implementation.

8.2 TCP and UDP as End User Transport Layer Protocols

In Chapter 3, user plane protocol stacks for the mobile network system were presented, showing how end user services are delivered. Due to its simplicity, versatility and efficiency, the TCP/IP protocol suite is now the dominant technology that is used by Internet based applications and services. The high data rate and low latency provided by the evolved radio access systems enable the migration of these services and applications to mobile environments; user traffic over the mobile backhaul is dominantly TCP/IP based.

Popular applications cover a wide range of services with specific QoS requirements such as voice (Skype™), web browsing, web-mail (Gmail™, Yahoo™ mail), instant messaging (MSN™, GoogleTalk™), social networking (Facebook™, MySpace™, LinkedIn™), image and video listing (Flickr™, Picasa™), micro-blogging (Twitter™), video sharing (Youtube™), on-line encyclopedia (Wikipedia™), virtual map and navigation (Google™ Maps), peer-to-peer (BitTorrent™), online gaming (WoW™), on-demand Internet streaming media (Netflix™), Internet radio, etc.

The TCP/IP protocol suite is a four-layer system consisting of the application (HTTP, FTP, e-mail, etc.), transport (TCP, UDP), network (IP, routing protocols, ICMP, etc.) and link layers (Ethernet, etc.). The role of the transport layer is to provide transport service to the application layer above, i.e., a flow of data between two equipments (UEs, UEs and servers, etc.). The two dominant transport layer protocols, TCP and UDP are specialized to serve two distinct, well defined sets of applications. TCP, which provides a reliable connection oriented service (despite the fact the IP protocol as such is unreliable) is used by data applications such as FTP, HTTP, e-mail, etc., requiring error free delivery of the whole data sent from one terminal to the other, whereas UDP, which transfers the data from one end to the other without any guarantee regarding reliable delivery is well suited for real time applications such as VoIP and VoD or for gaming and supporting short queries such as DNS lookup at web page download.

8.2.1 UDP

The UDP protocol provides a simple, connectionless, datagram based, unreliable transport service. UDP packets are referred to as datagrams. The UDP layer sends the data generated and handed over by the application without further delay and without any guarantee that the data will ever reach its destination or that the delivery of datagrams will be in order. The UDP header (refer to Figure 4.21 in Chapter 4) carries the source and destination port numbers, a checksum and information on the length of the UDP datagram. Based on the source and destination port numbers, the receiver is able to de-multiplex the incoming UDP datagrams and to deliver them to the corresponding application.

Data transfer over UDP does not require connection setup at the transport layer; however, the application is allowed to set up a connection for transferring data, for example VoIP is UDP based, where the connection set-up is done by specialized protocols such as SIP. UDP also allows point-to-multipoint communication. These characteristics of the UDP protocol make it suitable for real time applications (VoIP, VoD, etc.) where the timely delivery of data is more important than error free communication and over-delayed packets are dropped.

Another common use of UDP is when short queries are issued to network servers (for example DNS lookup).

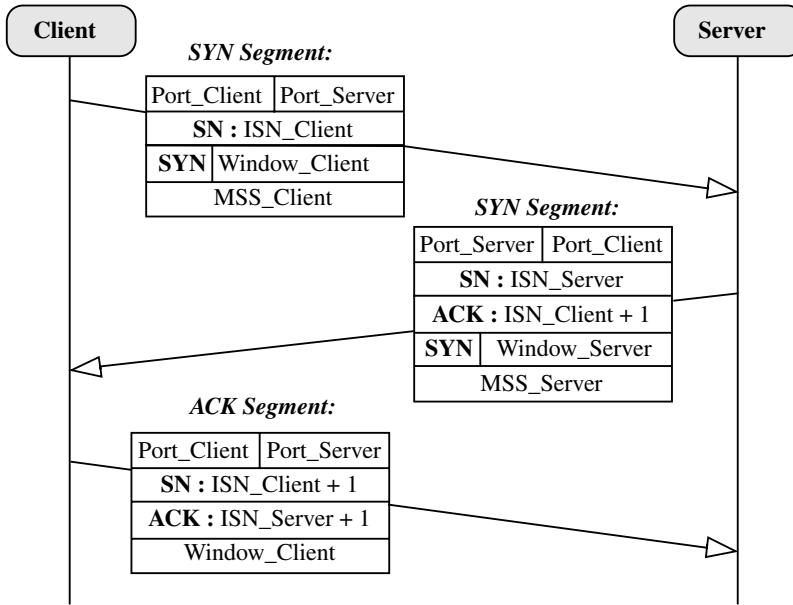
8.2.2 TCP

TCP offers a connection oriented, point-to-point, full-duplex and reliable service for data transfer over the Internet or over any IP based network that is by nature an unreliable media. The data sent in one packet is always referred to as segment. Before data transfer, the TCP connection must be established with the connection set up protocol (three-way handshake, see Figure 8.4), whereas after all the data has been transferred the connection has to be closed separately with the connection termination protocol.

During the three-way handshake, the two endpoints involved in the communication exchange information about the port numbers through which the communication will be performed, the initial sequence number (required for the reliable communication), the advertised window size that indicates the amount of data the endpoints can receive (this prevents buffer overrun) and the maximum segment size (*MSS*)-included as an option field to the TCP header – that an endpoint will be able to send/receive (required to define the best size for the TCP segments).

The header of each TCP segment contains the source and destination port numbers that identify the sending and receiving application. (The header is illustrated in Figure 4.23 in Chapter 4). These port numbers together with the source and destination IP addresses in the IP header of the packet carrying the TCP segment uniquely identify the TCP connection, which is referred to as a socket in the networking API. The connection is full-duplex as both endpoints can send and receive data.

Sequence numbering is used to facilitate in order delivery of segments, to identify duplicate reception of the same data and is used as a reference when correctly received data is acknowledged. Each TCP segment is sent with a sequence number. Acknowledgements are sent only for correctly received data; the source is not explicitly notified about missing or erroneously received data. Data is retransmitted if it is not acknowledged in time or is considered to be lost based on the received acknowledgements. With an acknowledgement, all



SYN – synchronize sequence numbers

SN – sequence number

ISN – initial sequence number

MSS – maximum segment size

ACK – acknowledgement

Figure 8.4 Client originated TCP connection set-up: three-way handshake.

the correctly received data up to the missing TCP segment is acknowledged (cumulative acknowledgement). TCP has efficient congestion control mechanisms that operate based on the assumption that packet losses are always due to congestion as in wired networks it is very unlikely that a packet is discarded due to bit errors. However, this assumption is not valid in case of the WCDMA or LTE air interface.

8.2.3 TCP Congestion Control

The first TCP implementations had no congestion control mechanisms until it was recognized that this would cause congestion collapse or sustained overload with a high packet drop ratio. Therefore, the first congestion control algorithm, referred to as Tahoe, was introduced in 1988 [4].

Nowadays, the TCP congestion control algorithm is the dominant end-to-end mechanism to control congestion on the Internet. The main scope of the TCP congestion control [5] is to apply efficient techniques whenever packet loss is detected or assumed in order to allow the system to recover and at the same time to guarantee efficient resource usage and data transfer.

There are several TCP versions that differ by the applied congestion control mechanisms. The most common version is TCP New Reno [6] that has the following congestion control elements: the Congestion Window ($cwnd$), the Additive Increase and Multiplicative Decrease (AIMD) mechanism that controls the value of the $cwnd$ (i.e., the maximum amount of in-transit data in octets between the two endpoints) and the Slow-Start, Congestion Avoidance, Fast Retransmit and Fast Recovery algorithms. These mechanisms are those that the source can use in order to provide reliable and efficient data transfer.

An established TCP connection starts sending data in Slow-Start mode: the sender initiates the data transfer by setting the size of the $cwnd$ to the MSS and sending the first segment. The size of the $cwnd$ is increased by one MSS each time the reception of a segment is acknowledged by the receiver. This means that during the Slow-Start, the size of the $cwnd$ is doubled every round trip time, i.e., its increase is exponential, as illustrated in Figure 8.5. The amount of data the sender is allowed to transmit is limited by the minimum of the $cwnd$ and the advertised window.

The exponential increase is continued until a predefined threshold (the slow-start threshold) is reached, a loss is detected or timeout occurs. When the slow-start threshold is reached, the

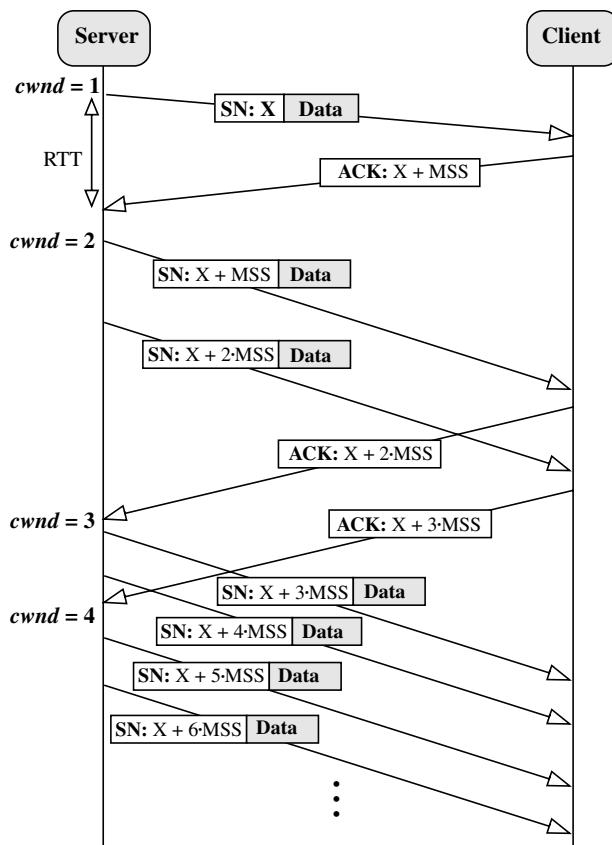


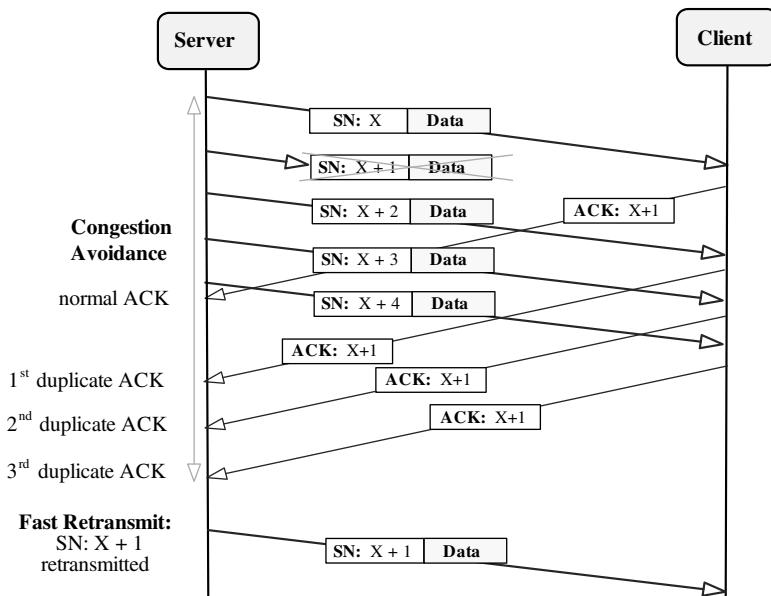
Figure 8.5 Example of Slow-Start.

TCP connection enters Congestion Avoidance mode. In this mode, the *cwnd* is increased additively with one *MSS* per each round trip time which is a much less aggressive increase mechanism than the one used during Slow-Start. The increase of the *cwnd* is continued until packet loss is detected. The receiver acknowledges the correctly received data with acknowledgements (ACKs) containing the sequence number of the TCP segment it is expecting to receive next.

In case of a loss, the receiver will be able to send an ACK only if a consecutive segment is correctly received. If that happens, the receiver notifies the source about the missing segment by sending an ACK with the sequence number of the first missing segment. This ACK is referred to as a duplicate ACK as it acknowledges the same segment as the ACK sent as a response to the last correctly received segment. Duplicate ACKs are sent without delay in order to assist the proper operation of the source.

Loss is detected by the sender when it receives multiple duplicate ACKs (ACKs with the same sequence number). A packet is considered to be lost when three duplicate ACKs are received, i.e., four consecutive ACKs acknowledging the same data. Upon this event, the packet that is considered to be lost (the one with the sequence number indicated by the duplicate ACKs) is retransmitted immediately. This procedure is called Fast Retransmit and depicted in Figure 8.6.

After the Fast Retransmit, the sender enters Fast Recovery state. Not executing the retransmission immediately after the first duplicate ACK helps the system avoid retransmission due to reordering of TCP segments. Triple duplicate ACKs indicate to the source (in addition to the fact that one segment is possibly lost) that the connection can still transfer data, therefore there is no point in drastically reducing the rate of the connection. Accordingly, at



Note: for the sake of simplicity, the sequence numbers are shown as multiple of the MSS (for example X + 2 means X +2·MSS).

Figure 8.6 Triple duplicate ACKs followed by Fast Retransmit.

Fast Recovery, the slow-start threshold is set to one half of the minimum of the current value of the $cwnd$ and the receiver's advertised window and then the $cwnd$ is set to the slow-start threshold plus three times the maximum segment size (this is called window inflation). The $cwnd$ is incremented by one segment after the reception of each duplicate ACK and if the value of the $cwnd$ (and advertised window) allows, a new segment is transmitted. The connection stays in Fast Recovery state until an ACK acknowledging all previously unacknowledged segments is received. Upon the reception of this ACK, the connection enters Congestion Avoidance mode with $cwnd$ reset to the slow-start threshold (this is called window deflation). In case of a partial acknowledgement (i.e., when only part of the segments sent before entering into Fast Recovery are acknowledged), the first unacknowledged segment is retransmitted.

For each segment sent, the sender maintains an associated retransmission timeout timer (RTO) which defines the time when an ACK is expected from the receiver. The value of the timer is estimated based on the round trip time measurements. In case a timeout occurs, i.e., no ACK is received before the timer expires, the segment is retransmitted; the slow-start threshold is set to the maximum of the following two values: two times the MSS and half of the actual $cwnd$ value; the $cwnd$ is set to MSS ; the RTO timer value is doubled and the TCP connection enters Slow-Start. After each unsuccessful attempt to transmit the segment, the RTO timer value is doubled until it reaches 64 seconds. This procedure is referred to as exponential backoff.

As discussed, the first TCP version with congestion control was TCP Tahoe that incorporates the Slow-Start, Congestion Avoidance and Fast Retransmit mechanisms. When triple duplicate ACKs are received, the source retransmits the missing segment, sets its $cwnd$ to one MSS and enters Slow-Start. This is not efficient in case of a transient packet loss as the rate of the source is unnecessarily reduced.

TCP Reno implements a simpler version of the Fast Recovery algorithm where after three duplicate ACKs the missing segment is retransmitted. The source will leave the Fast Recovery state when an ACK acknowledging new data is received, i.e., TCP Reno is able to retransmit only one missing segment per round trip time as the exit criteria does not mandate the reception of a full acknowledgement (i.e., for each segment sent before entering into Fast Retransmit/Fast Recovery). In case only one segment was lost, the new ACK will acknowledge the whole amount of data, whereas in case of multiple losses some but not all the segments are acknowledged. This is called partial acknowledgement.

TCP CUBIC [7] is the default congestion control algorithm of the TCP stack used in Linux and Android kernels. It was introduced in order to handle the TCP efficiency problem observed in case of high speed long distance networks. These networks have large bandwidth delay product which determine the required amount of in-transit packets between the source and destination in order to achieve high utilization of resources.

The $cwnd$ of the TCP versions with additive increase during congestion avoidance, e.g., after a congestion event, is not able to reach this value fast enough. Some connections last for a shorter time than the time that would be required to reach the bandwidth delay product. TCP CUBIC solves this problem by replacing the linear growth function with a cubic function. When a loss is detected, the regular Fast Retransmit, Fast Recovery mechanisms are executed; the $cwnd$ is reduced by a factor (with default value of 0.2); the window size before loss is recorded. This window size (W_{max}) is considered to be the saturation point. As shown in Figure 8.7, until W_{max} is reached, the $cwnd$ growth is according to the concave portion of the cubic function that has its plateau at W_{max} . The concave growth is followed by a convex increase after W_{max} is reached. This growth

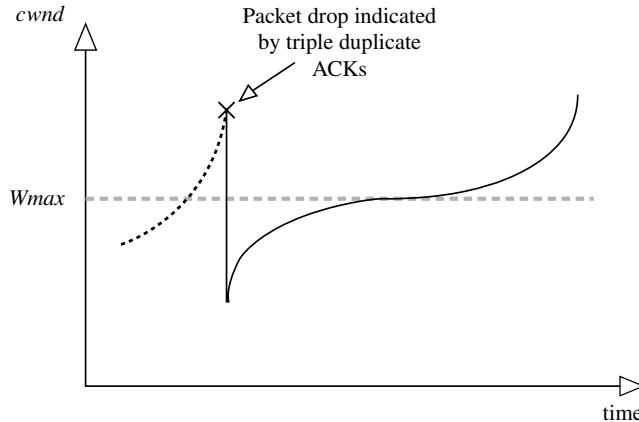


Figure 8.7 Schematic diagram of the window growth function of CUBIC.

function ensures that the saturation point is reached in a short time, the *cwnd* is kept almost constant around *Wmax* and finally that the *cwnd* is carefully increased in case no loss is detected. The *cwnd* growth depends on the elapsed time (*t*) from the last congestion control event (i.e., a Fast Recovery), this promotes fairness among competing connections as each can have approximately the same *cwnd* size even in case of different round trip times. The algorithm estimates the window size that TCP Reno would reach within the elapsed time (*t*) and sets *cwnd* to this value in case this is above the one calculated with the cubic function. This is the TCP friendly region of the CUBIC algorithm.

Compound TCP [10] is a congestion control algorithm developed for Windows (since Vista and Server 2008). It maintains a special *cwnd* that consists of two components: one that is increased in the same way as in TCP Reno and a delay based window that is increased when the network is underutilized, i.e., the delay is small, and decreased when the experienced delay grows, i.e., when congestion builds up. The algorithm tries to maintain the value of the *cwnd* (i.e., the sum of the two components) constant as much as possible.

The TCP congestion control mechanism has several limitations such as the reliance on packet loss in order to detect congestion, the inaccuracy of the delay or round trip time estimation and their inability to distinguish between possible reasons for packet drops. The latter triggers unnecessary congestion control actions when a drop is not due to congestion, which happens in wireless environments when a packet is dropped due to bit errors, mobility, etc. This deteriorates the performance of TCP over mobile networks as discussed in detail in the next section.

In order to allow the congestion control algorithm to take actions before a packet drop, i.e., to slow down the transmission rate before the network nodes are forced to drop packets, the Explicit Congestion Notification (ECN, RFC 3168) was introduced. The ECN mechanism sets notification bits in the IP header when congestion is detected, a solution that enables the two endpoints to reduce the rate of the connection.

One possibility to improve the delay and round trip time measurement is to apply the TCP timestamp option (RFC1323) by attaching a timestamp to each TCP segment. Based on the timestamp, the source can have an accurate round trip time measurement and *RTO* timer value estimation.

8.2.4 TCP Over Wireless

The various TCP versions discussed so far are all based on the assumption that bit errors and thus data discards due to erroneous receptions are extremely rare and the reason for packet loss is always network congestion. In mobile environments, the requirement of efficient TCP operation raises several issues as packet losses are frequently caused by bit errors occurring on the air interface. As the TCP source is not able to differentiate between the packet losses caused by bit errors and those caused by congestion, it will erroneously reduce the rate of the connection in case of bit errors too as if there was congestion. The solutions described in this section are examples from a wide range of proposals regarding TCP optimization. In practice, this type of proposals would need to be supported in the TCP/IP stack of the terminal and also commonly in TCP/IP SW on the servers, etc (depending on the proposal in question).

The fact that TCP was developed and optimized for transmission over wired links can become a serious issue whenever TCP based traffic is transmitted over wireless links. UMTS and LTE radio layer features such as Automatic Repeat-reQuest (ARQ) and Hybrid ARQ (HARQ) have been introduced in order to handle air interface errors via retransmission of the missing data. Air interface imperfections are not the only sources of packet loss as the radio conditions are changing due to user mobility: handovers, sudden coverage holes (transient coverage problems) and network asymmetry (in LTE, UL coverage might disappear near cell edges while DL coverage still exists) might result in packet loss as well.

In addition to packet loss caused by air interface problems and mobility, TCP might experience performance degradation due to unpredictable delays, i.e., sudden change of the TCP round trip time: due to the forwarding of TCP segments over X2 (LTE), ARQ and HARQ retransmissions or due to different circumstances (longer transmission path, larger buffer, higher load) after a successful handover. TCP retransmissions can be triggered by out of sequence delivery of data in case of handovers and network asymmetry. In these cases, the efficiency of data transfer is limited, the *RTO* timer of the TCP might expire, triggering TCP slow start that will result in long recovery times (the recovery time is the time required to reach the throughput experienced before the problem occurred) for the TCP sources.

The limited battery capacity of the mobile devices requires efficient data transfer in wireless environments that currently TCP is not able to provide. TCP efficiency is a major issue also from a resource usage point of view as the air interface and the microwave links are scarce resources compared to the high capacity wired links.

The TCP efficiency problem over the radio interface has resulted in a multitude of candidate solutions. One possible classification divides them into two groups: network supported solutions and end-to-end solutions. Network supported solutions require extra functionality at the network side whereas the end-to-end solutions are transparent to the underlying network.

Network side solutions (Snooping Protocols, Split TCP solutions, etc.) are based on the idea of implementing extra functionalities at given network nodes. The functionality intercepts the TCP segments of a given connection and performs actions with them on behalf of the receiver in order to prevent congestion control actions whenever packet loss is not due to congestion. These solutions may benefit from the immediate information about the air interface imperfections, therefore their ideal location is at the eNBs (LTE) or RNCs (WCDMA/HSPA). In case of handovers, the information on the status of the intercepted connection must be exchanged between the involved nodes that might result in increased handover latency especially in case of LTE, a fact that justifies the placement of the extra functionality in the gateway nodes.

Snooping protocols hide the air interface from the transport (TCP) layer without violating the end-to-end semantics of the connection.

In case of the Snoop Protocol [11], the packets of the connection are intercepted, copied by the Snoop module (located at the eNB or RNC). Copies of the packets are stored until an acknowledgement is received. Triple duplicate ACKs are discarded followed by a retransmission of the missing segment. WTCP [12] improves the Snoop Protocol by measuring the round trip time over the air interface with the help of the TCP timestamp option. This measurement is used to manipulate the sender in such a way that retransmission timeouts are avoided.

TCP SACK-Aware Snoop Protocol [13] uses the functionality provided by the TCP SACK (selective acknowledgement) (RFC2018) option. SACK reduces the amount of the retransmitted data as the receiver can inform the source in case non-continuous blocks of segments have been received. The Snoop module retransmits the missing segments indicated by the duplicate ACKs or the blocks of missing segments indicated by the selective ACKs.

Split TCP solutions divide the TCP connection into two (or more) separate connections. The node at which the connection is split, i.e., the interfacing node is referred to as the proxy. Packets are intercepted, buffered and acknowledged to the source by the proxy. The simplest solution is to create two TCP connections (Indirect-TCP [15]) in order to isolate the air interface and to handle the air interface problems with a shorter TCP connection.

A more sophisticated solution is to use Radio Network Feedback containing information on the instantaneous channel quality or available bandwidth when the window size of the connection between the proxy and the UE is calculated [16].

Instead of requiring and relying on extra functionalities at the network equipments, end-to-end solutions require specialized mechanisms at the sender and receiver that are able to handle both congestion and air interface error caused packet losses. This approach maintains the end-to-end semantics of the connection.

The negative impact on the connection performance caused by handovers, i.e., the timeouts due to temporary disconnection is handled by the Freeze TCP [17] by enabling the UE to send a zero window advertisement (ZWA) to the sender before the handover is executed. Upon receiving the ZWA, the sender enters in persist mode and stops sending data to the receiver. The data transfer is resumed when upon reconnection, the UE sends a non zero window advertisement to the sender. This mechanism prevents timeouts and *cwnd* shrinkage due to handovers and allows the sender to continue the data transfer with the *cwnd* value it had before handover.

TCP-Westwood [18] proposes a sender side enhancement that allows the estimation of the available bandwidth by monitoring the rate of the ACKs. Congestion window and slow start threshold are set to this value when the sender recalculates them as a result of a detected congestion event (triple duplicate ACKs or timeout).

8.3 DSCP, Traffic Class, and Priority Bits

8.3.1 Differentiated Services

In the mobile backhaul, transport level quality of service (QoS) solutions are based on the Differentiated Services (DiffServ) approach which provides a relative service differentiation for connections aggregated into a limited number of well defined QoS classes based on the

DiffServ code point (DSCP) encoded into each IP packet's header. 3GPP requires the use of DSCPs and a mapping of traffic categories into DSCPs as the means of transport layer QoS marking.

Traffic handling and scheduling decisions at the transport nodes are made based on the DSCP, that is, the network nodes differentiate the traffic based on its class marked by the DSCP. For each class a separate handling and forwarding policy is defined that is referred to as PHB (Per Hop Behaviour). Connections belonging to the same service class should be marked with the same DSCP in order to ensure the same level of service.

Differentiated Services (DS) classifies and conditions traffic at the network boundary (ingress), and assigns it to a behavior aggregate. The behavior aggregate is marked by a DS codepoint (DSCP). Network nodes then forward and schedule the packet according to the marked codepoint.

A DS domain consists of boundary nodes and interior nodes. Boundary nodes classify and condition traffic at the ingress. DS domain nodes then select the per-hop-behaviour according to the DS codepoint marking. A DS domain is managed typically by a single entity but may consist of multiple networks. To support Differentiated Services between DS domains, Service Level Agreements (SLAs) are needed.

A classifier/conditioner is shown in Figure 8.8. Classifier matches incoming packets against a rule. At its simplest it is the DS codepoint field but can consist of other header fields, such as source and destination addresses, port numbers, protocol info and so forth.

Traffic profile is needed to decide, whether the incoming traffic is in-profile or out-of-profile, with a meter, which measures the incoming traffic. Packets which are out-of-profile, may be subject for further conditioning, e.g. queueing, discarding, re-marking, or other actions. Shaper may queue the packet in order to make it comply with the profile. Alternatively a non-conforming packet can be dropped.

In general in backhaul, marking should be done by the radio access nodes as the markers should follow the mapping rules of radio QoS parameters to their transport counterparts. The required information is available at the radio access nodes. This is one of the means to ensure consistent end-to-end QoS provisioning.

Obviously, classification and conditioning has a lot to do with trust. If traffic sources are known to behave well and can be trusted, there is less need for classification and conditioning

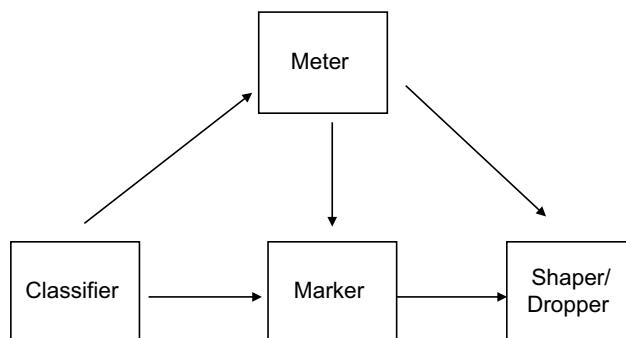


Figure 8.8 Classifier/conditioner [20].

at the subsequent nodes. When traffic crosses an organizational boundary or administrative domain, it is more likely to be classified, re-marked and maybe policed.

DSCP has meaning only over a given network domain, therefore the marking is done at the boundary. When traffic is handed over to the next domain, re-mapping/re-writing of the DSCP might be required in order to ensure end-to-end consistency. The subsequent network nodes can then use appropriate buffer and scheduler management functions and traffic conditioning, to realize the service. The specification does not intend to specify algorithms or functionality of how to realize the required behaviour, but the characteristics of the service.

DSCP itself is a number between 0 and 63 (decimal). Approximately half of the values are standardized (recommended), the other part is for local definition.

The starting point for QoS marking and differentiation is the DS field of the IP header for the packet based mobile backhaul. This is also specified in 3GPP. Mobile network elements, such as eNodeB, support DSCP marking. How the DSCP is derived, is discussed separately for each radio technology.

When the DS field of the IP packet is marked by the mobile network node, e.g. a eNodeB, subsequent DS nodes can then forward the packet accordingly. The eNodeB may also mark the related fields in the headers of other protocol layers, e.g. in the Ethernet frame using the priority bits (Assuming an IEEE802.1Q tagged frame).

8.3.2 IPv6

IPv6 supports the same DiffServ architecture as IPv4 does. In the mobile backhaul, the marking of DSCPs is performed as in the case of IPv4. For DiffServ, the field in the IPv6 header is called Traffic Class (TC).

The end user IP traffic, whether IPv4 or IPv6, is tunneled within GTP-U (and other protocols, depending on the radio network technology) transparently over the radio network.

One QoS related difference in a mobile backhaul IPv6 application is the larger header size of IPv6. This needs to be taken into account when calculating the required bandwidth on lower layers and on physical links. Without extension headers, IPv6 header is 40 bytes, as opposed to a 20 byte header (minimum) of IPv4.

8.3.3 Per-Hop Behaviours

Per-hop-behaviours include the Default or Best Effort Behaviour, Expedited Forwarding, and Assured Forwarding.

Default behaviour (Code 000000) marks Default Per-Hop Behaviour or the best effort class. No specific guarantees are given.

Expedited forwarding (EF) is meant for traffic of low latency, low loss and assured bandwidth. It should be able to replicate a point-to-point service end-to-end over the Differentiated Services domain, however the end-to-end behaviour is not in the scope of the definition of the EF PHB. EF traffic should not be queued (buffered) and in order to avoid buffering, the service rate should be at least equal to the arrival rate. EF PHB is meant to address guaranteeing the service rate. Traffic conditioning is needed to make sure that the

Class 1	Class 2	Class 3	Class 4	
AF11 001010	AF21 010010	AF31 011010	AF41 100010	Low drop prec.
AF12 001100	AF22 010100	AF32 011100	AF42 100100	Medium drop prec.
AF13 001110	AF23 010110	AF33 011110	AF43 100110	High drop prec.

Figure 8.9 AF forwarding classes.

arrival rate is bounded. EF is recommended to use the DS codepoint 101110. Packets marked with the EF codepoint must not be demoted, and if the code point is changed, the new codepoint must satisfy the EF PHB requirements.

EF PHB can e.g. be implemented as a strict priority queue, which is serviced whenever there are packets to be sent, but other queue implementations are possible, as long as the requirements of the EF PHB are met.

Assured Forwarding (AF) classes are shown in Figure 8.9. An AF PHB group consists of four classes, each of which supports three drop precedences. AF classes are not allowed to be aggregated together, and a minimum amount of resources must be allocated to each implemented classes. Excess resources may be allocated to the classes, the algorithm of how this is done is not covered in the specification. Traffic conditioner may modify the drop precedence, and also allocate traffic to another class.

Short term congestion (bursts) can be handled by queuing packets. An AF Implementation attempts to minimize long term congestion within each class, with an active queue management algorithm. Examples of such algorithms are Random Early Discard (RED)¹ and Weighted RED (WRED). Dropping algorithm should be insensitive to the packet bursts, and react to the long term congestion. AF PHB can be implemented e.g. with a weighted round-robin scheduler.

8.3.4 Recommended Use of DSCPs and Treatment Aggregates

For the DSCPs, a recommendation for usage is provided in RFC 4594. Recommended values help interoperation at the administrative domain boundaries. This does not yet provide a direct answer on how to map the mobile backhaul traffic. The basic principle with DSCP allocation however becomes clear.

Network control means routing and other network control traffic that is essential in keeping the network operational. OAM traffic is also network control traffic, although it is separated in Table 8.1. Network control traffic is network operator's (service provider's) traffic, instead of user traffic.

User traffic includes different types of services, and also user signalling (example is IP telephony signalling) with demands for loss, delay, and jitter that are specific to the service.

¹ RED was originally introduced in order to prevent TCP synchronization. The RED together with the TCP Congestion control are the two dominant congestion control mechanism of the Internet. The scope of RED is to drop randomly selected packets (i.e., arriving packets are discarded with a certain probability that depends on the load of the buffer) in order to trigger TCP CC control actions and thus to reduce the load.

Table 8.1 Recommendation for DSCPs [29].

Service class name	DSCP name	DSCP value	Application examples	Traffic characteristics	Loss tolerance	Delay tolerance	Jitter tolerance
Network control	CS6	48	Network routing	Variable size packets, mostly inelastic short messages, but traffic can also burst	Low	Low	Yes
Telephony	EF	46	IP telephony bearer	Fixed size small packets, constant emission rate, inelastic, low-rate flows	Very low	Very low	Very low
Signaling	CS5	40	IP telephony signaling	Variable size packets, somewhat bursty short-lived flows	Low	Low	Yes
Multimedia conferencing	AF41, AF42, AF43	34, 36, 38	H.323/V2 video conferencing (adaptive)	Variable size packets, constant transmit interval, rate adaptive, reacts to loss	Low-Medium	Very low	Low
Real-time interactive Multimedia streaming	CS4	32	Video conferencing and interactive gaming	RTP/UDP streams, inelastic, mostly variable rate	Low	Very low	Low
Broadcast video	AF31, AF32, AF33	26, 28, 30	Streaming video and audio on demand	Variable size packets, elastic with variable rate	Low-Medium	Medium	Yes
	CS3	24	Broadcast TV & live events	Constant and variable rate, inelastic, non-bursty flows	Very low	Medium	Low
Low latency data	AF21, AF22, AF23	18, 20, 22	Client/server transactions, Web-based ordering	Variable rate, bursty short-lived elastic flows	Low	Low-Medium	Yes
OAM	CS2	16	OAM&P	Variable size packets, elastic & inelastic flows	Low	Medium	Yes
High-throughput data Standard	AF11, AF12, AF13 DF (CS0)	10, 12, 14 0	Store and forward applications Undifferentiated applications	Variable rate, bursty long-lived elastic flows A bit of everything	Low	Medium-High	Yes
Low-priority data	CS1	8	Any flow that has no BW assurance	Non-real-time and elastic	High	Not specified High	Not specified Yes

Default forwarding (marked standard) means a best effort (BE) service. This service is defined for the background low priority data traffic that is elastic.

Assured Forwarding is an ‘enhanced best effort’ service. Multiple DSCP values are provided for identification of the traffic. A common queue stores the aggregate, and an active queue management (e.g. RED or WRED) is used to protect the network and limit the delays.

CS stands for Class Selector, defined for IPv4 precedence queuing in RFC1812. Differentiated Services maintains a level of backwards compatibility to the Class Selectors. Class selector uses three bits.

Another informational RFC, RFC 5127, proposes that the service classes as presented, are mapped into treatment aggregates. The benefit is in reduction of different behaviours that the network nodes need to support. The supported service classes are aggregated into treatment aggregates. RFC5127 discusses four treatment aggregates: Network control, Real-time, Assured elastic and Elastic. See Table 8.2.

Multiple DSCP values are in Table 8.2 collected into an aggregate, except for the Network control aggregate, which includes only Class selector 6. Real-time traffic maps into a per-hop-behaviour of EF (Expedited Forwarding). Assured elastic uses Assured Forwarding, while Elastic is a Default or Best Effort – class.

8.3.5 DSCP in IP Tunnels

IP is in some cases tunnelled, so that the DSCP marking of the inner IP packet is not visible. In the mobile backhaul, one such type of an application is with the use of IPSec. In the IPSec tunnel mode, the complete IP packet is encapsulated within a new IP packet with a new IP header, and an AH/ESP header. The inner packet DSCP value can be copied to the outer packet DSCP field, so that DS nodes are able to support QoS, even though the inner packet header is not available. RFC2983 provides a discussion of Differentiated Services and tunnels.

Note that the end user level IP packet is tunneled with GTP-U (and/or other radio network layer protocols, depending on the radio network technology) through the radio network. In this case, the DSCP from the inner IP packet is not directly copied to the backhaul layer IP header. The inner (end user) IP packet is carried transparently over the radio network. QoS marking for the backhaul layer IP header originates from the radio network layer QoS parameters.

8.3.6 Use of DSCPs for Mobile Backhaul

The traffic types existing depend on the radio network technology. In 3G, there are more traffic types than in LTE, due to the Iub interface and the functional split between the RNC and the NodeB.

The recommendations discussed for DSCP usage need to be adapted for the radio network application, and thus example mappings are studied in later sections further for each of the radio technologies.

Even with user traffic, there are additional requirements due to constraints set by the radio network architecture. In 2G, this is due to the BSC and the BTS, and the radio L2 protocols in

Table 8.2 Treatment aggregates [30].

Treatment aggregate	Loss tolerance	Delay tolerance	Jitter tolerance	Service class name	Loss tolerance	Delay tolerance	Jitter tolerance	Treatment aggregate behavior	DSCP values
Network control	Low	Low	Yes	Network control	Low	Low	Yes	CS6	EF, CS5, AF41, AF42, AF43, CS4, CS3
Real-time	Very low	Very low	Very low	Telephony	Very low	Very low	Very low	EF	
Assured elastic	Low	Low-Medium	Yes	Signaling	Low	Low	Yes	Yes	
				Multimedia conferencing	Low-Medium	Very low	Low	Low	
				Real-time interactive	Low	Very low	Low	Low	
				Broadcast video	Very low	Medium	Low	Low	
				Multimedia streaming	Low-Medium	Medium	Yes	AF	CS2, AF31, AF21, AF11, AF32, AF22, AF12, AF33, AF23, AF13
Elastic	High	High	Yes	Low latency data	Low	Low	Low-Medium	Yes	
				OAM	Low	Medium	Medium	Yes	
				High-throughput data	Low	Medium-High	High	Yes	
				Standard	Not specified	Not specified	Not specified	Default	Default (CS0), CS1

between over the Abis. As an example, carrying the traffic for a file download via GPRS on a pseudowire requires real-time delivery for the pseudowire, due to the scheduling in the BSC. The end user application itself – file download – has less strict requirements. The general mapping principles discussed are adapted to cope also with the constraints due to the radio technologies.

One topic is the mapping of voice traffic compared to the mapping of radio network control traffic – such as NBAP in 3G and S1-AP in LTE. Voice traffic has stringent requirements regarding delay and delay variation. To provide this service with good quality to the end user, voice traffic is given priority over the control plane traffic. The control plane traffic is tolerant for delay and delay variation, while voice needs deterministic service.

Control plane traffic however needs to have an amount of bandwidth, otherwise the signaling messages will not get through. When voice is marked with DSCP 46 (EF) and radio network control plane traffic with DSCP 34 (AF41), there must be an assurance that the radio network control plane does not starve.

Note that lately in RFC5865 an additional DSCP for voice traffic – 44 – has been proposed. The traffic would be treated as EF. RFC5865 describes different implementation possibilities regarding prioritization of traffic marked DSCP 46 and those marked DSCP 44. This may open up further configuration possibilities.

8.3.7 *MPLS Traffic Class*

MPLS header has a 3-bit field for marking the QoS. Originally the bits were named EXP after Experimental, which caused confusion. RFC 5462 (Feb 2009) corrected this and named the bits as TC (Traffic class).

There is no recommendation on how the eight possible values should be interpreted: which traffic type should be marked with what numerical code. The label switch routers at the edge either have a default mapping between DSCPs and traffic classes, or a configuration (mapping) table is supported.

Two possible MPLS QoS marking models are defined: marking of the QoS to the TC field of the MPLS header. This type of LSP is called Explicitly TC-encoded PSC LSP, or E-LSP. PSC stands for Per Hop Behaviour Scheduling Class. Other model is L-LSP (label-only-inferred-PSC LSP), in which case the L-LSP defines the PSC, and drop precedence is marked into the TC bits.

As the number of bits available in the TC field is identical to what Ethernet IEEE802.1Q priority bits can provide, one can reuse the logic followed while defining the 802.1Q recommended value settings. This would be aligned with Differentiated Services, as the IEEE 802.1Q defined priority bit usage is aligned with Differentiated Services.

8.3.8 *IEEE802.1Q Priority Bits*

In the previous sections we focused on marking the IP packets. In many cases the access lines are Ethernet-based, possibly without a capability to investigate IP layer headers. It is useful to mark the priority also to the Ethernet frames. The IEEE802.1Q frame includes three priority bits for this.

Table 8.3 Priority values and their usage [32].

Traffic type	Characteristics	Acronym
background	bulk transfers, permitted but should not degrade other applications	BK
best effort	default, for unprioritized applications	BE
excellent effort	'CEO's best effort', delivered to its most important service (still best-effort type)	EE
critical applications	guaranteed minimum bandwidth required, traffic admitted by admission control	CA
'video', < 100 ms latency and jitter	video, or other applications requiring low latency	VI
'voice', < 10 ms latency and jitter	voice, max delay and jitter, as one-way through the LAN on a single campus	VO
internetwork control	Supports maintenance of the network over administrative domains	IC
network control	Guaranteed delivery required, supports maintenance of the network infrastructure	NC

The IEEE standard informatively defines the use of priority bits as in Table 8.3. The purpose is to facilitate interoperation and have a well defined and commonly understood way of indicating priority in the Ethernet IEEE 802.1Q frame.

The number of classes Ethernet bridges must have, is not defined in IEEE, so it is an implementation issue. Table 8.4 shows an example of how to map traffic, as a function of the number of queues.

With only a single queue available, all traffic is best effort. With two queues, voice and other real-time traffic can be separated into its own queue. Third queue allows separating network control. With four queues, critical applications are separated from the best effort, and so on. Granularity is increased up to the maximum of eight queues (given the three priority bits). Many equipment support e.g. four to six queues, which already allows for a clear differentiation of traffic.

Some Ethernet switches support DSCP marking, i.e. IP layer information. Others may be able to only use the information encoded into the Ethernet frame header. Table 8.3 presented some guidelines. In the end, this is subject to the network operator's decision. Configuration can be expressed as a mapping from DCSPs to PCPs, since DCSPs are the source information for the marking.

A straightforward mapping is to associate each PHB with a PCP:

- BE PHB is mapped to PCP 0.
- AF1 PHB to PCP 1.
- AF2 PHB to PCP 2.
- AF3 PHB to PCP 3.
- AF4 PHB to PCP 4.
- EF PHB to PCP 5.
- Network control traffic to PCP 6.

Table 8.4 Use of traffic classes as a number of queues available [32].

Number of queues	Class	Traffic types
1	BE	Best effort, background, excellent effort, critical applications, voice, video, internetwork control, network control
2	BE	Best effort, background, excellent effort, critical applications, voice, video, internetwork control, network control
3	BE	Best effort, background, critical applications, excellent effort
	VO	Voice, video
	NC	internetwork control, network control
4	BE	Best effort, background, critical applications, excellent effort
	CA	critical applications, excellent effort
	VO	Voice, video
	NC	internetwork control, network control
5	BE	Best effort, background, critical applications, excellent effort
	CA	critical applications, excellent effort
	VO	Voice, video
	IC	internetwork control
	NC	network control
6	BK	background
	BE	best effort
	CA	critical applications, excellent effort
	VO	Voice, video
	IC	internetwork control
	NC	network control
7	BK	background
	BE	best effort
	EE	excellent effort
	CA	critical applications
	VO	Voice, video
	IC	internetwork control
	NC	network control
8	BK	background
	BE	best effort
	EE	excellent effort
	CA	critical applications
	VI	video
	VO	voice
	IC	internetwork control
	NC	network control

This leaves PCP 7 unused. In some networks Ethernet frames with a PCP value 7 are discarded in order to prevent attacks.

PCPs do not support a notion of a drop precedence as defined for the AF PHBs at the IP layer. It is also possible to define a color marking by indicating two drop precedences within PCPs by using some of the PCP values for a color indication.

8.3.9 VLANs

In some cases, instead of marking QoS to the priority bits of the Ethernet IEEE 802.1Q frame, QoS is indicated by the VLAN ID. An example is an Ethernet service, which may be defined to use the VLAN ID as the source for QoS. VLAN ID may also guide mapping of the user packet flow at a PE device to a suitable service (a suitable pseudowire).

8.3.10 QoS with MEF Services

Metro Ethernet Forum (MEF) services include QoS definitions at a detailed level. MEF services, with their performance objectives, were discussed in Chapter 5. The QoS for the mobile backhaul application is addressed by the definition of the attributes of the service.

How does the mobile operator then indicate the required Class of Service at the User-to-Network (UNI) interface?

Details depend on the service. An EVC is logically the Ethernet layer connectivity between two (or more) UNIs. For the EVC, the indication for CoS may be a VLAN ID only, in which case the VLAN indicates the CoS. In this case EVC corresponds to the customer VLAN.

Alternatively, the CoS is indicated by the VLAN ID and additionally either the priority bits of the Ethernet 802.1Q frame, or the IP layer DSCP field. In this case, single EVC includes two or more classes of service.

Single CoS (color-blind policer)

If policing at the ingress is color-blind the single EVC – single CoS application for the base station requires that a guarantee is given for all of the traffic, meaning the full aggregate rate from the base station. Full aggregate traffic means all of the traffic types combined. This rate shall equal the CIR rate guaranteed in the SLA. Otherwise it cannot be ensured that critical real-time and control traffic does not suffer from policing. This approach however leads to a high value for the CIR attribute, since it now also covers background traffic.

Two CoSs (color-blind policer)

Preferably background traffic is treated as non-guaranteed traffic with an Excess Information Rate (EIR). Assuming a color-blind policer, this requires two Classes of Service. One CoS supports the real-time/control traffic with a CoS H (High) and the other one non real-time traffic with a CoS L (Low). The two CoSs are identified at the UNI by either

- two separate VLAN IDs (two EVCs), or by
- a single VLAN ID and Ethernet p-bits or DSCPs (single EVC with two CoSs).

As an example,

- with CoS High, CIR is defined and EIR is 0. CIR is the value that real-time/control traffic requires;
- with CoS Low, CIR is 0 (or a low value) and EIR is according to the traffic volume of the NRT/background types.

With multiple CoSs as above bandwidth fragmentation may become an issue. This depends on how CIR and EIR are defined in each of the classes, and what type of policing is performed

by the service provider. A Hierarchical Bandwidth Profile (H-BWP) defined in MEF addresses this issue.

Single CoS (color-aware policer)

MEF 23 also specifies a way for the customer equipment (CE) to predefine traffic within CIR (green) and within EIR (yellow). This is useful, assuming that the policing function of the service provider at the PE device ingress is color-aware. Now bandwidth is used more efficiently than with color-blind policers, due to

1. compared to Single CoS (color-blind policer), there is no need to have the CIR covering all traffic (real-time/control and NRT/background combined);
2. compared to Multiple CoS (color-blind policer), no bandwidth fragmentation occurs, as the unused CIR is usable for the background traffic.

The indication for the color-aware policer occurs by an Ethernet priority bit value or a DSCP. An example for the indication according to MEF 23 follows, with a Class of service Low.

At the Ethernet layer, as an example,

- green marking is indicated at the Ethernet frame by the least significant priority bit value of ‘1’;
- yellow marking is indicated at the Ethernet frame by the least significant priority bit value of ‘0’.

At the IP layer, similarly as an example,

- green marking is indicated at the IP header with a DSCP value of 10 (decimal) (AF11);
- yellow marking is indicated at the IP header with a DSCP value of 12 (AF12), 14 (AF13) or 0 (Default).

Additional issue is that the above marking is not by MEF defined for CoS High. A suitable CoS with color marking as in the above example would need to be selected.

The real-time and control traffic is shaped and pre-colored green by the base station when entering the service. Assuming a color-aware policer at the service provider ingress, the green frames pass through as long as they do not exceed the CIR rate at the ingress of the policer. Because of shaping, they do not exceed the CIR.² Shaping should take into account all relevant protocol overheads. These depend on what layer information is referred to with the committed information rate attribute. See Section 5.5 concerning Bandwidth profiles within MEF services.

Currently most of the MEF services available support only color-blind operation at the UNI. If this is the case, then multiple CoSs are needed in order to separate real-time/control traffic from the background.

²For this purpose, it is useful to consider some margin at the shaping function at the egress of the base station port. The shaping rate at the egress should be somewhat less than the defined CIR rate, in order not to lose traffic due to inaccuracies in the shaping/metering on either side of the UNI.

8.4 Ingress and Egress Functions

In this section, ingress and egress functions are discussed. Classification and policing are typically related to the ingress, while scheduling, queuing, queue management and shaping to the egress. These functions are dependent on the network node capabilities.

8.4.1 Ingress Classification and Policing

At the ingress of e.g. a service provider's node, traffic may be subject to classification and ingress policing. An example is policing related to MEF services.

These ingress functions are related to trust. Often at the boundary between administrative domains, ingress functions are executed. If the backhaul network is administered by an entity separate from the mobile network operator, the backhaul network operator may wish to control the amount of traffic entering his network by means of policing.

Ingress policing also helps to avoid excessive traffic overloading the network element. This is related to security as well, since it mitigates some of the risks related to flooding attacks. Again, if the network can be trusted, there is less need for ingress policing.

Control plane policing refers to limiting the amount of ICMP control messages, spanning tree BPDUs, routing protocol messages, and similar control messages, which all typically require processing from a processor unit of a network node. If a processor is overloaded with incoming packets to be processed, it may crash or stall causing unpredictable behaviour.

If the policer is QoS-aware, it may selectively drop color packets which are less sensitive to loss. The input for this is Diff Serv code point field. Depending on the equipment capabilities, information in addition to the QoS marking of the Diff Serv bits can be used, e.g. IP source and destination addresses, L4 ports, upper layer protocols and VLANs. Simpler devices may not support this type of functions at all.

Dropping non-conforming packets unselectively by a service provider creates an issue for the mobile backhaul. This is the case with color-blind policers. Assuming e.g. that a single BTS access capacity of 2 Mbit/s is guaranteed, but typically 20 Mbit/s would be available, means that the mobile operator as a service user is only sure that 2 Mbit/s of traffic will be accepted to enter the service. Voice and control traffic, even if not exceeding 2 Mbit/s, is not guaranteed if the aggregate traffic exceeds 2 Mbit/s, and packets are dropped blindly (i.e. nonQoS-aware) by the service provider ingress function.

If the policer is QoS aware and voice and other critical packets are marked with a high priority DSCP, the ingress policer of the service provider drops other packets before the high priority traffic. As long as voice and control traffic do not exceed 2 Mbit/s, throughput is guaranteed for this traffic. Other traffic types might suffer packet loss, in case the ingress device decides to drop traffic over the 2 Mbit/s rate.

Having QoS-awareness requires that network administrators agree on the interpretation of the QoS marking to be used: which traffic should be serviced as a priority, and which should be the first candidates for dropping; i.e. DSCP marking is interpreted identically.

Related to the QoS marking as received in the ingress, if the traffic sources are not trusted by the service provider, he may decide to re-classify the traffic before it enters his network. This might mean demoting part of the packet flows to a lower priority.

8.4.2 Single-Rate Two Color Policer

A policer for a single rate can be compared to a fixed size bucket that is filled with tokens at a constant rate, i.e. the targeted rate. Whenever a packet passes the policer, the policer checks that an adequate amount of tokens exist in the bucket. If so, the packet is colored green and the tokens are removed from the bucket. Otherwise the packet is colored red and discarded.

The size of the bucket determines the size of possible bursts. The larger the bucket, the more can the rate be exceeded for a short period of time. The size of the bucket has to be at least as large as the largest packet, otherwise such packets would always be discarded.

8.4.3 Two-Rate Three Color Policer

In many cases policing with a single rate is insufficient. Often the service provider guarantees a transmit rate of packets, which is the committed/guaranteed information rate. Another rate, the excess information rate, will be transmitted if possible. In the case of high network load, or a link failure, there might be insufficient capacity for the excess traffic. A policer needs to identify when traffic is conforming to the committed rate, to the excess rate, or whether it is violating both rates and is to be discarded. This information is encoded in three colors.

Such a two-rate-three-color marker (RFC2698) can be compared to two buckets that are filled at a constant rate. The fill rate corresponds to the committed and excess information rates, respectively. Whenever a packet passes the policer, first the amount of tokens in the committed information rate bucket is checked. If the amount of tokens is sufficient, the packet is colored green and tokens are removed from the bucket. If the amount of tokens is not adequate in the first bucket, availability of tokens in the second bucket is checked. The second bucket corresponds to the excess information rate. If this bucket contains enough tokens the packet is colored yellow and again the tokens are removed. If the tokens in the second bucket run out, the packet is colored red, i.e. it is discarded.

Both green and yellow packets are allowed to pass the system, but in case of resource limitations in the network, yellow packets suffer first. The information of whether a packet has been colored green or yellow has to be made available even outside of the node coloring the packets. This information could be encoded e.g. by DSCP remarking, e.g. the DSCP is changed from AFx1 to AFx2. It is also possible to encode this information on the link layer, e.g. by agreeing on specific VLAN PCPs to express yellow Ethernet frames, or by using DEI (Discard Eligible Information) where available in Ethernet headers. Policing can be applied on both IP and on lower layers. Depending on the layer, different amount of overhead needs to be accounted for, due to different size of packet headers.

8.4.4 Egress Scheduling, Queue Management, and Shaping

If traffic is classified at the ingress, DSCP marking tells the scheduler how to treat the traffic at the egress. Scheduler types include strict priority schedulers, weighted round robin schedulers and weighted fair queuing schedulers. This topic is implementation dependent. Schedulers are closely tied to the queues. Each traffic type (behaviour aggregate) typically has its own queue, in order to allow a differentiated treatment. A scheduler may then serve the queues according to a predetermined policy.

8.4.5 Strict Priority Scheduler

A strict priority scheduler serves the queues in a priority order. First, all packets from the 1st – highest priority – queue are scheduled. Only if this queue is empty, packets from the 2nd queue will be scheduled. If both the 1st and 2nd queue are empty, packets from the 3rd queue are scheduled and so on.

An issue with this scheme is that it might cause starvation to the lower priority queues. As long as the amount of traffic in the higher priority queues can be kept under control or it is known from dimensioning the system that this cannot take all available bandwidth, a strict priority scheduler can be used. Often traffic that is scheduled as strict priority has a means for limiting traffic at the ingress: via an admission control function, or via policing and shaping.

Another limitation is that there is no ratio defined in how to split the bandwidth among the queues. The queues are served in the priority order, and whether lower priority queues receive any service, depends on the amount packets in the higher priority queue.

8.4.6 Weighted Round Robin Scheduler

A round robin scheduler will serve the different queues in a round robin fashion: one after the other. A weighted round robin scheduler (WRR) serves the queues according to a weight. If a queue is assigned a weight of 2 and another queue is assigned a weight of 1, the amount of packets scheduled will have a ratio of 2:1.

WRR schedulers are work conserving, i.e., if a queue is empty and there are no packets to be scheduled, then packets from other queues are scheduled, again taking the weights of those remaining queues into account. The available bandwidth is used efficiently, since as long as there is a packet to be sent, it will be sent.

The weights of a round robin scheduler determine the ratio among the amount of packets to be sent. The actually used bandwidth of the traffic types might differ as the traffic in different queues might have different average packet sizes.

As an example one queue might contain video data whereas another queue might contain traffic due to file downloads. It can be expected that the packet sizes of the file download traffic are larger on average than those of the video streams. Exact bandwidth ratios (in terms of bits/s) are difficult to realize.

8.4.7 Weighted Fair Queuing

A weighted fair queuing scheduler serves the queues by taking the packet sizes of the scheduled packets into account. A WFQ actually achieves a bandwidth ratio according to the configured weights. Weighted fair scheduling is a theoretical concept, but there are efficient implementations that achieve reasonably good approximations.

A weighted fair scheduler is able to provide bandwidth guarantees for the different traffic classes. Assume that the overall shaping rate is 50Mbps and the weights of three queues are 6, 3, and 1 respectively. Even in case of congestion each queue will be served according to its weight, meaning that traffic of the three queues will have a guaranteed bandwidth of 30 Mbps, 15 Mbps, and 5 Mbps, respectively. If one of the queues runs empty, the other queues are able to use the resource (work conserving operation).

8.4.8 Combined Schedulers

Each of the scheduler types has its advantages. Often a combination of a strict priority and either a weighted round robin or a weighted fair queuing scheduler is available. The use of strict priority schedulers allows the latency of traffic e.g. for voice, to be reduced.

To avoid the situation where this traffic uses all of the resources, admission control and/or policing is used. This way, a limit is imposed to the amount of this traffic. The remaining traffic types are served from the queues with weighted round robin or weighed fair queueing. This provides bandwidth guarantees to these remaining traffic types.

A combined scheduler with a single strict priority queue, and additional queues for weighted round robin or weighted fair queueing is shown in Figure 8.10.

Whenever there is a packet in Queue 1, Queue 1 is served. The remaining queues are served corresponding to their scheduling weights w_1 to w_5 . Due to the policing of the traffic from Queue 1, it is possible to avoid starvation of the remaining queues, thereby providing a bandwidth guarantee for Queues 2 to 6.

For the use of Queue 1, as an example consider a first mile access from a base station by a microwave radio, supporting 10 Mbit/s bandwidth. The scheduler at the egress of a base station, can now serve packets from its strict priority queue with the whole interface rate (10 Mbit/s), provided there are packets in Q1. Typically the rate to the Q1 would be limited by both admission control and/or policing, or by simply knowing a priori that the rate will not exceed a certain maximum. For example, voice bearers are subject to admission control, so voice can be limited to a certain rate. In the case of 3G, common channels and SRB DCHs (both used for signaling) may use Q1. However, the bandwidth these signaling messages need is modest. Even more, the maximum bandwidth can be estimated, e.g. based on the capacities these channels occupy in the air interface. So the capacity needed is known. Additionally policing can be used to limit the rate.

Q1 provides an absolute guarantee in a sense that it is served any time there is a packet waiting. This is independent on the situation on other queues. Even if all other queues would contain a large amount of packets, Q1 continues to be served until it is empty.

The other queues, Q2 to Q6, are served with a relative priority. They may be starved if Q1 occupies all bandwidth. This is why policing/admission control is important for Q1. Q2 to Q6 use the remaining bandwidth. It is also used efficiently as any remaining capacity can be used. If there are no packets in Q1, a packet from Q2, Q3, Q4, Q5 or Q6 can be served.

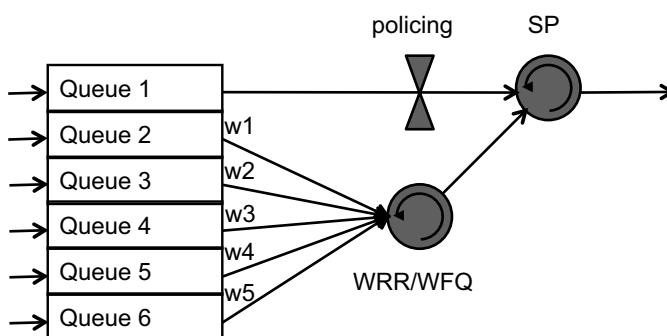


Figure 8.10 Combined scheduler.

Assume now Q1 is limited to 2 Mbit/s in our example. The bandwidth available for packets on Q2 to Q6 is between 8 Mbit/s and 10 Mbit/s, and this bandwidth is then shared according to the configured weights. Assuming that the weights are set in a decreasing order, Q6 has the smallest share of bandwidth, and will suffer first in case of congestion. However, Q6 still has its share.

Setting the weights is not necessarily trivial. Ideally, weights should be set according to the actual traffic amounts. As the traffic mix in practice varies, it is not possible to set the weights so that it would optimally match the traffic mix at every instant. Maintaining and managing scheduling parameters with a fine granularity in each of the backhaul nodes would be a difficult task. In general this is not needed either. As long as services can be supported with the QoS specified, it is not critical what the actual share of bandwidth utilized per each queue is.³ Backhaul complexity increases when the amount of queues and weights grows. This easily outweighs the achievable benefits.

The important target is that all of the traffic types in Q2 to Q6 are served according to their requirements. However, if the traffic mix changes permanently it may be a reason to modify the weights. Otherwise the basic setting should be simple and robust so that typical traffic mixes and patterns can be supported without modification.

8.4.9 Buffering

When the arrival rate to the egress queue exceeds the service rate, packets have to be either buffered or dropped. For short bursts, buffering helps to reduce packet loss. What is the size of buffers that is needed?

As buffering reduces loss, one might first consider long buffers as beneficial. However this depends on the traffic type. An example is voice traffic. For voice, quality suffers more from many delayed packets than from a few occasionally dropped packets. As the overall end-to-end delay should be in the order of 150 ms, the buffer lengths for voice traffic need to hold only a few 10 ms of traffic.

As another example, consider file download, using TCP as the transport protocol. TCP suffers from packet drop: dropped packets manifest congestion and TCP acts accordingly. Lost packets also have to be retransmitted. Longer buffers are useful for this type of traffic.

8.4.10 Tail Drop

Even if large buffers are used, they may fill up completely until there is no more memory to buffer additional traffic. A simple strategy is to discard the packet. A disadvantage of this is that several consecutive packets of one traffic flow are dropped. In case this is a TCP flow, it will trigger a slow start. Bandwidth of this flow will drop drastically and only recover slowly. Relying only on tail drop means that traffic is reduced only when the buffer is already full.

³ For planning purposes, monitoring of the actual bandwidth used per traffic class in the backhaul may reveal useful information. It clearly can be a trigger for modifying configuration of the weights. Often, it may as well indicate a shortage of capacity, which is addressed by expanding capacity in the backhaul network.

8.4.11 Active Queue Management

Active queue management tries to avoid the disadvantages of tail-drop. It intends actively to keep the queue within a certain limit of its buffer length. The best known algorithm is Random Early Discard (RED). When a certain percentage of the queue is filled up, RED will randomly discard packets from the queue. The probability of discarding packets increases with the fill level of the queue.

As packets are discarded randomly, the probability of hitting several consecutive packets of one flow is rather low. Nevertheless, TCP will recognize that packets are missing and will reduce the used bandwidth. Thereby a rather steady fill level of the queues can be achieved.

Weighted RED is a variant of RED where the probability of packets to be discarded depends also on other factors, such as their DSCP. This allows drop precedences to be implemented, e.g., to drop packets marked with DSCP 12 (AF12) before packets marked with DSCP 10 (AF11).

RED applies to a traffic mix where a large portion of the traffic is based on TCP. For other types of traffic such as voice or video streams it is less beneficial. For some types of mobile backhaul traffic it does not have an effect. As an example, a 3G radio bearer with an RLC acknowledged mode compensates for a dropped packet by retransmitting it on the RLC layer. The end user TCP flow does not see the dropped packet, since RLC retransmissions hide the packet loss.

8.4.12 Shaping

Schedulers and queue management mechanisms have an impact only when there is more traffic than can be handled. In some cases it is simply the bandwidth of physical links that is limited. The available physical link bandwidth may be less than the port rate. As an example, consider 1 Gbit/s Ethernet port connected to a Microwave Radio link of 10 Mbit/s. A similar case exists with the use of Ethernet services, and the ingress policers of EVCs.

To avoid dropping of packets by the policer, which might not take the QoS marking of traffic types into account, the traffic delivered to an EVC should be shaped to a rate defined by the EVC. With the microwave radio link example, it is also useful to shape the traffic at the source to match the physical link rate.

In order to take QoS marking into account, typically a shaper and a scheduler with its corresponding queues are combined. The shaper ensures that the defined bandwidth is not exceeded. The scheduler decides which of the buffered packets is sent first.

In the case of 3G, the RNC may shape the downlink traffic to known bandwidth limitations towards each of the NodeBs. Typically this is the bandwidth supported by the NodeB first mile access link. The first mile link may be e.g. a microwave radio hop, or an EVC, but also a TDM line.

The RNC physical interface is shared among several NodeBs. The combined aggregate traffic can be shaped additionally to the capacity of the physical link connecting the RNC to the backhaul network. This results in hierarchical shaping: Traffic is shaped according to the aggregate rate (NodeBs combined), but also to the bandwidth separately defined for each of the NodeBs.

8.5 2G

8.5.1 Native PCM-Based Abis

In 2G, BSC is managing the radio resources, and schedules traffic over the Abis interface. With native Abis (PCM-based), statistical multiplexing is not possible. Air interface timeslots directly map to the Abis timeslots. When a timeslot is allocated, there should not be an impact from the TDM-based Abis to the end user perceived QoS. TDM impairments, such as frame slips, may occur.

8.5.2 Abis Over Pseudowire

When Abis is emulated over a packet based network (MPLS pseudowires), Abis traffic is encapsulated into IP packets and further labelled by MPLS. In the mobile backhaul, Abis pseudowire is then subject to queuing and scheduling and other QoS related features of switches and routers. The pseudowire needs to be mapped to a suitable per-hop-behaviour in order not to compromise 2G service quality.

Assuming structure-agnostic emulation, a complete TDM frame is carried within the pseudowire, and there is no differentiation according to Abis traffic types. When Abis traffic is mapped to the pseudowire, headers add overhead and increase capacity needs. A trade-off can be made between this overhead and packetization delay, by configuring the amount of TDM frames that are collected into a single packet. While doing this, some additional delay is introduced due to the packetization.

Abis pseudowire is delay critical, since 2G radio network layer channels are scheduled by the BSC. There is not a tolerance for loss either, due to the GPRS based data services which are carried within the same pseudowire. GPRS includes a retransmission capability at the LLC layer between the SGSN and the MS. Packet loss on Abis leads to retransmissions. These quickly reduce throughput and increase latency for the GPRS service.

The Abis specification (TDM-based) includes a value for the maximum delay over the Abis. However, the specification is dated. In practice, maximum delay supported depends on the implementation. Delay variation has to be considered as well. The requirements originate on one side of the Abis implementation (in the BSC and in the BTS), and of the end-to-end delay budget; how much can be allocated to Abis.

The pseudowire needs to have the QoS marked so that the traffic receives a suitable treatment within the mobile backhaul. Marking of the IP packet/Ethernet frame needs to be done in the element originating the pseudowire.

8.5.3 Abis Example

In this example, a standard compliant TDM-Abis is mapped into a pseudowire. IP Abis (vendor specific) allows more alternatives if traffic types can be separated. Native IP based Abis ('IP Abis') is not standardized.

Abis pseudowires have stringent delay requirements due to the nature of the Abis interface. Pseudowire emulation requires the same constant bandwidth independent of the amount of actual traffic. All TDM frames are carried over the network, whether the timeslots are used or not. There is no differentiation between traffic types.

The packets of a pseudowire are marked with DSCP 46 and treated as EF to satisfy the requirements. The backhaul has to be dimensioned to provide sufficient capacity for this traffic. Statistical multiplexing or overbooking cannot be used as the bandwidth per BTS is static.

With a vendor specific IP Abis, different traffic types can be marked with different DSCPs, allowing for differentiation in the mobile backhaul. This, as said, is implementation specific as IP Abis is not standardized so the marking possibilities also depend on the implementation.

The used bandwidth in the IP Abis case depends on the actual amount of traffic to be carried. Statistical multiplexing becomes possible. Bandwidth utilization on the backhaul is improved. DSCP marking and corresponding treatment of packets has to ensure that the system remains operational and the QoS for the end-user is provided even in cases when congestion occurs.

8.6 3G/HSPA

8.6.1 Bearers and Their Attributes

End-to-end service in 3G, with the related bearers is shown in Figure 8.11. The attributes of the UMTS bearers are listed in Table 8.5.

Core network sets up the radio access bearers, and during the set-up procedure by RANAP (Radio Access Network Application Protocol), QoS parameters are signalled to the Radio network (to the RNC). These parameters indicate the required QoS for the new RAB.

The list of attributes in Table 8.5 is quite long. For HSPA, the system is simplified by specifying a scheduling priority indicator (SPI), which is carried over the Iub to the NodeB-SPI is used in the air interface scheduling, and also for mapping of the HSPA traffic into transport DSCPs.

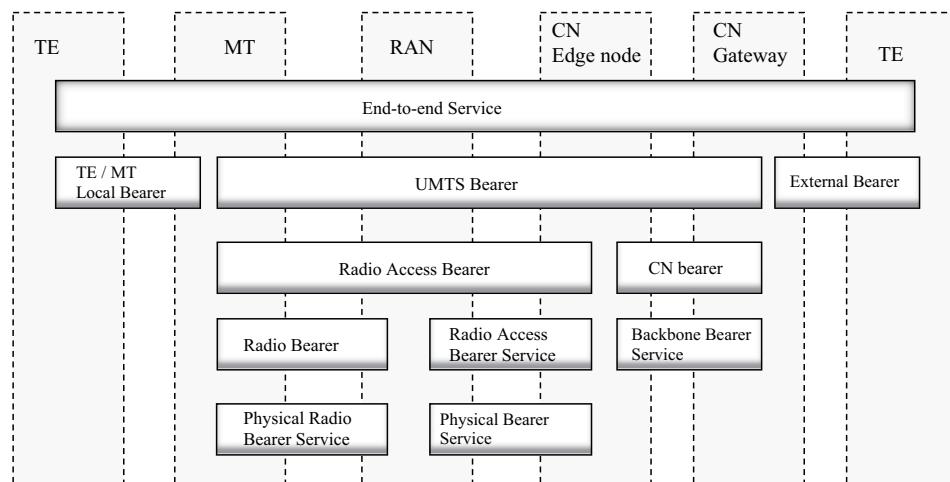


Figure 8.11 3G bearers [35].

Table 8.5 UMTS bearer service attributes [35].

Attribute	Description
Residual bit error rate (BER)	Indicates the undetected BER in the delivered SDUs. If no error detection is requested, residual BER indicates the BER in the delivered SDUs.
SDU format information	Indicates the exact size of SDUs
Sdu error rate	Indicates the fraction of SDUs lost or detected as erroneous.
Delivery of erroneous SDUs	Indicates whether erroneous SDUs are delivered, or discarded.
Max SDU size	Maximum SDU size for which the network shall satisfy the negotiated QoS.
Delivery order	Defines whether the UMTS bearer provides in-sequence SDU delivery, or not.
Transfer delay	Maximum delay (95th percentile of the distribution of delay) for all delivered SDUs during the bearer service.
Traffic class	Defines the type of application for which the bearer service is optimized. Values: conversational/streaming/interactive/background
Traffic handling priority (THP)	Relative importance of SDUS compared to SDUs of other bearers within interactive class. Values: 1, 2 or 3. Used in scheduling, as an alternative to absolute guarantees.
Allocation and retention priority (ARP)	Priority of the bearer, for admission control and pre-emption. Based on subscription. Values 1,2, .., 15.
Maximum bit rate	Maximum of bits delivered in a period of time. Upper limit a user or a application can accept or provide.
Guaranteed bit rate	Number of bits guaranteed to be delivered in a period of time. Service attributes (e.g. delay) are guaranteed up to the GBR.
Source statistics descriptor	Defines specific characteristics of the source of submitted SDUs (e.g. speech)
Signalling indication	Indicates the signalling nature of submitted SDUs, only for interactive class.

8.6.2 Iub

Channels that are carried over the Iub are shown in Figure 8.12.

In addition to user channels (DCH, E-DCH, HS-DSCH), common channels (RACH, FACH, BCH, PCH), radio network signaling (NBAP), O&M and synchronization is needed over the Iub. Radio network signaling also is carried in SRBs which are DCHs.

For the transport network itself, control traffic like routing protocols and Ethernet layer control protocols, may exist.

User data can be further classified, starting from a division into CS/PS domains and traffic class. THP and ARP parameters can be taken into account, as well as SPI for HSPA channels. SPI information element is included into the NBAP signaling. SPI is an integer value between 0 (Lowest priority) and 15 (Highest priority), and indicates the relative priority of the HSDPA or HSUPA data frames. SPI is set by the RNC, and used in the air interface scheduling by the NodeB.

Figure 8.13 shows the radio network traffic types divided into HSDPA, HSUPA, Rel-99 PS and CS domains, and into common channels. In the mobile backhaul from the RNC to the

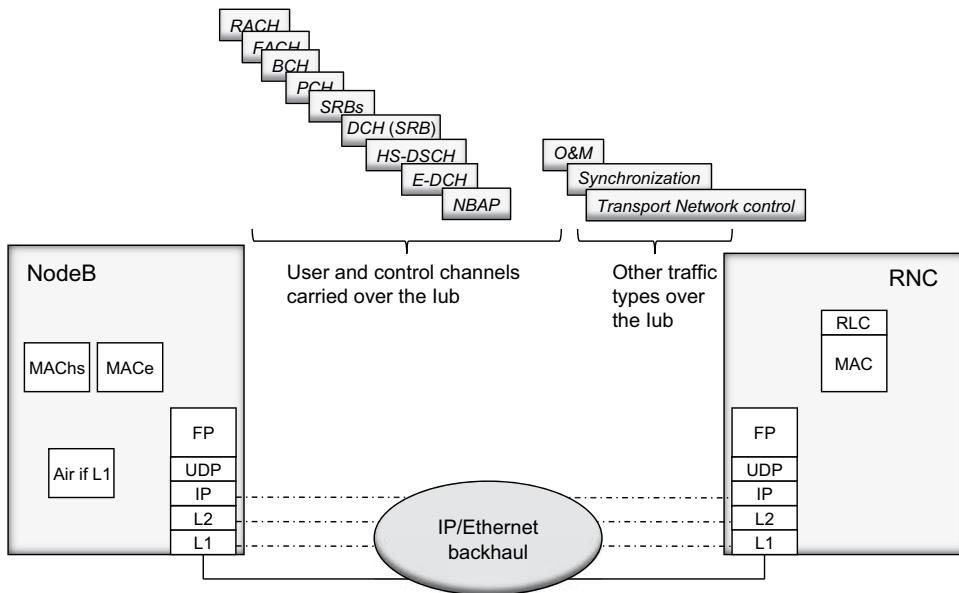


Figure 8.12 Traffic channels over the Iub.

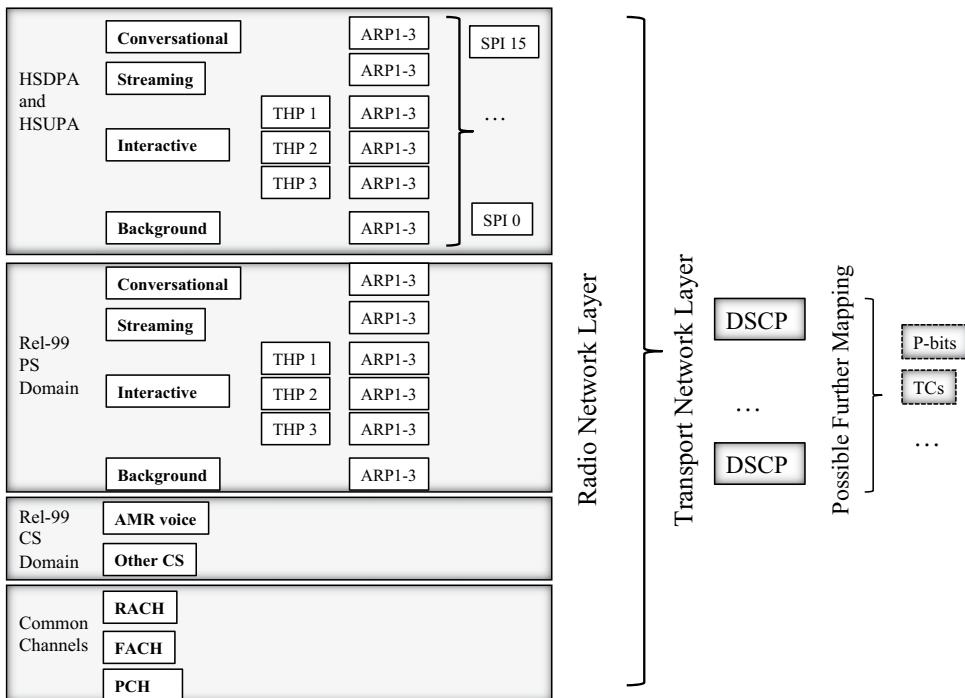


Figure 8.13 Example of mapping between radio and transport QoS.

NodeB, the received parameters are used as an input to mark the Diff Serv code point field in the IP packet in the RNC.

Additionally, SRB DCHs and NBAP signaling channels need to be mapped, as well as O&M, synchronization, and possible routing protocols and other IP control plane packets. Furthermore, the Diff Serv code point may be mapped into the Ethernet 802.1Q p-bits field, when Ethernet is used at the L2.

The IP layer QoS information in the Diff Serv field, derived as explained, may then be used in the ingress policing and egress scheduling/shaping functions in the mobile backhaul network elements. Allocation and retention priority (ARP) may be used by the RNC e.g. in deciding which bearers should be released in case of resource shortage. Guaranteed bit rate information can be used in admission control and related resource reservations in the mobile system, including transport interfaces, by the mobile network elements.

The bearer service attributes are not visible for the mobile backhaul directly. The backhaul elements identify the intended QoS from the packet marking done by the RNC in the downlink direction. In the uplink direction, NodeB marks packets similarly for transport over the backhaul network.

With HSPA, Frame protocol layer now includes a flow control function (credit allocation) between the NodeB and RNC. In the downlink (HSDPA) direction, NodeB allocates credits to the RNC, according to the air interface resource availability. When air interface has room for more HSDPA packets, the NodeB gives RNC further credits, which are consumed when the packets have been received over the Iub interface at the NodeB. The algorithms for credit allocation are not defined by 3GPP.

FP Flow control is a way of controlling the amount of information that needs to be transmitted over the Iub interface, based on the air interface resource situation. Another feature, HSDPA congestion control, additionally has the means of detecting and informing that there is congestion over the Iub interface. HSDPA congestion control detects both delay build-up and packet loss. HSDPA and HSUPA congestion control are discussed in a dedicated section.

In transport the delay increase is attributable to the buffer levels of the transport elements, typically to the egress buffer. Within mobile backhaul, the first buffers on the transport layers exist at the mobile elements, NodeB and RNC. Depending on the mobile backhaul type, further buffers typically are found in all packet network elements, IP routers, multilayer switches, Ethernet switches, and congestion may occur at any of these elements.

8.6.3 *Iub Example*

An example QoS mapping for an IP based Iub interface is considered in this section, by mapping control plane, network control, user plane, O&M and synchronization into the DSCPs for the transport service.

8.6.3.1 **Control Plane (NBAP, Common Channels, SRBs)**

Radio network signaling does not require a constant bandwidth. However, failing to service radio network signaling on common channels and on SRBs causes issues at the

radio network level. Delays on radio network signaling impact call set-up/bearer set-up times and operation of handovers, the success rates for these, etc. These are difficult to trace to the backhaul, as radio network impairments are another potential cause. For mapping the QoS in the mobile backhaul, this dependency is a specific consideration.

The radio network control plane on NBAP is marked with DSCP 34 (AF41). The service for common channels (RACH, FACH, PCH,...) and signaling (SRBs) may be marked similarly with DSCP 34 (AF41), assuming however that adequate transport service is provided by the AF41 PHB.

For a stronger guarantee especially for common channels and SRBs, DSCP 46 (EF) can be used. This ensures that radio network signaling receives a guaranteed throughput. By this, the risk of compromising radio network performance due to delays and packet loss in the transport layer service is minimized. Traffic volume on SRBs and common channels is typically not high.

8.6.3.2 Network Control

Network control (e.g. routing protocols) is marked with a DSCP value of 48 (CS 6). This helps to ensure that the transport service and the backhaul network itself stays operational, can recover rapidly from failures, and maintain the transport service for the radio network layer.

8.6.3.3 User Plane

In the user plane, bearers are either Rel-99 DCH or HS-DSCH (HSDPA) and E-DCH (HSUPA) bearers. DCH in general has more stringent requirements on delay and jitter than HSPA and HSUPA. This is due to the network architecture and the location of radio network layer protocols.

However, HSDPA also includes a mac scheduling function in the RNC. HSUPA supports macrodiversity combining and fast power control in the RNC. Data from a UE is received via several NodeBs and should arrive at the same time to the RNC. This means in practice, that user traffic on HSPA still has radio network originated requirements for the Iub backhaul, even though the fast scheduling (mac-hs, mac-e) function is located in the NodeB.

Often the data volume in the uplink is lower than in the downlink. As backhaul typically provides symmetrical bandwidth in downlink and uplink directions, congestion rarely occurs in the uplink.

As presented in Figure 8.13, user traffic may be divided into a CS and a PS domain. Further, the PS domain is categorized into Conversational/Streaming/Interactive/Background. A starting point is a division into real-time and non real-time traffic types. The amount of data traffic is growing faster than the amount of voice traffic, therefore the amount of NRT traffic dominates the amount of RT traffic.

For the non-real time traffic, new WCDMA mobile phones typically provide the ability to use HSPA. Therefore it can be expected that more and more NRT traffic will use HSPA radio bearers instead of NRT DCHs. Also, it can be assumed that a substantial amount of background NRT traffic exists on HSPA bearers. This elastic traffic can be used to reduce bandwidth requirements on the transport network by statistical multiplexing.

In this example user plane is mapped to three different treatment aggregates:

- All real-time traffic, whether Rel99 or HSPA, is marked with DSCP 46 (EF).
- Non real-time traffic on Rel 99 dedicated channels is marked with DSCP 26 (AF31).
- Non real-time traffic on HSPA is marked with DSCP 0 (BE).

8.6.3.4 O&M

O&M traffic is marked with DSCP 16 (CS2). A special characteristic of the mobile network O&M traffic is that occasionally the traffic volume is high, while on average it is much lower. Consider a SW download as an example of high data transfer needs. On average the amount of traffic is low. As alarms are transferred via the O&M channel, a minimum capacity is needed for the O&M, so that this information can be forwarded.

8.6.3.5 Synchronization (by Packet)

How synchronization is obtained, is subject to implementation. Synchronization was discussed in Chapter 6. Even with a standardized approach for packet timing, e.g. IEEE1588v2, algorithms are implementation specific. These may differ with respect to their requirements for the backhaul. A constant bit stream is served by an EF class.

A specific issue is delay jumps. If delay changes abruptly (which could happen in a recovery after a failure), algorithms may react differently, again depending on the implementation.

As an example, assuming IEEE1588v2, the traffic can be marked with DSCP46 (EF) to treat it with a deterministic throughput and low delay.

Table 8.6 presents this as a summary.

8.6.3.6 Summary

In the transport network each treatment aggregate could be mapped to a separate queue, based on the DSCP marking of the packets. If four or less queues are available, then network control

Table 8.6 DSCP marking for the Iub.

	Traffic type	DSCP
Control plane	NBAP	34 (AF41)
	Common channels	34 (AF41) ^a
	SRBs	34 (AF41) ^a
Network control		48 (CS 6)
User plane	real-time	46 (EF)
	HSPA/HSUPA	46 (EF)
O&M	non real-time	26 (AF31)
Synchronization	DCH	0 (BE)
	HSPA/HSUPA	16 (CS2) ^b
		46 (EF) ^c

^a Common channels and SRBs alternatively as DSCP 46 (EF)

^b If CS2 is not used, AF21 is an alternative

^c Depending on packet synchronization requirements.

and real-time traffic could be mapped to the same queue. O&M traffic could be mapped either to the same queue as non real-time Rel99 traffic or as to the queue with non real-time HSPA traffic. Either strict priority schedulers or WRR/WFQ schedulers can be used.

A further topic is in mapping different classes of users in the radio network and in the backhaul. Approaches for the backhaul are easily vendor specific, as there is no guidance by 3GPP. A starting point is the radio network layer ARP parameter.

8.6.4 Congestion Control in MBH

Despite the capabilities of the backhaul transport protocols (resilience, high data rate, low latency, QoS differentiation, etc.), transient congestion may occur due to the capacity limited first mile links such as microwave radio or due to the overbooking of the high capacity aggregation links.

During congestion, connections experience increased delay, reduced throughput and packet drops. TCP is the dominant transport protocol used by the majority of data applications; it has its own efficient congestion control mechanism that reacts by reducing the rate of the connections and by retransmitting the data that is assumed to be lost. Adaptive video streaming applications are also able to adapt their rate to the available bandwidth.

In full packet based systems such as LTE, these end-to-end mechanisms might be enough but not for HSPA. Transport congestion deteriorates the overall HSPA system performance as it can trigger unnecessary RLC AM (Radio Link Control Acknowledged Mode) retransmissions as dropped packets are retransmitted by the RLC AM entity. Moreover, in cases where the transport is shared by the LTE and HSPA traffic, congestion may cause fairness problems as HSPA traffic is not TCP friendly and thus LTE connections can starve.

Although the functionality of the HSPA systems has been extended by 3GPP with a means of detecting congestion at the RNC and Node B and with the capability of notifying the source node that congestion has occurred, the congestion control algorithm itself is not specified. Detection and indication are possible via additional IEs (Information Elements) included in the HS-DSCH and E-DCH frame protocols' data frame headers or via special control frames. LTE has no similar standardized functionality.

8.6.5 Congestion Control in HSPA Systems

Data applications such as file transfer or web browsing are TCP based, which provides error free delivery of the data. The TCP congestion control mechanism was developed based on the assumption that the bit errors caused by the physical media imperfections are very unlikely thus the transport congestion is the only reason for data loss in the system. This is true for wired systems but not for radio links, where bit and block errors are frequent. The TCP mechanisms do not handle data loss well due to air interface imperfections. In legacy WCDMA systems, the user plane traffic between the RNC and Node B is mostly carried over dedicated transport channels via the Iub and Iur interfaces.

To overcome the negative impact of the air interface errors on the data traffic, NRT (non-real-time) bearers (carrying data traffic) are handled by the RLC AM entities (one is located in the RNC and the other is in the UE) that have an ARQ mechanism responsible for retransmitting the missing (lost) data in case negative acknowledgement (NACK) is received

from the peer RLC AM entity. As the RLC AM entities are located in the RNC and the UE, there is a significant latency (referred to as Layer 2 Round Trip Time) in the reaction time to air interface errors or packet drops on the transport network that has a negative impact on the TCP performance.

Transport congestion is hidden from the TCP as the RLC AM entity retransmits the missing data. In-sequence delivery, another feature of the RLC AM, prevents the RLC AM entity at the receiver side from delivering any correctly received consecutive TCP segment to the upper layer before the missing data is received or the incorrectly received TCP segment is discarded. That is, in case the maximum number of allowed retransmissions is reached, the corresponding TCP segment (that is, the RLC AM SDU) is discarded and the RLC continues with a consecutive TCP segment. If that is successful, a duplicate ACK is sent to the TCP source that might already be in slow start due to timeout caused by the long latency of the RLC retransmission mechanism. In order to improve the system performance, i.e., to minimize the amount of retransmissions, the transport should be dimensioned so that the likelihood of transient congestion is low. This results in an over-dimensioned and costly transport network.

HSPA has introduced additional system features (fast scheduling, adaptive coding and modulation, HARQ, HSDPA flow control) and Iub protocols (MAC-hs, MAC-e, MAC-es) that improve the system performance in terms of data rates and latency.

The role of the HARQ mechanism is to reduce the latency of the Layer 2 retransmissions due to air interface errors: it executes retransmissions when negative acknowledgements are received or when the acknowledgement is not received in time. This mechanism handles the air interface errors efficiently, but packet drops due to transport congestion are still handled by the RLC AM entity residing at the RNC and UE. On the other hand, the HSDPA flow control algorithm considers only the air interface (when the capacity allocations are calculated) thus it can easily overload the transport network causing congestion and eventually packet drops. This leads to performance degradation as packet drops at the transport network trigger RLC AM retransmissions.

HSDPA/HSUPA congestion control functionality was proposed by 3GPP TR25.902 with the scope to handle the efficiency problems caused by transport congestion. The approach was to reuse the existing network features and, despite the technical differences, to provide similar solutions in HSUPA and HSDPA.

The congestion control entity is located at the Node B as in both HSDPA and HSUPA it controls the rate of the connections either via capacity allocations sent to the SRNC (HSDPA) or via grants issued to the UEs (HSUPA).

The congestion detection is based on additional information attached to each frame sent in UL or DL: a reference time that indicates the time when the frame was sent and a sequence number. The reference time is used in order to detect delay build-up whereas the sequence number is used in order to detect frame loss. Delay build-up indicates that the frames are queued in transport buffers due to overload whereas the frame loss indicates that the frames are dropped due to overload. The role and functionality of the Iub protocols during a web page download over HSDPA is shown in Figure 8.14.

8.6.6 HSDPA Congestion Control

The HSDPA congestion control functionality is located at the Node B as a complementary mechanism to the already existing HSDPA flow control functionality of the MAC-hs layer that

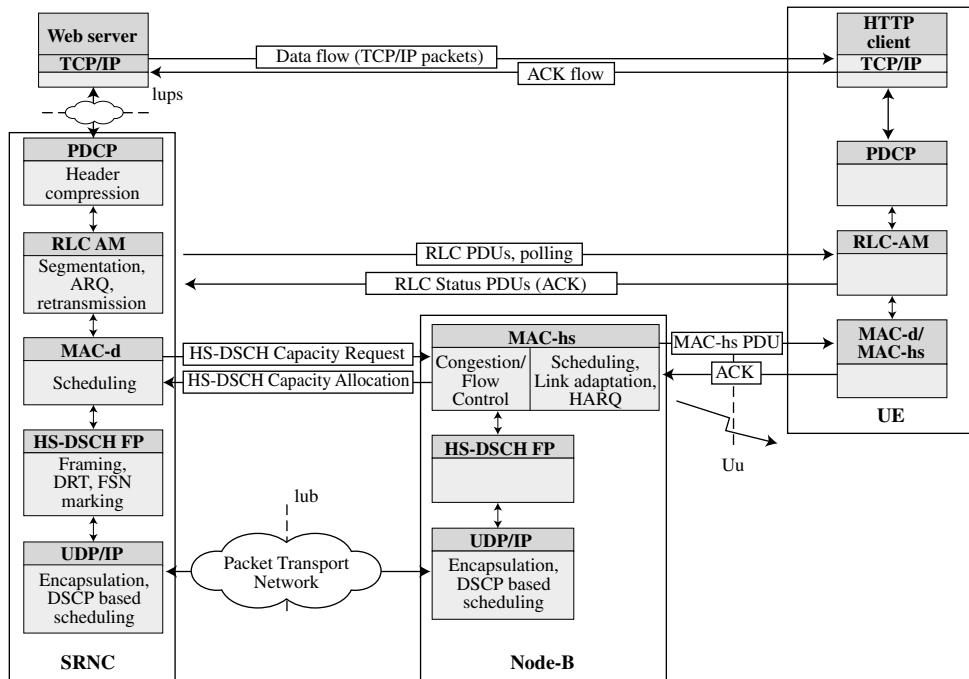


Figure 8.14 The role of the Iub protocols during a web page download over HSDPA.

defines the rate of the MAC-d flows via capacity allocation messages sent to the SRNC. The reference time and sequence number enables the Node B to detect downlink transport congestion, based on which information the congestion control mechanism calculates the amount of resources granted to the MAC-d flows.

The rate of the MAC-d flows is controlled via the *HS-DSCH FP Capacity Allocation Control Frames* sent to the SRNC through the Iub interface. There are two types of control frames: HS-DSCH Capacity Allocation Control Frame Type 1 and Type 2. Type 2 was introduced with the Flexible RLC, whereas Type 1 is used in earlier releases with fixed RLC. This message contains the allocation that the SRNC may use, i.e., the following IEs: Maximum MAC-d PDU Length Type1/Type 2, HS-DSCH Credits and HS-DSCH Interval. Additionally, the HS-DSCH Repetition Period IE (within both Type 1 and Type 2 control frames) defines the validity period of the allocation. If its value is set to zero, it means that the allocation can be repeated without a limit.

A further possibility is to inform the SRNC about the detection of congestion in the downlink direction by setting the Congestion Status bits within the Capacity Allocation Control frame. The possible values of the 2 bits carrying the congestion status are: no TNL congestion (0), TNL congestion detected by delay build-up (2), TNL congestion detected by frame loss (3), whereas value 1 is reserved for future use.

The Maximum MAC-d PDU Length IE indicates the maximum allowed PDU size from the MAC-d PDU sizes configured via RNSAP (Type 1 frames) or it is a factor in the granted amount of MAC-d PDU data the SRNC may transmit during one HS-DSCH Interval (Type 2

frames). In the latter case, the amount of data is obtained by multiplying the content of the MAC-d PDU Length Type 2 IE with the content of the HS-DSCH Credits IE.

An HS-DSCH Credits IE value set to zero means that no resources are allocated to MAC-d flows and the transmission should be stopped. In case of Type 1 frames the content of the IE indicates the amount of PDUs the SRNC is allowed to transmit during one HS-DSCH Interval, whereas in case of Type 2 frames the IE is used when the granted amount of MAC-d PDU data is calculated, as described above.

The HS-DSCH Interval IE indicates the scheduling interval of the allocated amount of data at the SRNC. The first interval starts right after the reception of the allocation, subsequent intervals start after the first interval elapsed, until the number of intervals specified by the HS-DSCH Repetition Period IE is reached.

The capacity allocation messages are sent either as a response to an HS-DSCH Capacity Request sent from the SRNC or whenever the HS-DSCH flow control or congestion control algorithm decides to reduce or increase the rate of a particular connection. The amount of allocations should be calculated so that the air interface resources are not wasted, i.e., when the packet scheduler selects a bearer/connection for scheduling, there is enough data in the connection's buffer at the Node B, but on the other hand this buffer is not overloaded and at the same time the transport congestion is reduced.

At each HS-DSCH scheduling interval, the SRNC schedules the amount of MAC-d PDU data specified in the last HS-DSCH Capacity Allocation control frame. The MAC-d PDUs are sent in HS-DSCH FP data frames (Type 1 and Type 2, respectively) that contain the FSN and DRT information elements and the buffers' status report (size) included into their headers. The FSN is incremented for each HS-DSCH data frame of a given MAC-d flow. The length of the IE is four bits where the value zero is not used at wraparound. The DRT (length 16 bits) is a 40960 counter with 1 ms resolution. In addition the FP data frame header contains an FSN/DRT reset bit; when it is set, the Node B should reset any state of congestion estimation based on earlier FSNs and DRTs.

HS-DSCH Capacity Request control messages are sent by the SRNC whenever the SNRC considers that the buffer status needs an increased reporting frequency, i.e., to signal an event such as data discard or arrival, compared to the buffer reporting via data frames. The request can be sent in order to trigger reconsideration of the allocation size, for example in the case of zero allocation when the SRNC side buffer content is increasing or when after a longer idle period new data has arrived at the SRNC.

An example of HSDPA congestion control architecture is provided in Figure 8.15.

8.6.7 HSUPA Congestion Control

Similarly to the HSDPA, the HSUPA Congestion Control functionality is located in the Node B, whereas the detection functionality is located in the SRNC. The congestion information is signaled to the Node B with the E-DCH FP Congestion Indication Control Frame.

The MAC-e (air interface) packet scheduler at the Node B calculates the allocation (serving grant) sent to the UEs based on the congestion information received from the SRNC. This allocation defines when and with which bit rate the UEs are allowed to transmit in the cell. The received data is assembled into E-DCH Data Frames with the following congestion control related IEs: frame sequence number (FSN), connection frame number (CFN) and sub-frame

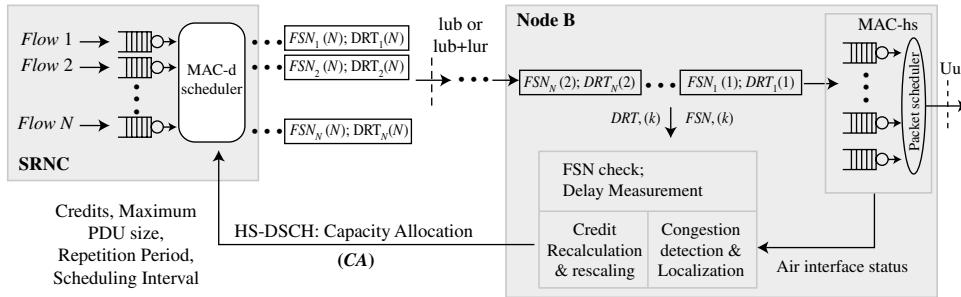


Figure 8.15 An example HSDPA congestion control architecture.

number. The FSN IE is incremented by one (modulo 16) for each transmitted data frame. The CFN indicates the radio frame when the HARQ process correctly decoded the data. In the case of 10 ms HSUPA TTI, the delay is calculated by the SRNC based on the CFN. When the HSUPA TTI is 2 ms, the same CFN is shared among five sub-frames; therefore the sub-frame number is used to calculate the delay variation between consecutive frames.

In order to detect congestion, the SRNC analyzes the content of the relevant IEs of the received E-DCH FP Data Frames: a packet discard on the transport is detected by following the value of the FSN IE of the consecutive data frames of the same connection, whereas delay build-up is detected by calculating the delay or delay variation based on the content of the CFN or sub-frame IEs.

The congestion status is signaled to the Node B via E-DCH FP Congestion Indication Control Frames that can indicate that the congestion is over (or there is no congestion on the transport network) or that there is congestion. The reason for congestion, i.e., whether delay build-up or frame loss was detected is signaled with different reason codes included in the congestion status messages.

The Node B reduces at least the rate of the MAC-d flow for which the congestion indication control frame was received. After the congestion is over, the rate of the MAC-d flows is gradually increased. An example of HSUPA congestion control architecture is provided in Figure 8.16.

8.6.8 Co-existence of Radio Networks

In most of the cases LTE systems will be deployed in order to provide additional radio access possibilities to the already existing WCDMA/HSDPA systems thereby creating

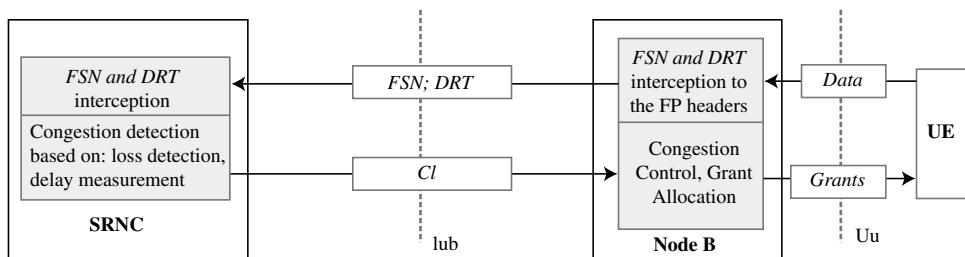


Figure 8.16 An example HSUPA congestion control architecture.

heterogeneous coverage over the same areas and providing the possibility of offloading the data traffic from the WCDMA/HSDPA systems. Heterogeneous radio access networks will use the same packet based transport infrastructure, i.e., limited transport resources will be shared by WCDMA/HSDPA and LTE systems. Transport links carrying the traffic of the co-existing radio access systems might experience transient congestion.

The HSDPA Congestion Control located in the Node B will detect the congestion based on the delay measurements and sequence number of the HS-DSCH transport frames and it will take action in order to solve the congestion by reducing the rate of the connections. The feedback loop of the HSDPA congestion control is short. In most of the cases, the transport congestion will be transparent to the TCP congestion control mechanism.

In case of LTE, transient congestion is detected by the TCP congestion control mechanism. The rate of the connections is controlled by the round trip time. In the case of single packet drops, the TCP fast retransmit and fast recovery algorithms retransmit the missing TCP segment and halve the rate of the connections. In cases of severe congestion, when no acknowledgement is received before the TCP timeout timer expires, the TCP source transits into slow start phase and retransmits the data for which an acknowledgement was not received.

As the HSDPA congestion control feedback loop is shorter than the end-to-end TCP flow/congestion control loop, the rate of the connections using HSDPA is reduced first. The unused bandwidth will be taken by TCP connections over LTE. This will continue until the total starvation of the HSDPA connections.

When the HSDPA congestion control is not applied, the rate of the HSDPA connections is controlled by the HSDPA flow control mechanism that aims to achieve optimal air interface usage without considering the transport congestion. The HSDPA data traffic in this case is not TCP friendly as upon congestion the data rate of the MAC-d flows is not reduced but the lost data is retransmitted by the RLC AM entities. In cases where the congested link is shared by HSPA and LTE, this will cause the starvation of the TCP connections over LTE. Moreover, turning the HSDPA congestion control off in homogeneous networks or when there is only HSDPA traffic in the system can lead to inefficient resource utilization due to RLC AM retransmissions even if there are no other connections but HSDPA.

As the topic is not addressed in 3GPP, solving the co-existence efficiently becomes an implementation issue, and different solutions exist.

8.7 LTE

8.7.1 *QoS Architecture*

With LTE, the quality of service concept is simpler than that of the 3G system. A default bearer supports IP connectivity for basically any service. When the indicated quality of service need is different than that which the default bearer offers, a dedicated bearer is set up – provided that the system has resources for the new bearer.

A traffic flow template (TFT) defines which user flows can be mapped to which bearers. User flows that have similar requirements, can use the same bearer. User flows with differing requirements, will be mapped to different EPS bearers. Traffic flow template exists in the UE for the uplink, and in PDN GW for the downlink.

In addition to traffic flow template, service data flow (SDF) template is defined. This is related to Policy and Charging Control (PCC) rules. SDF may e.g. provide a finer granularity than TFT. SDF and TFT may also in practice be identical. SDF only exists in the core network.

Traffic mapped to the same EPS bearer receives the same QoS treatment. QoS functions in the LTE system include scheduling and queue management policy, RLC configurations (e.g. RLC mode) and a shaping policy.

With LTE, Quality of service Class Indicator (QCI), as the name implies, is one key indication for the QoS of the bearer. In transport Differentiated services codepoint (DSCPs) are available for carrying this information in the IP packets, based on the QCI.

8.7.2 Packet Flows and Bearers

EPS bearer provides the packet data network (IP network) connectivity to the UE. It is comparable to the PDP context concept from 3G.

End-to-end service consists of the EPS bearer (between UE and the PGW) and of the external bearer (typically to the internet, or e.g. to an enterprise VPN). E-RAB is comparable to the RAB in 3G. An EPS bearer is realized with the radio bearer, S1 bearer, and the S5/S8 bearer. S8 interface is for roaming cases. See Figure 8.17.

The connectivity to the external network (PDN, the internet) consists of transport of traffic flow aggregates. These are defined to be collections of multiple service data flows. Further, service data flow is a set of packet flows matching a service data flow template. Service data flows are bound to the bearer (IP-CAN bearer, IP connectivity access network).

PCC rules are needed when EPS bearers are to be established. PCC rules can be statically configured in the PDN GW. Alternatively they can be provided by PCRF (Policy and Charging Rules Function). Rules may also be modified during the lifetime of the EPS bearer.

PCC rules are used to detect these service data flows. Based on the rules, parameters for charging and policy control (including QoS policy) are determined. PCC rules include the service data flow template, and this template possibly contains multiple service data flow filters. Further, the filters may be separate in both uplink and downlink directions. This is illustrated in Figure 8.18.

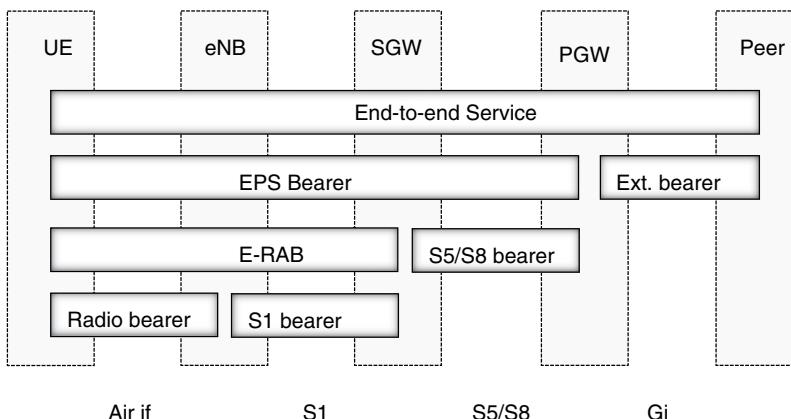


Figure 8.17 Bearers in LTE [42].

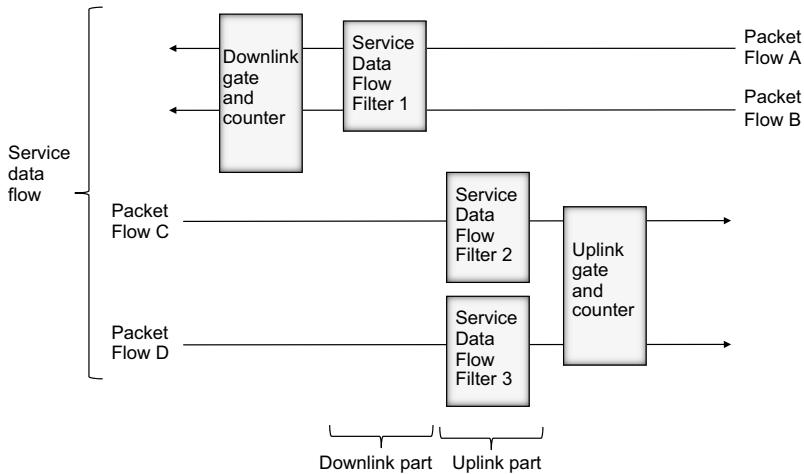


Figure 8.18 Service data flow filters and detection [36].

In practice, detection means e.g. matching with the source and destination IP addresses, source and destination port numbers, and the protocol used above IP. IP address information may contain a prefix mask, and instead of a single port number, port ranges may be defined. Additionally, the QoS marking on the IP header may also be used. It is also possible to look further into the packet, and examine the transport and application protocol layers.

Due to the detection (packet inspection) capabilities, it is possible to identify packet flows and then apply a policy for charging and QoS control. The EPS bearer, with the QoS characteristics defined, is then established accordingly. The bearer establishment may be subject to admission control on all nodes (eNodeB, SGW, PGW) and other resources (air interface, transport, processing capacity etc.).

A default EPS bearer is established for the lifetime of the PDN connectivity ('always-on connectivity'), when the UE connects to the PDN network. The default bearer is always a non-GBR bearer, which is set up by the MME. For the default bearer, the QoS parameter values of the default bearer are based on subscription data. MME queries the necessary information from the Home Subscriber Server.

Additional bearers are dedicated bearers that may be set up by either UE initiating the request, or by data arriving from the external network. UE may e.g. initiate a VoIP session that triggers setup of a dedicated EPS bearer. Dedicated bearers can be either GBR or non-GBR bearers.

If the bearer set up is due to the external network, the bearer is established via the PCRF and the PDN GW – also e.g. in the case of a VoIP session.

Gx interface is supported between PCRF and the PDN GW. In cases where the S5/S8 interface is Proxy Mobile IP based, the QoS mapping is done by the SGW, with the Gxc interface.

Each dedicated EPS bearer includes a traffic flow template, for both uplink and downlink directions. Traffic is matched against the template, and then mapped accordingly to an EPS bearer. UE maps the traffic in the uplink direction, and the PCEF (in the PDN GW) maps the traffic in the downlink direction. With EPS bearer modifications, the PDN GW delivers traffic flow related information (e.g. source and destination IP addresses, source and destination port numbers, and protocol used) to the UE, so that the UE can associate the application with the correct EPS bearer.

Table 8.7 QCIs [36].

QCI	Guarantee	Priority	Packet Delay Budget	Packet Error Loss Rate	Example Services
1	GBR	2	100 ms	10^{-2}	Conversational voice
2	GBR	4	150 ms	10^{-3}	Conversational video (live streaming)
3	GBR	3	50 ms	10^{-3}	Real-time gaming
4	GBR	5	300 ms	10^{-6}	Non-conversational video (buffered streaming)
5	Non-GBR	1	100 ms	10^{-6}	IMS signalling
6	Non-GBR	6	300 ms	10^{-6}	Video (Buffered streaming), TCP based (e.g. www, e-mail, chat, ftp, p2p file sharing, progressive video, etc)
7	Non-GBR	7	100 ms	10^{-3}	Voice, Video (live streaming), Interactive gaming
8	Non-GBR	8	300 ms	10^{-6}	Video (Buffered streaming), TCP based (e.g. www, e-mail, chat, ftp, p2p file sharing, progressive video, etc)
9	Non-GBR	9	300 ms	10^{-6}	Video (Buffered streaming), TCP based (e.g. www, e-mail, chat, ftp, p2p file sharing, progressive video, etc)

8.7.3 QoS Parameters

The QoS parameters for the EPS bearer are QCI (QoS Class Identifier), ARP (Allocation and Retention Priority), GBR (Guaranteed Bit Rate) and MBR (Maximum Bit Rate). GBR and MBR are applicable only for GBR-bearers. Per UE (aggregate of EPS bearers), additionally QoS parameters APN-AMBR (Aggregate Maximum Bit Rate) and UE-AMBR are defined.

QCI classes are standardized in 3GPP (TS23.203), and shown in Table 8.7.

For Guaranteed Bit Rate (GBR) services, dedicated resources are allocated typically with Admission Control (AC). Non-GBR services do not have a similar guarantee for resource availability.

Priority level is meant for differentiating between multiple service data flows of a single UE, and also for differentiating service data flow aggregates between UEs. Via the QCI, an SDF aggregate is associated with a packet delay bound, and a priority level. Packet delay bound is primarily intended to be used for differentiating scheduling between SDF aggregates. If the packet delay bound cannot be met for all SDF aggregates, priority level is used in guiding which SDF aggregate is scheduled first. Priority level N has preference over priority level N + 1.

Packet delay budget is interpreted as the maximum delay bound, with a confidence level of 98%. It is defined as the delay between the UE and the PCEF (Policy and Charging Enforcement Function). PCEF is a function of the PDN Gateway. Packet delay budget is the same in uplink and downlink, for a certain QCI.

Packet Error Loss Rate (PELR) specifies an upper bound for packets that are sent by the link layer (e.g. RLC) but not received at the upper layer (e.g. PDCP) at the receiver end. PELR tells the packet loss rate without congestion. PELR is used in selecting the link layer configuration (e.g. RLC acknowledged mode). For a certain QCI value, PELR is the same in uplink and downlink directions.

The PELR value also assumes that the loss in the mobile backhaul (from the base station to the PDN GW) is negligible. So, for the system to meet the PELR loss bound, mobile backhaul should not have a contribution. In non-congested cases, normal conditions and well-designed links, the packet loss ratio should also in practice be minimal in the backhaul. For the mobile backhaul not to have a contribution, the packet loss ratio caused by the backhaul, could be e.g. one decade less than the PELR loss bound.

To estimate for the backhaul induced PELR, an analysis is needed to estimate the packet loss rate when the physical layer has a certain, residual BER (Bit Error Rate). Layer-2 protocols such as the Ethernet and the PPP, include a checksum, which is calculated over the complete frame. If the checksum fails, the frame is discarded. Thus, bit errors lead to frame discards. So the physical layer Bit Error Rate (BER) value has to be considered to assure that there is no significant contribution from the backhaul.

In the mobile backhaul, congestion in the network (packet drops in the egress queues of the routers and switches) may cause further packet loss. The PELR target may be violated in these cases, if the backhaul is congested, even if the air interface has capacity and good radio conditions. Similarly, during temporary backhaul connectivity losses (e.g. due to failing links, and reconvergence of the network), packets may be lost in the mobile backhaul.

Transport level QoS marking is done based on the QCI value of the EPS bearer. eNodeB is responsible for marking in the uplink direction. SGW and PGW mark the packet in the downlink direction and in the uplink directions, similarly based on the QCI value of the EPS bearer. So, at bearer level, EPS bearer is the level of granularity of QoS, and the QCI value of the EPS bearer is further marked to the transport layer.

The bearer establishment/pre-emption priorities are identified by the ARP parameter. It is also needed in the eNodeB, so that pre-emption at the radio interface can be supported.

An example of ARP usage is given in TS23.401, where voice has a higher ARP than video. In cases of congestion, video service is dropped before the voice service. Similarly, ARP may be used to relieve network load. In the case of high traffic load due to e.g. a natural catastrophe, lower priority ARP bearers may be torn down.

GBR is the expected bit-rate provided by the EPS bearer, while the MBR may limit the bit rate with a shaping function. GBR parameter can be utilized e.g. in admission control.

APN-AMBR and UE-AMBR are both subscription parameters which are stored in the HSS. APN-AMBR limits the bit rate of all Non-GBR bearers within the APN. UE-AMBR is the upper limit for all active APN-AMBRs, up to the value of subscribed UE-AMBR. The parameter is set by the MME. So GBR bearers are not included in the UE-AMBR calculation and potential shaping.

Bit rates for GBR, MBR and AMBR, are calculated as the bit stream at S1 excluding GTP-U and IP headers. The QoS parameters have a component for both uplink and downlink directions.

8.7.4 Admission Control

In LTE, for GBR bearers, dedicated resources are allocated, e.g. with admission control. Admission control needs to take into account all functions, radio resources, hardware (processing) resources and also transport.

eNodeB is responsible for the radio admission control. Before establishment of a new GBR bearer, radio resource availability needs to be ensured, both for UL and for DL. Priority levels and QoS required for the new bearer are considered, as well as those of the already established bearers. If resources are available, the bearer is admitted. Otherwise the bearer set-up request will be rejected, unless priority levels and pre-emption indicators suggest that another bearer should be pre-empted in order to allow the new bearer to be established.

8.7.5 S1 Interface

S1 interface includes a user plane connection between the eNodeB and the SGW, a control (signalling) connection (S1-AP) between the eNodeB and the MME, O&M channel towards network management, and often also synchronization (such as IEEE 1588v2). Additionally, transport control plane, such as IP routing protocols, may be carried.

As opposed to 2G and 3G BTS access interfaces, eNodeB S1 interface does not need to transfer radio layer protocols, as those are all terminated in the eNodeB. The requirements for delay, delay variation and loss originate from the needs of the end user experience rather than from any mobile system architecture constraints.

In the user plane, mandatory E-RAB level QoS parameters are QCI, and Allocation and Retention Priority. GBR QoS information is an optional parameter (not needed for non-GBR bearers), consisting of Guaranteed and Maximum bit rates uplink and downlink. Without the GBR QoS information, GBR bearer set-up fails. For the user plane, it is mandated to support configurable DSCP marking. The input information for the marking are the QCIs and other parameters.

For user plane packet loss, the S1 interface behaves differently than packet data on the Iub (assuming acknowledged RLC mode). Lost packets on Iub are retransmitted at the RLC layer. With LTE, RLC layer terminates at the eNodeB. S1 is comparable to the Iu interface of 3G also in this respect. This means that packet loss on S1 is visible to the application layer, as there is no protocol hiding the loss by retransmissions. If the application uses reliable transmission (e.g. TCP), packet loss on S1 causes TCP layer retransmissions.

8.7.6 S1 Example

In LTE the requirements on the backhaul are driven by the applications. There are no technical requirements on delay and delay variation from a controller, since there simply are no radio controllers in an LTE network.

8.7.6.1 Control Plane (S1AP)

Radio network control plane in the case of LTE and eNodeB is again simpler than in 3G, since now the radio layer protocols are not carried over the S1. Correspondingly, cell common channels also terminate in the eNodeB, and are not a concern in the mobile backhaul.

Radio network signaling, S1-AP, is carried over the S1 logical interface. This is marked with DSCP 34 (AF41).

8.7.6.2 Network Control

Network control (e.g. routing protocols) is marked with a DSCP value of 48 (CS 6). There is no difference between radio technologies concerning the transport network control protocol marking. It is needed as CS6 for the same reasons in all radio networks.

8.7.6.3 User Plane

The different end-user applications are indicated by their Quality of Service Indicator (QCI). 3GPP has defined 9 QCIs in TS23.203. Therefore the QoS mapping for LTE in the user plane is determined mostly by a mapping of QCIs to DSCPs.

If DSCPs indicating higher drop precedences are used (such as AF32 as an example), then there are enough DSCPs so that each QCI can be mapped to a unique DSCP value. The amount of QCIs is, however, larger than the amount of PHBs. Therefore one has to expect that several QCIs will be treated similarly in the transport network. In this example only DSCPs corresponding to low drop precedence are considered.

From the user plane traffic, voice and real-time gaming traffic have the most stringent requirements, therefore QCI 1 and QCI 3 are marked with DSCP 46 (EF).

IMS signaling traffic is treated as AF41. So QCI 5 is marked with DSCP 34 (AF41).

Regarding the video services we distinguish whether guaranteed bit rates are provided and mark these services differently. QCIs 2 and 4 use GBR-bearers and are marked with DSCP 26 (AF31). QCIs 6 and 7 use non-GBR bearers and are marked with DSCP 18 (AF21).

The default bearer (QCI 9) is marked with DSCP 0 (BE), allowing any premium data traffic to be marked with DSCP 10 (AF11).

QCI 8 and 9 are used to distinguish between interactive and background data traffic, e.g. web serving and file download. Alternatively they differentiate between user classes (ordinary/premium, etc).

8.7.6.4 O&M

O&M traffic is marked with DSCP 16 (CS2), for the same reasons as in the Iub example.

8.7.6.5 Synchronization (by Packet)

Synchronization for LTE (FDD) follows the reasoning discussed in the Iub example.

Assuming IEEE1588v2, the traffic is marked with DSCP46 (EF).

8.7.6.6 Summary

The DSCP marking is presented as a summary in Table 8.8.

A mapping to four queues is considered, as there may not always be as many queues in the backhaul network as there are DSCPs used.

Q1, a strict priority queue, is used for:

- Network control.
- Voice.
- Real-time gaming.
- Synchronization.

Table 8.8 Example QoS mapping for the S1.

	Traffic type	DSCP
Control plane		34 (AF41)
Network control		48 (CS 6)
User plane		46 (EF)
	QCI 1 (voice)	26 (AF31)
	QCI 2 (streaming, live video, GBR)	46 (EF)
	QCI 3 (real-time gaming)	26 (AF31)
	QCI 4 (streaming, buffered video, GBR)	34 (AF41)
	QCI 5 (IMS signaling)	18 (AF21)
	QCI 6 (streaming, live video, non-GBR)	18 (AF21)
	QCI 7 (streaming, buffered video, non-GBR)	10 (AF11)
	QCI 8 (premium data)	0 (BE)
	QCI 9 (default bearer)	16 (CS2)
O&M		48 (CS 6)
Synchronization		

The GBR and non-GBR video services should not be merged together. GBR video services might be subject to call admission control. Typically this would not be the case for the non-GBR services. Treating these services together would make the call admission control pointless.

Q2 is used for:

- Control plane (S1-MME).
- QCI 2 and QCI 4 (GBR-bearers for video).
- QCI-5 (IMS signaling).

Q3 is used for:

- QCI 6 and QCI 7 (non-GBR bearers for video).
- QCI 8 (premium data).
- O&M.

The expected behaviour for the AF PHBs can be implemented using WRR/WFQ schedulers. Although e.g. non-GBR video services have more stringent requirements than premium data, this is not a strict priority relation. Instead each of the traffic types should get a defined share of the bandwidth in cases of congestion.

Finally, Q4 is for

- Default bearer.

8.8 Summary

Requirements for the backhaul QoS originate from the end user services, radio network layer operation, transport network control plane, synchronization and management plane. Backhaul service takes the above into account by supporting the treatment required for each of the traffic

types. Instead of having a differentiated treatment for each service separately, traffic types are aggregated in the mobile backhaul, which simplifies the backhaul implementation.

In the radio network layer, scheduling occurs for an individual user bearer. In the backhaul, multiple users are aggregated and the aggregate is scheduled. Radio network and transport network QoS functions need to be aligned. In this way the user is served with the required quality and network resources in both the air interface and in the backhaul network are used efficiently.

All radio networks, 2G, 3G and LTE, have their own needs for the transport service. Each radio network technology specific channel or traffic type is taken into account and corresponding packet flow is marked accordingly. The starting point in marking is the Differentiated Services Code Point (DSCP) field of the IP packet header. The DSCP value is typically configurable, and deduced from the radio network layer channels and the bearer attributes, and configured for other traffic types, such as transport network control and packet based synchronization.

From DSCP, the mapping is carried over to other network layers such as MPLS or Ethernet (L2). Once the mapping to the DSCPs is performed, the backhaul network QoS is based on the DSCP value and on the corresponding Per-Hop-Behaviour/Treatment aggregate.

References

- [1] 3GPP TS22.105 Service aspects; Services and service capabilities, v10.0.0
- [2] IETF RFC 768 User Datagram Protocol (UDP)
- [3] IETF RFC 791 Internet Protocol (IP)
- [4] Jacobson, Karel: ‘Congestion Avoidance and Control’, Proceedings of the Sigcomm ’88 Symposium, vol.18(4): pp. 314–329. Stanford, CA. August, 1988
- [5] IETF RFC 2581 TCP Congestion Control
- [6] IETF RFC 3782 The NewReno Modification to TCP’s Fast Recovery Algorithm
- [7] Rhee: CUBIC for Fast Long-Distance Networks. Internet-Draft, 2008
- [8] IETF RFC 3168 The Addition of Explicit Congestion Notification (ECN) to IP, IETF, 2001
- [9] IETF RFC 1323 TCP Extensions for High Performance, IETF, 1992
- [10] Sridharan: Compound TCP: A New TCP Congestion Control for High-Speed and Long Distance Networks. Internet-Draft, 2008
- [11] Balakhrisnan, Seshan, Katz, ‘Improving Reliable Transport and Handoff Performance in Cellular Wireless Networks’, ACM Wireless Networks, vol.1, no. 4, Nov. 1995, pp. 469–481
- [12] Ratnam, Matta, ‘WTCP: An Efficient Mechanism for Improving TCP Performance overWireless Links,’ IEEE Symposium on Computers and Communications (ISCC), 1998.
- [13] Vangala, Labrador, ‘The TCP SACK-Aware-Snoop Protocol for TCP over Wireless Networks’, IEEE VTC, Orlando, FL, vol. 4, Oct. 2003
- [14] IETF RFC 2018 TCP Selective Acknowledgement Options.
- [15] Bakre, Badrinath: ‘I-TCP:Indirect TCP for Mobile Hosts’, Proc. IEEE ICDCS’95
- [16] Moller, Molero, Johansson, Petersson, Skog, Arvidsson, ‘Using Radio Network Feedback to Improve TCP Performance over Cellular Networks,’ Proc. of the 44th IEEE Conference on Decision and Control, December 2005.
- [17] Goff, Moronski, Phatak, Gupta, ‘Freeze-TCP: A true end-to-end TCP enhancement mechanism for mobile environments’, Proc. of IEEE Infocom 2000, Tel-Aviv, pp. 1537–1545, 26–30. Mar. 2000
- [18] Casetti, Gerla, Mascolo, Sanadidi, Wang, ‘TCP Westwood: end-to-end congestion control for wired/wireless networks’, Wireless Networks, v.8 n.5, p. 467–479, September 2002
- [19] IETF RFC 2474 Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers
- [20] IETF RFC 2475 An Architecture for Differentiated Services
- [21] IETF RFC 2597 Assured Forwarding PHB Group

- [22] IETF RFC 2598 An Expedited Forwarding PHB
- [23] IETF RFC 3246 An Expedited Forwarding PHB (Per-Hop Behavior)
- [24] IETF RFC 3260 New Terminology and Clarifications for Diffserv (Informational)
- [25] Wang: Internet QoS. Architectures and Mechanisms for Quality of Service. Morgan Kaufmann Publishers, 2001.
- [26] IETF RFC 5462 Multiprotocol Label Switching (MPLS) Label Stack Entry: 'EXP' Field Renamed to 'Traffic Class' Field
- [27] IETF RFC 3270 MPLS Support of Differentiated Services
- [28] IETF RFC 5865 A Differentiated Services Code Point (DSCP) for Capacity-Admitted Traffic
- [29] IETF RFC 4594 Configuration Guidelines for DiffServ Service Classes
- [30] IETF RFC 5127 Aggregation of DiffServ Service Classes
- [31] IETF RFC 1812 Requirements for IP Version 4 Routers
- [32] IEEE 802.1Q-2005 IEEE Standard for Local and metropolitan area networks, Virtual Bridged Local Area Networks
- [33] IETF RFC 2698 A Two Rate Three Color Marker
- [34] Soldani, Li, Cuny: QoS and QoE Management in UMTS Cellular Systems. Wiley, 2006.
- [35] 3GPP TS23.107 Quality of Service (QoS) concept and architecture, v10.1.0
- [36] 3GPP TS23.203 Policy and charging control architecture, v10.4.0, v10.4.0
- [37] 3GPP TS23.207 End-to-end Quality of Service (QoS) concept and architecture, v10.0.0
- [38] 3GPP TS 25.401 UTRAN overall description (Release 10), v10.2.0
- [39] 3GPP TR 25.902 Iub/Iur congestion control, v7.1.0
- [40] Kaaranen, Ahtiainen, Laitinen, Naghian, Niemi: UMTS Networks, Architecture, Mobility and Services, Second Edition, Wiley 2005
- [41] 3GPP TS 23.401 General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access, v10.5.0
- [42] 3GPP TS 36.300 Evolved Universal Terrestrial Radio Access (E-UTRA), Overall description, v10.5.0
- [43] 3GPP TS 36.401 Evolved Universal Radio Access Network (E-UTRAN); Architecture description, v10.3.0
- [44] MEF 22 Mobile Backhaul Implementation Agreement – Phase 2, MEF, 2012
- [45] MEF 23 Carrier Ethernet Class of Service – Phase 2, MEF, 2012

9

Security

Esa Metsälä and José Manuel Tapia Pérez

In the mobile network, security is covered in a number of 3GPP specifications. Most of the security features of the mobile system do not directly involve backhaul. Before designing protection for a mobile backhaul application, it is useful to study what is offered by the mobile system itself. A review of this is provided in Section 9.1.

Also for the backhaul 3GPP sets the framework. In many cases, IPsec is required for the protection of the backhaul. 3GPP view concentrates on protecting the IP layer with IPsec. This is reasonable, as the layers below IP are not specified. Without cryptography, a level of protection can be achieved by keeping traffic types separate. Traffic separation and L2 specific protection is the topic of Section 9.2.

IP layer protection with firewalls, access lists, and with the cryptographic protection achieved with IPsec, is discussed in Section 9.3. IPsec protocols and related tools that are relevant for the mobile backhaul are reviewed.

Finally, Section 9.4 considers issues related to IPsec VPN deployment: QoS, resilience, fragmentation, etc. In this chapter, a case with LTE S1 and X2 protection with an IPsec VPN is explored. Section 9.5 presents a summary.

9.1 Security in 3GPP Mobile Networks

The objectives for security, as documented in 2001 by 3GPP [1] are that:

- user information is protected against misuse;
- network resources and services provided by serving networks and home environments, are adequately protected against misuse;
- security features are standardized, have worldwide availability (considering e.g. export restrictions) and are interoperable so that roaming can be supported between different serving networks;

- level of protection for both user, and for providers of service, is better than that provided by contemporary fixed or mobile networks;
- security features and mechanisms are extendable as required, to address potential new threats and services.

The experiences from 2G security were the basis for creating the objectives for 3G, especially the identified shortcomings, real and perceived, of 2G security were addressed. Otherwise the security elements and components from 2G that were perceived to be robust were kept as a basis.

The documented 2G weaknesses include:

- Possibility for a ‘false BTS’ attack (since terminal is authenticated, but not the network). This is addressed with mutual authentication.
- Transmitting keys as clear-text. Between networks (and network domains), network domain security functionality has been included in 3G.
- Weaknesses in authentication, related to IMEI (International Mobile Equipment Identity).
- Lacking data integrity protection. Integrity of air interface signaling is protected in 3G.
- Lawful interception and fraud information gathering, was not initially considered in 2G.
- Encryption over the air interface was not extending deep enough into the network.
- Lacking flexibility for subsequent release new functionality (e.g. encryption algorithms and key lengths not sufficiently strong, and not easily extendable).

Security threats in the 3G system are according to TS33.120:

- Unauthorized access to sensitive data: e.g. by eavesdropping.
- Manipulation of sensitive data: e.g. by modifying the content of communication.
- Disturbing network service: e.g. by triggering traffic flows leading to denial of service.
- Repudiation: e.g. claiming that actions of a user or a network have not taken place.
- Unauthorized access to services: e.g. by misusing access rights.

As an exception to the confidentiality of the communication between users, in some cases law enforcement may be entitled to interfere with the communication between users. This happens technically through lawful interception (LI) functions. The rules of when and to what extent law enforcement has the right for lawful interception, are subject to legislation.

Many of the threats above are related to handling user information such as authentication of the mobile user, billing and accounting of the user, etc. These are all transparent to the radio network and to the mobile backhaul. Thus they are addressed by the mobile system standards. Clearly there are threats originating from the mobile backhaul; user data can technically be accessed and potentially manipulated, and also network service disturbed. These threats can be addressed by network security functionality, using cryptography (IPsec) and other features.

3GPP security architecture is based on features grouped into five areas, shown in Figure 9.1.

Network access security protects the wireless access to 3G services. *Network domain security* makes the mobile nodes’ exchange of sensitive information secure, over the wireline network. *User domain security* concerns of access to the mobile stations. *Application domain security* secures communication between user and network domains. Finally, *visibility and*

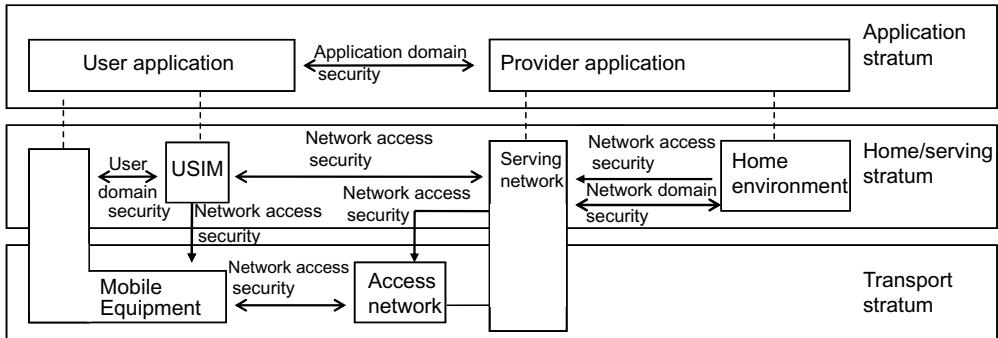


Figure 9.1 3GPP Security architecture [2].

configurability of security inform users of whether security features are in use, and, whether the service should depend on the availability of a security feature.

Features for the network access security exist in areas of user identity confidentiality, entity authentication, confidentiality on the access link, data integrity on the access link and mobile equipment authentication. User domain security features are in providing user to USIM authentication and ensuring an access to the terminal only via an authorized USIM. Application domain security includes e.g. securing applications that reside on the USIM.

Network domain security was addressed starting from 3GPP Rel-4, with the focus on core network signalling and securing the Mobile Application Part (MAP) protocol of the SS7 signalling system. IP layer security was included with 3GPP Rel-5. The resulting specification in Rel-5 is TS 33.210, defining the use of IPsec protocols. Security within the mobile backhaul falls into the network domain security area. The services needed are confidentiality, integrity and authentication.

9.1.1 Network Domain Security

IP networking is increasingly used in all radio technologies, 2G, 3G, HSPA, HSPA + and LTE. This means that open and well-known protocols (TCP/IP protocol suite) are used to backhaul the mobile network traffic. Having an access to the mobile backhaul network, means thus that there is IP connectivity to the mobile network elements, unless this access is restricted.

This opens the mobile network up to threats and risks that did not exist with TDM or ATM networks. IP is also used for all of the traffic types: User plane traffic, signalling/control plane, management plane and synchronization plane traffic. All of the traffic types have their own vulnerabilities.

3GPP defines in TS33.210 a security domain. This means a network that is managed by a single administration. So, typically, a network that is operated by a single operator, is also a security domain. Within a security domain typically the same level of security is maintained.

TS33.210 is written with the mindset of protecting the control plane. Radio access network internal interfaces (such as Iub), are not covered. Principles can be reused for interfaces that are not explicitly mentioned in the specification. In LTE, S1 and X2 interfaces are covered in LTE security architecture explicitly (TS33.401), so a standards-based view is clear.

With the network domain security, the interface to other security domains is designated as interface Za. A security gateway (SEG) is an element that is used at the interface to another security domain. A peer entity over the Za interface is SEG of the other security domain. An interface within a security domain is interface Zb. This is between a network element (NE) and another NE, or between NE and SEG.

Additionally, a transit security domain is defined as a domain which is transmitting NDS IP traffic between other security domains. The interface used between a transit security domain and other security domains, is Za.

Figure 9.2 matches directly e.g. a roaming case where two different operators interconnect and need to protect the common control plane (e.g. GTP-C) with a commonly agreed set of features.

A mobile network element (NE) may include an integrated SEG. In this case the SEG logically interfaces Za. Multiple SEGs may be operated in parallel, for load sharing or for resiliency. ESP tunnels may be available constantly, or established dynamically when needed.

For the interface between security domains (Za), authentication and integrity protection is mandatory, and encryption is recommended, all these functions using IPsec Encapsulating and Security Payload (ESP) protocol. The ESP tunnel between SEGs is negotiated, established and maintained using IKEv2 (according to 3GPP Rel-11 specification). Previously, IKEv1 may have been implemented. Rel-11 mandates SEGs to support both versions (IKEv1 and IKEv2), to ensure interoperation over the Za.

Zb interface is optional. If implemented, it supports ESP tunnel mode, and IKEv2 (according to 3GPP Rel-11). IKEv1 may be supported, as well as ESP transport mode. Authentication and integrity protection is mandatory and encryption is optional. In particular, control plane protection needs to be considered over the Zb.

The concept presented assumes that inter-domain communication always occurs through the use of SEGs, so that there is no direct NE-NE inter-domain communication. In the actual implementation, SEG functionality may be integrated with an NE, which is also discussed in

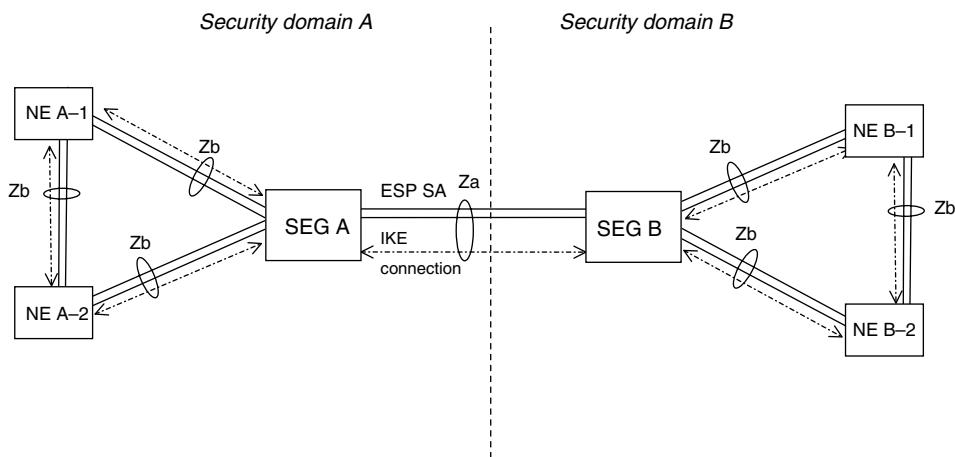


Figure 9.2 Za and Zb interfaces [3].

3GPP TS33.210. Secure inter-domain NE-NE communication is supported, if the combined NE/SEG node interfaces another combined NE/SEG node in the other security domain. The features supported in this case depend on the security policy. The combined NE/SEG node can interface either a combined NE/SEG node, or a SEG-only node in the other security domain.

For protecting the traffic in the mobile backhaul, both dedicated SEGs, and combined NE/SEG nodes are potentially relevant. In cases where the access tier of the backhaul network requires cryptographic protection, mobile system interfaces like Iub and S1/X2 have to be considered. Deciding whether the backhaul lacks adequate protection, has to be assessed case by case. If cryptographic protection is needed, then SEG is deployed for the traffic leaving the security domain.

At the BTS site, IPsec protection can be deployed either as a cell-site gateway with IPsec functionality (acting as a dedicated SEG, shown as Alternative 1 in Figure 9.3), or as implementing the needed IPsec functions into the eNodeb (LTE) (shown as Alternative 2 in Figure 9.3). The BTS site interfaces the unsecure network via Za. At the other end (core site), similarly a Za interface is used.

Both alternatives protect the S1 over the originally not-adequately-secured transit network. In Alternative 1), BTS site consists of two physical nodes, eNodeB and SEG. In this case, protection of the intra-site connection between eNodeB and SEG needs to be considered in addition.

In Alternative 2), SEG is integrated with eNodeB into a combined node. In this case, the connection from the eNodeB to SEG is node-internal. In the combined node case, the S1 interface is only available in an encapsulated mode (ESP tunnel). (As a side effect, troubleshooting S1 e.g. with protocol analyzers, is not possible. If an unprotected port is offered by the combined node, protection is compromised.)

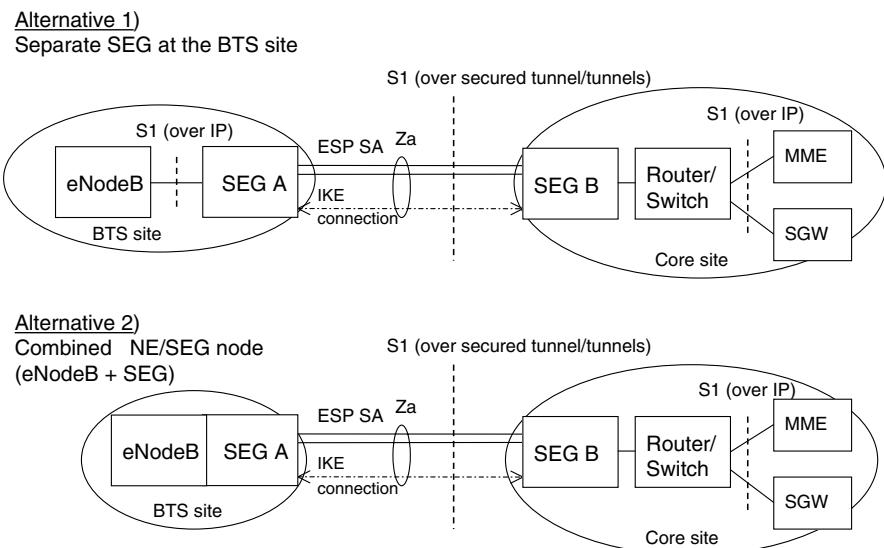


Figure 9.3 Dedicated SEG and combined NE/SEG node [3].

In the case of combined NE/SEG node, a further limitation is imposed, in that the combined NE/SEG node should not be used by other NEs towards other security domains. In Figure 9.3 as it is presented, in Alternative 2 this has the consequence that according to 3GPP, IPsec transport service is not allowed to be used for a possible co-located 2G or 3G BTS (or a chained BTS).

9.1.2 2G

At the mobile system level, three key security features of 2G are: subscriber authentication, encryption of the air interface, and the use of temporary identities. Subscriber authentication means that the subscriber has no access to the network unless successfully authenticated. (Note the exception of emergency calls). Temporary identities protect the identity of the user. International Mobile Subscriber Identity (IMSI) is the permanent identity of the subscriber. In addition to IMSI, temporary identities are defined.

Encryption is discussed further as it is relevant for considering the protection needs of the 2G backhaul.

The mobile voice service over GSM is designed to be no more vulnerable to eavesdropping than a fixed voice service. Air interface is considered a weakness compared to the voice service over fixed lines, and thus encryption is specified over the air interface between the mobile station and the BSS.

In the 2G system, ciphering is supported in the user plane in BTS for voice, and in LLC layer for the packet switched data (GPRS). This means that user plane traffic is ciphered over the air interface, from the mobile station to the BTS (for voice) and from the mobile station to the SGSN (for packet switched data). For the encryption function in the BTS, BSC is responsible for downloading the encryption key to the TRX. The key is downloaded using control messages between the BTS/TRX and the BSC.

The cryptographic algorithms used in 2G are variants of A5. A5/0 means no encryption. A5/1 is considered medium level encryption, A5/2 weak, and A5/3 as strong encryption. A5/3 is based on the algorithm used in 3G. For 2G/GPRS services, algorithms GEA1, GEA2 and GEA3 are used. Selection of the algorithm depends on the terminal and network capabilities. The algorithms are used for encryption only.

9.1.3 Abis, A and Gb

Abis interface itself (as a PCM-interface), is not providing any specific security feature. This is also comparable to voice on the fixed access systems. PCM-based Abis has not been viewed as a security vulnerability. The control traffic on Abis (LAPD links) is similarly in clear-mode over the TDM, unless protected separately.

Microwave radio transmission, commonly used for Abis (as well it is also used for Iub and S1), was perceived as a 2G system vulnerability, since wireless transmission is used. Point-to-point microwave radio links have a vendor-specific air-interface and they are typically not encrypted. Abis traffic is in clear-mode so it is open to eavesdropping and other security vulnerabilities.

Microwave point-to-point links, however, differ from the air interface transmission. Air interface of the BTS consists of omnidirectional, or sectorized cells. With point-to-point

microwave radio, line-of-sight is required, and the signal attenuates rapidly outside of the line-of-sight. Eavesdropping is naturally still technically possible. In 3G, the RLC/MAC layers are responsible for encryption, and they are located in the 3G RNC, and thus the radio bearers over the Iub, whether using microwave radios or some other backhaul, is encrypted.

When Abis is circuit emulated over the packet mobile backhaul, it is similarly based on the IP protocol stack, and exposed to the same vulnerability as native IP based Iub of 3G (Rel-5 onwards) or LTE S1 and X2. Specific concerns related to the access transport in general (Abis, Iub, S1, X2) are the high amount of BTSs, and the physical location of BTSs. Many BTS sites are not physically as secure as controller or core network sites.

The case of carrying Abis over packet mobile backhaul is implementation specific, as circuit emulation is out of the scope of 3GPP. IPsec protocol suite can be deployed and the guidance from other 3GPP specified IP interfaces for NDS can be adopted for the emulated Abis interface.

A interface connects the BSS to the circuit switched core. The user plane interface is either PCM as initially defined, or it may be A over IP. Gb interface connects the BSS to the GPRS core network. Initially frame relay was defined, and Gb over IP was added as an option.

3GPP specifications support both A and Gb with IP transport option for both user and control traffic. Although not identical, the A and Gb interfaces are comparable to the Iu-interfaces of 3G: The A interface is functionally comparable to the Iu-cs interface, and the Gb to the Iu-ps interface.

The protection of UTRAN/GERAN IP based protocols is discussed in Annex of TS33.210, stating that traffic shall be protected properly and that the NDS/IP protection principles are to be followed. Za interface is to be implemented between security domains. The document explicitly mentions Iu-mode interfaces, however it can be assumed that the same principles apply to GERAN A and Gb interfaces, as the protection needs of A and Gb are comparable to those of Iu-cs and Iu-ps.

9.1.4 3G

For 3G, the 2G security features that were assessed as robust and needed were taken as a basis. So, for example, subscriber authentication, air interface encryption and the use of temporary identities continued to be supported, with enhancements. At the same time, the weaknesses identified in the 2G system were addressed. 3G security features are an evolution of the corresponding 2G functionality, with enhancements where needed. For further information refer to [9].

For the traffic types, user plane traffic from the UE is encrypted for both CS and PS domain by the radio network layers. Now the encryption is extended, as the RLC/MAC layers support the encryption, and these are located in the RNC. This addresses one of the documented 2G weaknesses (Encryption not extending deep enough in the network). Similarly, control plane traffic from the UE is encrypted between the user equipment and the RNC. Additionally, UE control plane traffic is also integrity protected, which is again an enhancement compared to 2G (and again a documented 2G weakness). User traffic is not integrity protected.

A separate encryption key (as in 2G) is used for the CS and PS domains for the user traffic over the air interface. These keys are derived by the core network (SGSN and MSC/VLR) using information in the user equipment (UE, USIM) and in the core network subscriber

registers. The keys are delivered to the RNC by RANAP signaling. Appropriate security mode is selected based on terminal and network capabilities, and possible further information about the needs of the application.

UMTS AKA (Authentication and Key agreement) procedure is used to authenticate the user. Now also the network is authenticated, which is an improvement compared to 2G.

9.1.5 *Iub*

Similarly to the PCM-based Abis interface of 2G, with initial 3G RAN, the ATM based interface of Iub was not viewed as a real vulnerability from a security standpoint. With IP defined in 3GPP Rel-5 for the 3G Iub, the situation changed.

3GPP Rel-5 IP Iub can be used for all traffic from a NodeB, for user traffic on DCH and HSPA channels, and for signaling and management traffic. Also, packet based timing may be carried over IP. With Rel-5, the assumption is that of a closed IP network, in which case the network can be considered secure and trusted, and IPsec solution is not mandated by 3GPP. With this assumption, there is no need to specify IP Iub within network domain security. Whether the backhaul network deployed for the Rel-5 compliant IP is secure and can be trusted, has to be assessed case by case. In LTE, physical security for the eNodeB is the exception given in 3GPP, so as not to mandate cryptographic protection.

In addition to user and control plane traffic from the UEs, NBAP signaling traffic between nodeB and the RNC exist in the Iub interface. NBAP/SCTP/IP carries control messages necessary for the NodeB operation. There is no specific security functionality specified related to NBAP itself.

Similarly, other possible traffic types extending over the Iub to the NodeB but not over the air interface, are by default not protected by any specific security features. Network management and packet timing are examples of this kind of traffic. Network management traffic is critical and is typically protected even with ATM based interface, as over the ATM IP is used, and because of the risk level associated with O&M procedures.

Even though the mobile network is authenticated to the user equipment, UMTS authentication procedure is not authenticating any network elements (nodeB and RNC). This has to be arranged for separately, at the IP layer e.g. by using IPsec.

User plane traffic is encrypted up to the RNC due to the RLC/MAC layers, as well as the control plane to the UE (RRC messages). As mentioned, integrity protection is also added to the radio layer signaling (RRC).

Threats related to IP transport on Iub in general are similar to the other IP interfaces. Specific concerns related to the Iub are the high amount of NodeBs, and the physical location of NodeBs. Many NodeBs sites are not physically as secure as controller or core network sites. If the Iub backhaul is not closed (physically secured), IPsec can be applied. Following 3GPP NDS/IP specification for eNodeB access (LTE case), IPsec is mandated, unless the site is physically protected.

9.1.6 *Iu-CS, Iu-PS and Iur Interfaces*

Iu-CS, Iu-PS and Iur interfaces support both user and control plane traffic. Iu-CS user plane connects to the CS core, with RTP/UDP/IP stack (RTCP optional). Iu-CS control plane stack

consists of RANAP/SCTP/IP. Iu-ps user plane is based on GTP-U/UDP/IP, and the Iu-ps control plane is RANAP/SCTP/IP. Iur connects to another RNC, with FP/UDP/IP in the user plane and RNSAP/SCTP/IP in the control plane. Iu-cs or Iu-ps user and control planes do not include any specific protecting functionality on RANAP or Iu user plane layers. (Note that since the serving RNC terminates the radio Layer 3 (RRC) and Layer 2 (RLC/MAC) protocols, Iur is comparable to Iub in this aspect). RNSAP on the Iur does not include any specific protecting functionality.

The protection of these interfaces is based on NDS/IP (IPsec), when needed. As mentioned, NDS formally covers only control plane (RANAP and RNSAP).

Iu interface control plane (both Iu-cs and Iu-ps) carries sensitive information, e.g. the encryption keys, which is why 3GPP specifies the use of encryption along with integrity protection with IPsec ESP, when traversing inter-security domain boundaries. 3GPP documents the reason for requiring this so as not to limit the applicability of IP protocol to a closed network case. Logically this is the Za interface with a SEG. If the Iu interface does not cross a security domain boundary (Zb interface), protection with IPsec is optional.

9.1.7 LTE

LTE security architecture is defined in TS33.401, and it is further built on the 2G and 3G security features, with enhancements. Compared to 2G and 3G specifications of 3GPP, security functionality has been covered in further detail, also related to the eNodeB interfaces, S1 and X2.

For the authentication, with LTE it is possible for the UE to authenticate the network (serving network identity).

In the user plane, air interface is ciphered between UE and eNodeB. Integrity protection is not defined for the user plane. For the S1 and X2 interfaces, user plane protection relies on the use of IPsec (unless physical protection can be assured). In the control plane, protection of signaling includes both ciphering and integrity protection with replay protection.

The algorithms specified are SNOW 3G (128-EEA1 and 128-EIA1) or AES (128-EEA2 and 128-EIA2). These algorithms are used for both AS (access stratum) and NAS (non-access stratum) signaling, and for user plane, and the algorithm is selected based on UE and network capabilities and allowed security algorithms.

For the EPS AKA, LTE introduces new functionalities compared to the UMTS AKA. 3G USIMs are allowed, but 2G SIMs are not supported, so LTE maintains a backwards-compatibility only to the previous generation (3G) but not to 2G. MME is in the role of VLR/SGSN when compared to 3G. With LTE, key derivation includes more steps and intermediate keys than 3G.

Temporary identities are supported in a way comparable to 2G and 3G. Additionally, terminal identity confidentiality is introduced by conveying the terminal identity only after security mode for NAS is enabled.

9.1.8 S1 and X2 Interfaces

LTE differs from both 2G and 3G since in the LTE architecture, eNodeBs connect directly to the core network. This also means that there is IP connectivity from each eNodeB site to the core network, unless this connectivity is limited. The eNodeB is the only element in the radio

network, so that the S1 interface is logically comparable to the Iu interface, as it is the interface between the radio network and the mobile core network. In the user plane, the protocol stack is GTP-U/UDP/IP for S1 and X2, in the user plane. In both S1 and X2 control plane, the protocol stack is SCTP/IP.

From the mobile backhaul technology viewpoint, LTE is an IP network from the start, so network domain security is a key concern.

For the user plane (S1-U and X2-U), TS33.401 mandates integrity, confidentiality and replay-protection using IPsec (clause 12), unless the eNodeB is in a physically protected environment. In the latter case S1-U and X2-U links are considered to belong to an extended secure environment. Similar requirements are set for the control plane (S1-MME and X2-C): integrity, confidentiality and replay-protection (clause 11) is mandated, with the same exception allowed for a physically protected environment.

The user plane (S1-U and X2-U) protection in practice is implemented with IPsec ESP (RFC4303) with a profile defined in TS33.210. Tunnel mode is mandatory, transport mode is optional. At the core network side, a SEG may terminate the tunnel. IKEv2 certificate based authentication is specified, with a certificate profile and IKEv2 profiles defined in TS33.310.

The control plane (S1-MME and X2-C) requires IPsec ESP (RFC4303) as specified in TS33.210. IKEv2 certificates based authentication is required (TS33.310). Tunnel mode is mandatory and transport mode is optional.

eNodeB is responsible for ciphering the air interface, and the S1 and X2 interfaces (S1 and X2 via IPsec). Ciphering has to be implemented in a secure environment.

A secure environment is defined in LTE security specifications as a group of sensitive functions within the eNodeB, and a secure storage for sensitive data. Functions are e.g. encryption and those phases of authentication that use long-term cryptographic secrets. Sensitive functions in the boot process also belong to the secure environment. Integrity of the environment itself must be ensured, and the access to the environment must be limited to authorized personnel.

9.1.9 Management Traffic

For LTE, TS33.401 requires that attackers:

- Are not able to modify eNodeB settings or configurations via local or remote access.
- Security associations between the eNodeB and the EPS core are required, and they have to be mutually authenticated.
- Communication between O&M systems and the eNodeB have to be mutually authenticated.
- ENodeB software and data change attempts have to be authorized.
- Confidentiality and integrity protection of software transfer have to be ensured.

Often, management plane traffic is carried together with the S1 traffic, over the same physical links. The same protection mechanisms as defined for the S1 interface, can be reused for the management plane traffic (denoted S1-M).

Tunnel mode IPsec is mandated for the S1-M on the eNodeB. At the other end of the management plane, a SEG can be used, or the functionality may be implemented in an element manager. IPsec ESP, with the profile as specified in TS33.210, including confidentiality,

integrity, and replay protection. IKEv2 with certificate based authentication is specified for the eNodeB for the S1-M. Profile is defined in TS33.310.

If the management plane is carried on separate links than S1, essentially an equal level of protection is required. Typically management plane shares the physical links to the eNodeB with other traffic, as arranging separate links is costly. If the management plane interfaces can be trusted (they are physically secured), protection with IPsec can be omitted.

Often there is an additional level of protection at a higher layer by the network management system, since it is essential that communication to the management system is protected. The O&M is vendor specific. As an example, TLS (Transport Layer Security) can be deployed end-to-end between the NMS and the Network Element (NE). TLS can also use IPsec as the IP layer service.

9.2 Protection of the Backhaul

9.2.1 Cryptographic Protection Compared to Other Protection

As discussed, 3GPP in a number of cases mandates the use of IPsec for the mobile backhaul, unless the network is physically secure. IPsec is addressed in a separate section (9.3).

In addition to the services provided by IPsec protocols, further protection within the mobile backhaul may need to be considered. Backhaul service may be disturbed even if IPsec is deployed, if the backhaul is open to threats by other means of attack, or by attacks on other protocol layers. 3GPP does not specifically address the need for protection on other layers, or protection of the mobile backhaul in general.

As an example, a router may forward the IPsec encapsulated traffic to a false destination, if the routing table has been manipulated maliciously. Similarly, a Layer-2 bridge may be overloaded with non-legitimate traffic, so that the IPsec encapsulated packets cannot be forwarded further.

This type of issues and threats are not specific to the mobile network backhaul only, nor are the design practices and features that address them. Some of the key issues, mostly related to L2, are included in this section. They are specific to a *packet* mobile backhaul, as similar threats did not exist in the TDM era.

Protection of the backhaul against various threats is closely related with QoS and resilience. QoS features, such as ingress policing or admission control, prevent unauthorized or excessive use of network resources. Resilience targets maintain the service. Many threats aim at making the service unavailable. Security and resilience have thus a common goal of keeping the network operational.

Traffic separation, as well as other non-cryptographic protection, can be considered as additional and complementary measures to IPsec. They are not replacing the protection IPsec offers. On the other hand, IPsec alone does not address all of the threats within the mobile backhaul.

9.2.2 Leased Service and A Self-Deployed Backhaul

A mobile operator may have his own self-managed mobile backhaul or it may be leased from a service provider. Often it is a combination of the two.

Security has to do with trust, and protection is often deployed at the administrative boundaries, between two operator's domains. These operator domains are also different security domains. If the network service is not trusted, a security gateway can be implemented at the network border. Traffic can be carried within an IPsec VPN over the service provider's network.

The situation is, however, not as simple. Threats originate also from the inside of a security domain, not only from an external attacker. Misconfigurations take place. They may be unintentional, or done maliciously. Many of these threats can at least partly be mitigated by sound network design guidelines and operating practices. Besides, many topics are vendor and node technology specific. A few of these topics are addressed in the subsequent text.

9.2.3 *Traffic Separation*

Separating traffic logically isolates the backhaul network into separate sections or domains. Communication over a domain is typically restricted and controlled, which increases robustness of the system.

At the Ethernet layer, VLANs limit the broadcast domain so that unknown unicast and broadcast frames are not flooded into other VLANs. This already increases the level of protection, as issues (whether due to attacks or misconfigurations) in one VLAN, do not necessarily impact the service on other VLANs.

For traffic to pass on from one VLAN to another, a router is required. With a router, an IP layer control point is introduced. Traffic flows may be allowed or blocked according to a security policy, e.g. by using an Access control list (ACL).

Often, management plane (O&M) traffic is first separated from any other traffic type. As O&M allows many kinds of procedures with a remote configuration, it is specifically a concern. There is also no legitimate traffic need between O&M and user or control plane within the mobile system. Connectivity is only needed to the management system, and thus should be blocked to other destinations.

Metro Ethernet service definition include Ethernet Private Line (EPL), Ethernet Private LAN (EPLAN), Ethernet Virtual Private Line (EVPL) and Ethernet Virtual Private LAN (EVPLAN) services. EVPL and EVPLAN allow an Ethernet virtual connection (EVC) to be separated according to VLAN, so that VLAN based separation is also supported over the service.

Similarly with self-deployed Carrier Ethernet with MPLS/IP, VLANs can be used at the PE ingress node when mapping the attachment circuit to the service. In the MPLS core, traffic from different VLANs is kept separate from each other.

9.2.4 *Ethernet Services*

With MEF services, E-Line defines a point-to-point connectivity while E-LAN supports multipoint connectivity. (E-Tree is a variant of E-LAN as a rooted multipoint with a restricted connectivity between spokes, which can only reach the hub node).

All of these services logically extend the Ethernet LAN service over the wide area network, using e.g. MPLS/IP. While the MPLS/IP core itself is not vulnerable to the Ethernet LAN related threats, the customer service (e.g. E-LAN), basically has the vulnerabilities of a LAN.

With an E-LAN, the provider network looks like a virtual bridge: MAC address learning is supported, with bridging functions like unknown unicast and broadcast flooding.

E-LAN service, if utilized for the mobile backhaul, offers Ethernet level connectivity between all sites connected to the same service instance (VLAN). Logically this is a single VLAN where every station can reach every other station. Allocating stations to different VLANs support further traffic separation.

Point-to-point service (E-Line), or rooted multipoint (E-Tree) provide better isolation, as the spoke stations are not able to communicate with each other. With 2G and 3G, there is also no need for BTS to BTS connectivity, as the logical topology is a hub-and-spoke (E-Tree -like).

LAN related security risks, as well as proposed countermeasures, are discussed e.g. by Vyncke and Paggen [12]. Many topics are vendor specific, in that the configuration options and parameters available to harden the bridges vary. Not all Ethernet LAN related threats and risks can be covered here. A few examples, based on [12], are given in order to illustrate the topic.

Basically plugging in a new station to the LAN allows direct connectivity at the Ethernet layer, via the flooding, bridging/MAC learning processes. The first topic is to disable unused bridge ports. They can also be configured to an unused VLAN. This isolates the port.

DoS attack is easy if the LAN can be accessed via some active port. Traffic can be flooded, either as broadcast frames, or as control traffic. Control frames, such as spanning tree BPDUs or other Ethernet control protocols need processing, and control processor may crash unless protected. Ingress policing limits the maximum amount of traffic that is allowed at ingress of any port, and control plane policing limits the amount of control traffic that is allowed to enter the control processor.

Spanning tree protocol may be needed in the customer sites for redundancy. When spanning tree initially was developed, security vulnerabilities were not considered.

One topic in spanning tree is the root bridge selection. If a bridge is added to the LAN with a better (lower) bridge ID, it claims the root role. All traffic is thus passed via this new bridge. Bridges may have configuration options that limit the possibilities of accepting better bridge IDs from certain ports, or even accepting any type of BPDUs. There typically are also further (vendor specific) parameters to harden the bridge against misconfigurations and malicious attacks.

A simple attack consists of flooding BPDUs that need to be processed by the control processor. This can be mitigated similarly by not accepting BPDUs from a certain port. Also, the incoming frames can be policed so that a certain maximum rate is not exceeded.

Similarly as spanning tree, ARP protocol vulnerabilities were not considered at the time of the design. ARP is not authenticated. Also, all stations connected to the VLAN will learn the IP/MAC address mapping. As all stations also receive ARPs due to the broadcast nature, they also need to process it.

With an ARP spoofing attack, fake ARP replies (gratuitous ARP) are generated by an attacker. Consequently, the attacker receives all frames that were directed to the corresponding IP address. If this is the address of the default gateway, the attacker receives all traffic that is forwarded to the default gateway. For the mobile network, this may in practice mean all traffic from a certain radio access area, depending on the backhaul architecture.

One simple way is to not to allow hosts and devices to accept gratuitous ARP. This, however, slows down recovery in the case of failures. Another way is to maintain a table of ‘valid’ IP – MAC address tables and detect if an attempt is made to modify the bindings.

What is the level of protection IPsec can then provide? Clearly, IPsec does not mitigate Ethernet level vulnerabilities. However, the IP layer remains protected. DoS attacks are possible at the layer below. Application (radio network layer in the case of mobile backhaul), however, remains protected.

9.2.5 IEEE 802.1x and IEEE802.1ae

IEEE is addressing L2 security in recent work, in IEEE802.1x Port-based Network Access Control, in IEEE 802.1ae Media Access Control (MAC) security, and in 802.1ar Secure Device Identity.

Previously, wireless LANs (IEEE Std 802.11) support mutual authentication, generation of cryptographic keys, and using those keys for protecting the frames over the wireless LAN.

Port-based Network Access control supports authentication of devices attached to the LAN. Generation of the Secure Association Keys (SAKs) is defined in 802.1x. Secure Device Identity supports a cryptographic identity, that binds the identity to a device (such as a LAN station). Secure Device Identity supports authentication by IEEE802.1x.

MAC security means cryptographic protection. The default cipher suite is GCM-AES-128. Security relationships are maintained by MACsec Key Agreement. A definition of a secure Connectivity Association (CA) is used. Each CA supports Secure Channels (SCs) with symmetric keys. Each SC in the CA uses the same Cipher Suite. Within SCs, Secure Associations (SAs) are supported. These SAs use a Security Association Key (SAK). MAC security introduces a new Ether type (88E5H), with a new security tag.

802.1ae has been amended in 2011, with IEEE Std 802.1AEbn. The amendment adds a GCM-AES-256 Cipher Suite as an alternative.

When native Ethernet (or an Ethernet service) is used in the mobile backhaul, these protocols are candidates for increasing protection on the first mile from the BTS to the IP edge device.

9.2.6 MEF

In MEF service specifications, security is based on traffic separation, with EVCs defining how subscriber's traffic is kept separate (MEF10.2). Additional functionality is subject to subsequent phases in MEF 10.2. In the mobile backhaul implementation agreement (MEF22) for EVP-Tree, essentially a similar type of notion is given. A level of protection is achieved by the use of traffic separation.

9.3 IP Layer Protection

9.3.1 IPsec

IP Security (IPsec) is a set of protocols designed to provide cryptographic protection of a communication flow. It provides confidentiality, integrity protection and authentication services. It operates at the IP layer and as such it is able to protect any protocol carried by IP. For the same reason, it is not able to provide protection at the physical layer.

In mobile networks, the network domain security relies on the use of IPsec, which, according to the 3GPP is mandatory for all the communications between security domains (Za interface) and optional for the communications within the same security domain (Zb interface). Mobile backhaul, although not matching exactly either the Za or the Zb interfaces, is subject to a number of security threats which could be mitigated by using IPsec.

9.3.2 IPsec SA

An IPsec Security Association (SA) is a unidirectional logical connection established between two or more peers which defines the protection to be applied to the packets it carries. The IPsec SA has an associated list of parameters which governs how the packets are processed, the algorithms to be applied, the security protocol to be used, etc. These parameters are stored in the Security Association Database.

Provided that most of the communications are bidirectional, a pair of SAs is usually established. Packets carried by both SAs in a pair receive the same kind of protection.

IPsec SAs can be established manually by using a management interface of the implementation. However, this approach does not scale well when the number of SAs in a system increases, and it does pose the additional problem of how to securely distribute and maintain the session keys.

A much more practical approach for a mobile network is the use of a signalling protocol, such as IKE.

9.3.3 IPsec ESP

The Encapsulating Security Payload (ESP) [9] is one of the protocols in the IP Security suite used to protect the traffic. ESP supports confidentiality, data origin authentication, data integrity and anti-replay services. The exact set of services to be applied can be negotiated during the establishment of the Security Associations, depending on the capabilities of the peers as well as the security policies. Figure 9.4 illustrates the fields of the ESP packet format.

The SPI is used as an index to refer to the specific IPsec SA. The Sequence Number is an increasing number (and unique during the whole lifetime of the IPsec SA) which enables the

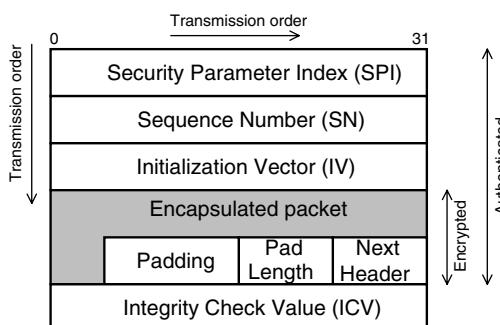


Figure 9.4 ESP encapsulation with encryption and authentication.

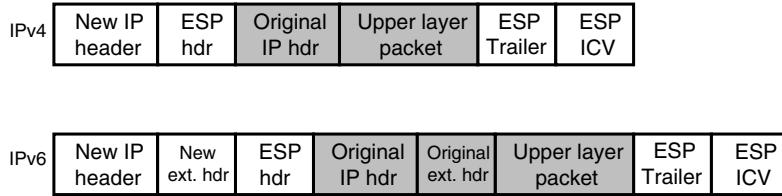


Figure 9.5 ESP encapsulation in tunnel mode.

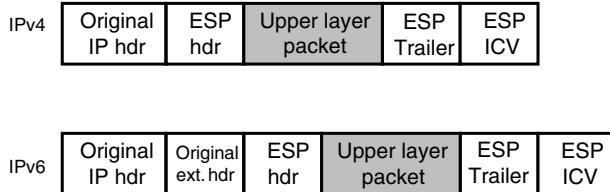


Figure 9.6 ESP encapsulation in transport mode.

anti-replay protection by the receiver. The Initialization Vector provides the value that initializes the encryption algorithm. The Next Header indicates the protocol of the encapsulated packet. The Integrity Check Value is used for the authentication and it contains the HMAC of packet.

Figures 9.5 and 9.6 illustrate the two possible ESP encapsulations, tunnel mode and transport mode, for IPv4 and IPv6.

As can be seen, the tunnel mode adds an additional IP header which is used to hide the internal IP packet. It is worth mentioning that the IP version of the inner packet and the IP version of the outer packet can be different so IPv4 can be encapsulated in IPv6 and vice versa.

As ESP does not protect the outer IP header, it is possible for a network device to perform address translation (NAT) without affecting the integrity of the packet. However ESP encapsulation does not use port numbers so port translation is not possible (NAPT). To overcome this limitation, an additional UDP encapsulation, as defined by [24] can be used for NAT traversal (NAT-T).

9.3.4 IPsec AH

The Authentication Header (AH) [26] is the other security protocol of the IP Security suite, with the goal of providing authentication, integrity and anti-replay services but without confidentiality.

Figure 9.7 shows the fields of the AH header.

The functions of SPI and SN are the same as for ESP. The Next Header indicates the protocol type of the packet following the AH header.

Figures 9.8 and 9.9 illustrate the encapsulation formats for IPv4 and IPv6.

The ICV in AH is calculated not just over the content of the IP datagram but also over the immutable fields of the IP header, including the outer header in tunnel mode. So, version, header length, packet length, identification, protocol, source IP address and destination IP address are protected.

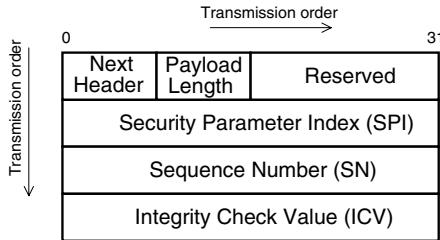


Figure 9.7 AH format.

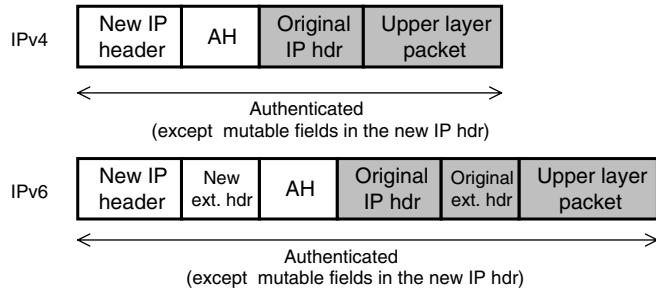


Figure 9.8 AH encapsulation in tunnel mode.

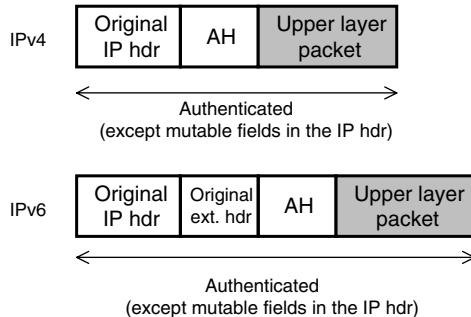


Figure 9.9 AH encapsulation in transport mode.

This capability presents interoperability problems in those networks where the addresses are translated, like with NA(P)T, and which prevent AH from being used it in those cases. Furthermore, there is no NAT traversal technique that allows AH to be used together with NA(P)T.

From the security point of view, it does not have any significant benefit over ESP as ESP can also authenticate the IP payload and in tunnel mode also the inner IP header. In addition, the algorithms used for authentication are the same in ESP and AH. Hence the application of AH in general is rather limited and it is not even a mandatory part of an IPsec compliant implementation.

In the mobile backhaul the only application of AH would be to protect those traffic types which require just authentication. The same can be achieved with ESP, which in addition can be used to provide confidentiality.

9.3.5 IKE Protocol

The Internet Key Exchange (IKE) is the IETF protocol to establish, release and maintain the IPsec SAs. IKE is a point-to-point connection oriented protocol, which creates a secure channel between two trusted peers to exchange the information required to manage the IPsec SA, as well as the IKE itself. This secure channel is the IKE Security Association.

The cryptographic algorithms used by IKE, ESP and AH for encrypting and authenticating the packets are based on symmetric cryptography so the same key is used by the sender and by the receiver to protect the packet. In order to establish a secure channel, IKE is required to create a common key without having prior knowledge of the peer. This is based on the Diffie-Hellman key exchange [24] where the peers exchange their public values and are able to derive a common shared secret which is used to encrypt further exchanges. Figure 9.10 illustrates how the exchange works.

The parties involved in the exchange first agree on which values of p and g are going to be used. Then each of the parties generates a random private value (l, j), calculate the public values (L, J) and send them to the peer. Based on the peer's public value and their own private value, the parties are able to calculate a shared secret K . Given that the prime number p , and the random numbers l and j are large enough, it is virtually impossible for an observer to derive the private values. This problem is referred to as the discrete logarithm problem.

The Diffie-Hellman key exchange can be performed with any peer, so it is also possible for an attacker to intercept the exchange and establish a separate one with each of the peers. To avoid this situation and to guarantee that the peer is legitimate, peer authentication is required once the secured channel is established. IKE supports several types of authentication, being the most common digital certificates and pre-shared keys (see Section 9.3.7 for details).

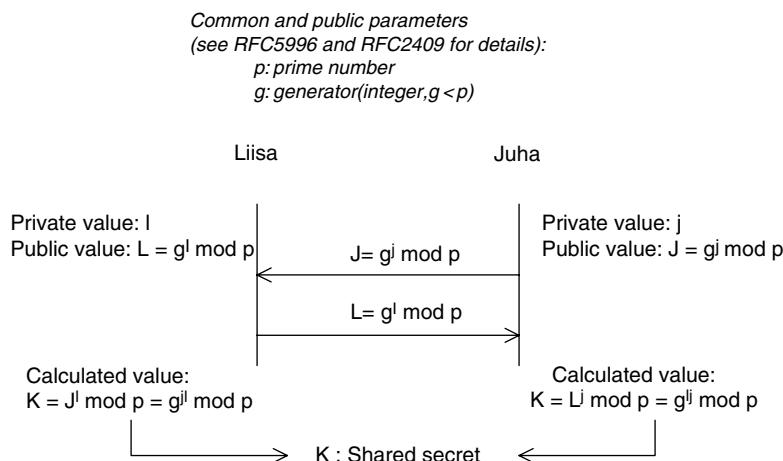


Figure 9.10 Diffie-Hellman key exchange.

Once both peers have a shared secret, further shared keys can be derived by each of the peers to be used by the encryption and authentication algorithms.

Following the establishment of the secure channel, IKE is used to negotiate what type of traffic will be secured by IPsec and what kind of protection is to be applied. Typically the parameters which are exchanged during negotiation are the IP addresses, protocol type and port number of the packet, whether authentication or encryption is used (or both) and cryptographic algorithms for the selected security services. Once the negotiation is completed, the IPsec SAs are created locally by the devices and the packets to be protected are carried over the SA.

It should be noticed that even though the IPsec SAs are unidirectional, IKE always negotiates bidirectional communication so two IPsec SAs (one in each direction) are established.

Currently there are two versions of IKE available:

- IKEv1 [24]
- IKEv2 [33]

IKEv1 specification, published in 1998, is widely implemented as an integral part of many operating systems, and it has been used already for many years in IPsec deployments.

IKEv2 was later introduced by RFC4306 and updated by RFC5996. A number of changes were introduced compared to IKEv1, simplifying the protocol and options, improving the robustness against DoS attacks and at the same time extending the protocol capabilities.

3GPP TS 33.210 mandates the use of IKEv2 for the NDS/IP networks while IKEv1 is optional so it is expected that an increasing number of implementations will support IKEv2 as the main protocol option. However, the number of products and networks which are already running IKEv1 is vastly superior to IKEv2, so the protocol is better known and some of the devices can be reused. Therefore, it could be expected that initial deployments as per TS 33.210 will be based on IKEv1. Eventually IKEv2 should become more common although it might take a long time before it supersedes IKEv1.

There are four exchanges in IKEv2:

IKE_SA_INIT:

- This is the first exchange when an IKE SA is established. The Diffie-Hellman key exchange takes place and the IKE SA parameters are negotiated (Figure 9.11).

IKE_AUTH:

- This exchange performs the peers authentication, and establishes the first Child SA pair (IPsec SA in IKEv2 terminology) (Figure 9.12).

CREATE_CHILD_SA:

- This exchange is used to (re)-establish additional Child SAs and to re-establish the IKE SA (Figure 9.13).

INFORMATIONAL:

- This exchange is used to perform supporting functions such as notifications and error handling (Figure 9.14).

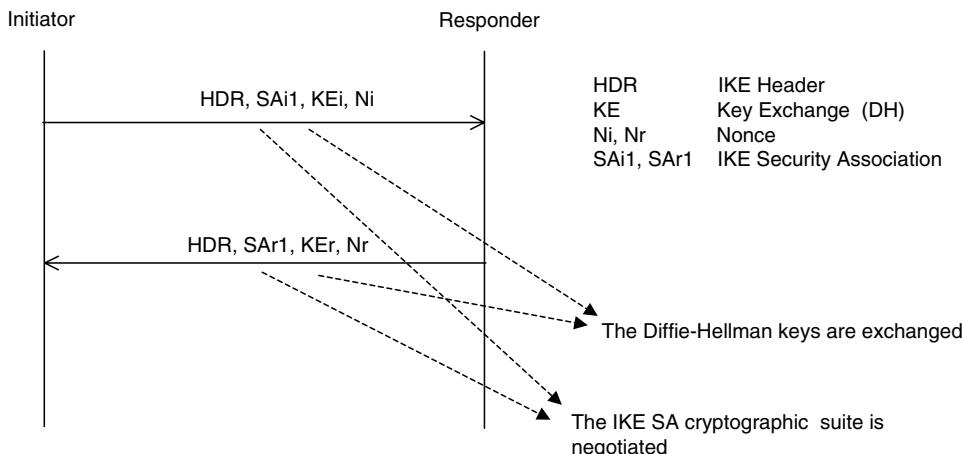


Figure 9.11 Example of IKE_SA_INIT exchange.

9.3.6 Anti-Reply Protection

One of the services supported by the ESP and AH protocols is the anti-replay protection. Anti-replay protection avoids a genuine packet that has been sent earlier from being accepted by the receiver if sent again. Without anti-replay protection, it would be possible to modify the content of a conversation by adding extra information, or in a management connection to interfere with the behaviour of the system by inserting extra commands.

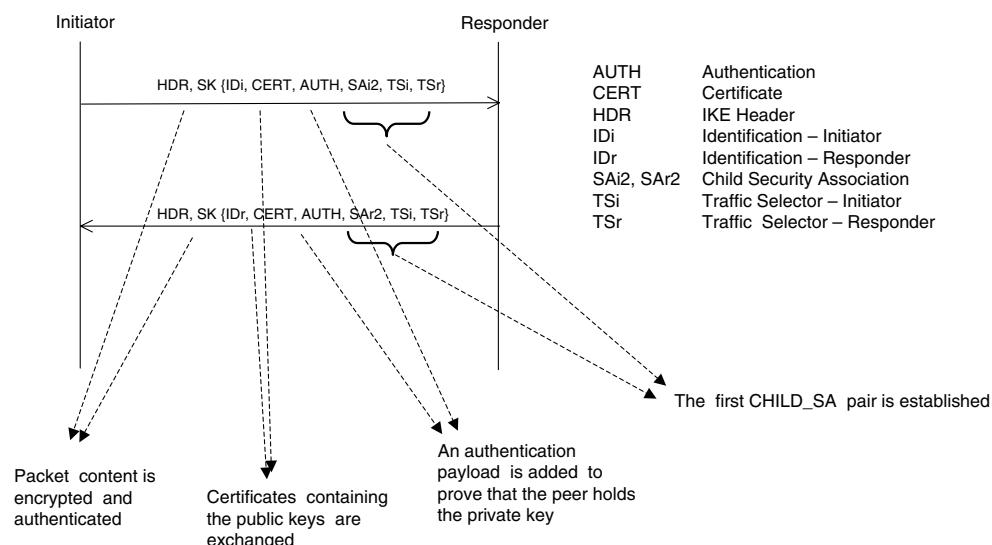


Figure 9.12 Example of IKE_AUTH exchange without EAP.

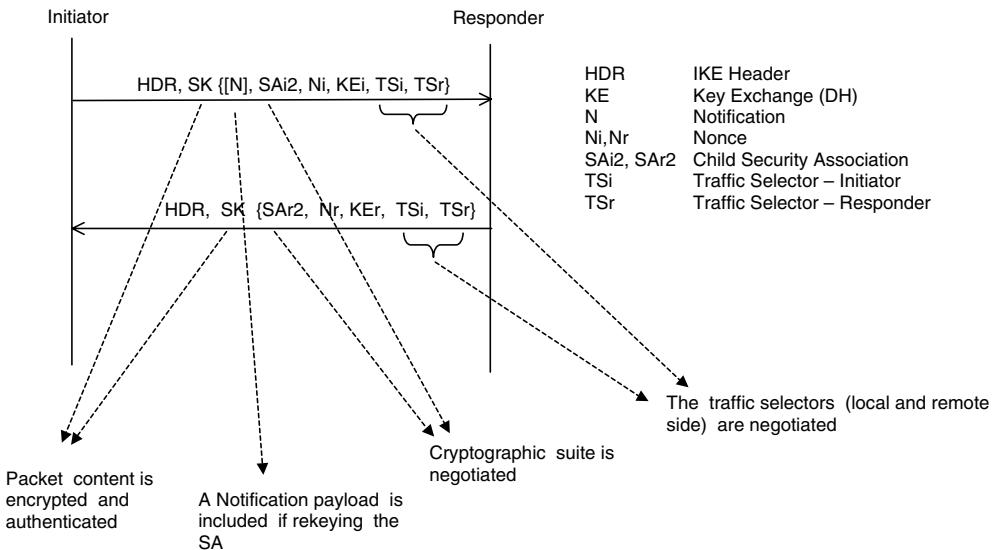


Figure 9.13 Example of CREATE_CHILD_SA exchange for creating an Child SA.

To prevent that kind of attack, each of the packets is assigned an authenticated sequence number which is unique for the whole life of the IPsec SA. The receiver, if it has anti-replay enabled, will check whether that sequence number has been received earlier or not within a window of sequence numbers (the anti-replay window). If the sequence number is within the window limits and it has not been received earlier, the packet is accepted. If the sequence number is ahead of the window, the window slides forward and the packet is accepted. Finally, if the sequence number is behind the window or it has been received earlier, the packet is dropped.

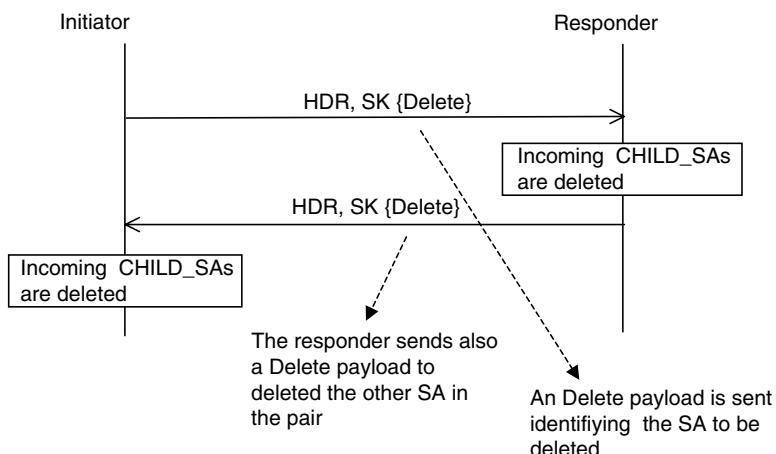


Figure 9.14 Example of INFORMATIONAL exchange to delete Child SAs.

As can be seen, there is a dependency between the size of the window and how far apart the numbers can be in given packet sequence. It is also worth noticing that the security protection is the same for small windows and for large ones so in principle it is advantageous from the system behaviour point of view to use large windows. The drawback of large windows is that additional memory is required.

9.3.7 Network Element Authentication

When a network element is connected to the network and it initiates the communication with another device, there is usually no way for each of the communication partners to know if the peer is who they claim to be. A packet is just sent or received through a network interface, hopefully with the right addresses which provide an indication of who the peer is. However it is not difficult for a man-in-the-middle to impersonate any communication party, even if the channel is secured by means such as a Diffie-Hellman key exchange (see Figure 9.15).

In a closed network, trust is usually granted to any device in the same network domain so there is no need for any authentication (essentially there is no man-in-the-middle threat).

However, if the network is not fully trusted, some authentication mechanism is required for a secure communication to take place.

Mobile networks already use a number of authentication mechanisms for the authentication of the mobile stations. Those mechanisms are usually bound to a SIM card installed in the terminal and the authentication target is therefore the user (the owner of the SIM card) rather than the device itself.

For mobile backhaul, authentication is focused on the devices as there is no concept of separate user (with the exception of Femto system and hosting party, not covered in this book). Installation of the credentials in the devices (BTS and Security Gateways) is under direct or indirect control of the operator and performed at the manufacturing phase by the equipment vendor, during commissioning by the installation engineer or when the device connects to the network.

Two main authentication mechanisms are used in mobile backhaul networks:

- Pre-shared keys
- Digital X.509 certificates

These mechanisms can be complemented by additional ones, if further authentication is required, such as EAP (for IKEv2 only). 3GPP mandates the use of pre-shared keys with the option of digital certificates [3][11] for NDS/IP.

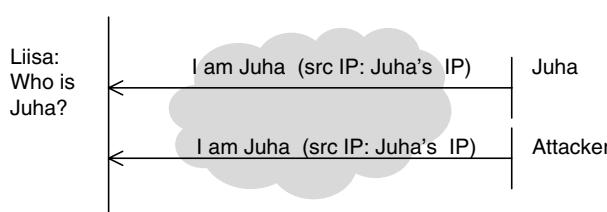


Figure 9.15 Authentication cannot rely just on IP addresses.

Pre-shared keys (PSK) is a simple and well-known mechanism to provide mutual authentication, where the BTS and the security gateway share a secret. Deployment of PSK's would be typically done during commissioning phase.

The secret can be unique for a BTS or it can be shared by a number of BTS's. At the same time, more than one PSK might be required to allow seamless migration when necessary, which quickly increases the number of PSKs in use. The operator needs to keep track of all the shared secrets, and they should be changed periodically according to the security policy to reduce the risk of intrusion. It can be seen that managing PSKs in a large system is complex as the number of keys grows and manual maintenance is required.

- Digital X.509 certificates

An alternative solution for authentication of the mobile backhaul is the use of digital certificates, in particular according to X.509 [34]. Certificate based authentication relies on public-key cryptography [19] as the device to be authenticated is bound to a key pair, the secret key is kept securely in the device and the public key is exchanged with any peer it wants to communicate with. The public key is delivered in a data structure, the certificate, which binds the key to the device identity, and it is signed by a trusted party. During the authentication process, the sender and owner of the certificate will use its private key to sign a string of data, which will be verified by the receiver with the public key and confirm that the sender is in possession of the private key.

Figure 9.16 shows the structure of the X.509 certificates.

The most relevant fields are explained in the list below [19] [34]:

- Version: three version numbers exist (1, 2 and 3). Version 3 is currently used.
- Serial Number: serial number of the certificate which is unique within the issuing authority.
- Signature: contains the identifier of the algorithm used by the issuing authority to sign the certificate.
- Issuer: contains the identity of the signing authority.
- Validity: contains the time interval during which the certificate is valid (unless otherwise revoked).
- Subject: contains the identity of the certificate owner.
- Subject public key info: contains the public key, associated with the certificate owner.
- Digital signature: contains the identifier of the algorithm and the digital signature of the rest of the fields of the certificate, as generated by the issuer.

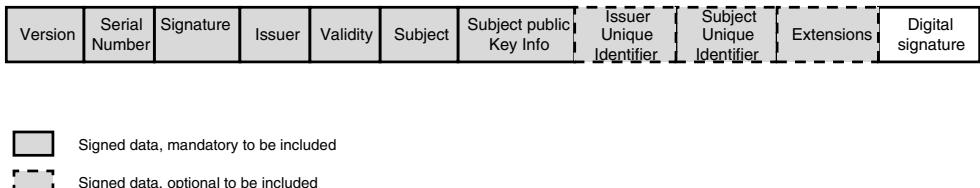


Figure 9.16 X.509 certificate structure.

When using certificates, typically there is only one end entity certificate instance per BTS and one per SEG which are used to establish all the secure connections to other peers. As the certificate is based on public-key cryptography, there is no secret to be shared with other devices so in cases where one of the devices (BTS or SEG) is compromised, that is, the private key is exposed, only the connections for the affected device are affected, but not the whole network.

- Chain of trust

It is important to note that the value of the certificate (and the public key it contains) relies on the trust on the entity signing the certificate. In the simplest model, the signing entity is the trust anchor, which is a party the operator chooses to trust. The trust anchor is referred to as root Certificate Authority (CA). When, for instance, a BTS has to authenticate to a SEG and it sends its own certificate (along with the authentication data), the SEG will use the certificate of the CA that issued the BTS certificate to verify the signature in the BTS certificate and validate it. The process is similar in the other direction.

In more complex models, the root CA does not sign directly the end entity certificate, but instead it signs the certificate of an intermediate CA. Multiple intermediate CAs can exist, the lowest one being the one signing the end entity certificate. Both models suit the case when all the devices that need to establish a trust relationship belong to the same organization as there is one single trust anchor. See Figure 9.17 for an example of trust chain with intermediate CAs.

In this case, the device which has to authenticate the peer requires not only the root CA certificate, but also the certificate from the intermediate CA that signed the peer's end entity certificate.

3GPP provides the possibility of using different intermediate CAs to issue certificates for Network Elements (BTS) and for SEGs. However, in simpler setups the same intermediate CA would be used to issue the certificates of both types of entities, or the intermediate certificate would not be used at all.

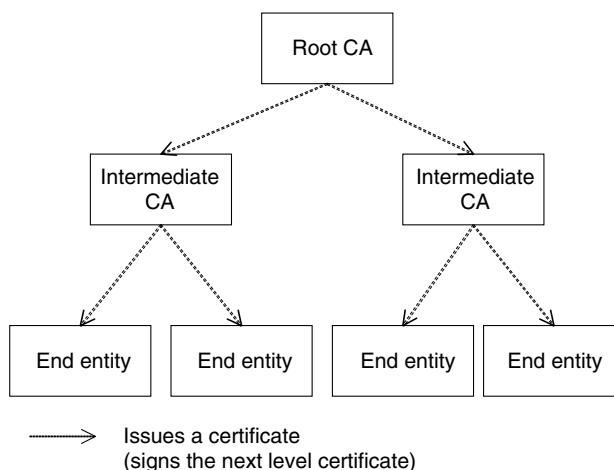


Figure 9.17 Trust chain example, with one level of intermediate certificates.

For those cases where there is a connection with another organization, it is necessary to provide some mechanism to trust certificates issued by the other organization. There are several possibilities available to perform this function, but the one required by 3GPP relies on the use of cross-certification. Each of the organizations will have at least one instance of the so called interconnection CA, a CA which signs the certificates of intermediate CAs in the other organization. Therefore, if two organizations A and B want to interconnect their networks, the connection will be done via two SEGs at the Za interface. The interconnect CA from organization A will sign the certificate of the intermediate CA issuing the certificate of organization B SEG, and vice versa, the interconnect CA from organization B will sign the certificate of the intermediate CA issuing the certificate of the SEG of the organization A. The certificate from interconnect CA from organization A is installed in the SEG from organization A, and the certificate from interconnect CA from organization B is installed in the SEG from organization B, providing in this way the required trust chain. Figure 9.18 illustrates the relationship between the entities.

In the mobile backhaul environment, cross-certification would be typically required when the BTS belongs to one operator and when the controller or the Core Network belongs to a different one.

It should be noticed that other trust models are possible, with multiple CAs and different relationships between them, depending on the requirements of the operator.

- Signing and signature verification processes

The signing process carried out by the CA (for any level of CA) is done using the CA's private key. The process is repeated for each level down to the end entity certificate. The process is depicted in Figure 9.19.

When the end entity has to be authenticated, the signature of its certificate has to be verified as well as the signature of any certificate in the chain of trust. In order to do that, the verifying

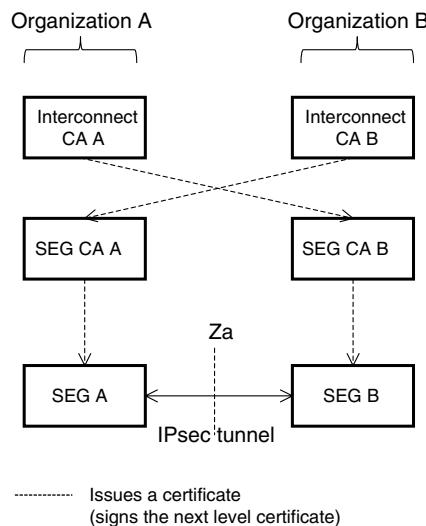


Figure 9.18 Cross-certification for NDS/IP according to TS33.310.

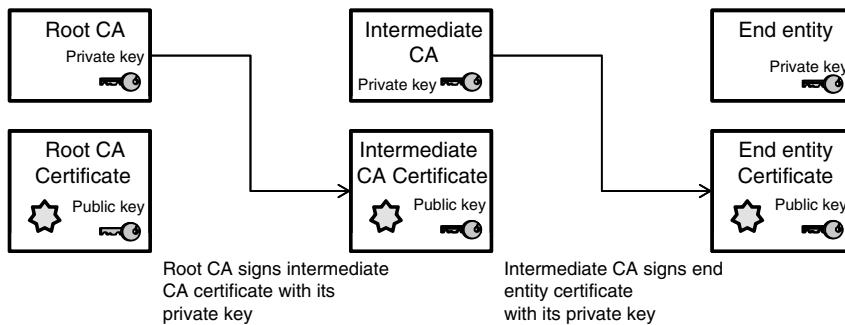


Figure 9.19 Certificate signing process.

entity has to possess all the certificates in the chain, which contain the public keys. The verification starts by the lower level certificate and moves recursively up to the root certificate, using the public keys of each of the signing entities. The process is illustrated in Figure 9.20.

- Certificate lifecycle management

There are two main operations to be performed during the lifetime of a certificate:

- Certificate enrolment.
- Certificate revocation.

The certificate enrolment is required whenever a new certificate has to be issued to the network element. Typically the first enrolment will take place during the initial contact with the operator network, and during the same operation, the CA certificate will be retrieved as well.

Additionally, certificates have a limited lifetime so when the certificate is about to expire, a new certificate has to be enrolled.

Another important element of the certificate management is the revocation. Occasionally it could happen that the private key associated to a certificate is compromised. This means that whoever is in possession of the key is now able to authenticate themselves to the network with the same identity as the rightful owner of the key. Once this incident has been detected, the certificate should be revoked, i.e. it should be marked as not valid anymore as a means of

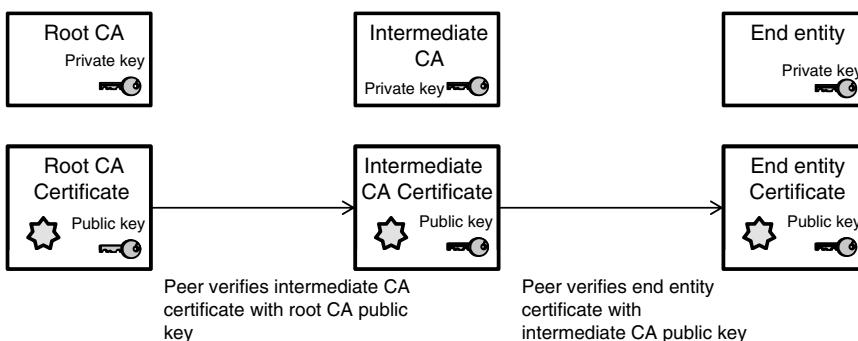


Figure 9.20 Certificate verification process.

authentication. Other possible situations where the operator might decide to revoke a certificate are when the equipment is decommissioned or the trust chain is modified. In those cases, the operator would also be interested in preventing any further use of the certificates. There are different mechanisms to publish the list of revoked certificates, being Certificate Revocation List (CRL) and Online Certificate Status Protocol (OCSP), two common protocols used for this purpose [19].

With CRL the revocation list is downloaded from a server and cached in the network element. The list can be downloaded when the previous one expires, on a periodic basis or it can be pushed. Therefore it is not necessary that the server is available all the time for the CRL to be used. The revocation list contains the serial numbers of all the revoked certificates and it is signed by the same CA which signs the certificates. This mechanism is the one selected by 3GPP.

With OCSP, on the other hand, the network element should request the server for the status of the certificate every time it has to be validated. Only the status of the requested certificates is delivered by the server and potentially it is more up-to-date information than with CRLs. However this largely depends on how often the information is updated in the server itself.

Given that OCSP is an online revocation method, the performance of OCSP is rather dependent on the performance of the network and the server. As the server needs to digitally sign its responses to the client, the delay to handle the requests can be significant. It is also important to take into account the availability of the connection to the server and the server itself so it might not be a suitable mechanism in a network with poor availability [19].

- Public Key Infrastructure as a service

In previous sections a number of entities and services have been introduced which support the use of certificates. They together constitute the so called Public Key Infrastructure (PKI). Various different degrees of complexity have also been discussed, depending on trust relationships and on the offered services.

When the operator faces the challenge of using certificates for authentication, they have to consider among other issues what kind of PKI is required, which services are going to be needed, whether an own CA is needed and who is going to provide the services.

If the operator already has a PKI which is used for other systems, they possess the competence and know-how that enables a new CA to be deployed for mobile backhaul authentication, or to use the existing one. This approach usually offers a better fit to the service requirements as it can be fully customized.

However, in some cases that competence is not available, the network is very small for an own CA to be cost efficient, or for some other reasons, the operator does not want to deploy it. In these cases there are a number of trust providers which offer a range of services suitable to support the mobile network authentication, such as issuing end entity certificates, providing a trust anchor, revocation services, etc.

9.3.8 Firewalls and Access Control Lists

IPsec plays an important role in protecting the mobile backhaul traffic, and due to the traffic filtering imposed by the inbound packet processing it also plays a role in protecting the nodes behind the SEG against DoS attacks, attacks based on malformed traffic, etc.

However, IPsec is not effective against some attacks directed against the IPsec implementation itself, for instance overloading IKE daemon, against attacks based on protocols below IP, such as ARP based attacks, or attacks conducted through traffic which is bypassed by the SEG, such as ICMP. Also IPsec does not inspect the content of the encrypted packet, so if the peer is compromised and malicious traffic is sent through the tunnel, IPsec is unable to filter it.

For the reasons stated above, SEG deployments are typically complemented by firewalls, either integrated in the same device or in a separate one. Some of the functions that the firewall would perform are the following:

- Inspection of ARP packets (to prevent ARP table poisoning).
- Rate limiting.
- Access control list.
- Deep packet inspection (for the clear text packets).
- Stateful filtering.
- Proxies.

Depending on the manufacturer, the firewall can be built into the same appliance as the SEG, or in a different one. While the independent appliance provides more flexibility in the kind of solution, it is an additional device that needs to be deployed in the SEG site, increasing the overall complexity in terms of routing and high availability.

The location of the firewall in respect to the SEG will be dictated by the kind of protection it needs to provide. When deployed at the public interface as a first line of defense, it will protect the SEG and other possible site devices (site routers) against DoS attacks.

When deployed after the SEG, the main role is to protect the core network assets. Packets which have been received through the tunnel or outside the tunnel will be inspected and filtered according to the security policies. The firewall can also analyze higher layer protocols to detect attacks hidden in the user data.

The firewall can also be deployed in interfaces where IPsec is not available or suitable. For instance, in some deployments the BTS can be located in a non trusted environment, or even in public premises with relatively easy accessibility to the BTS hardware. Local BTS interfaces are a possible way to penetrate the device and a firewall can be used to mitigate the risk of intrusion to the device and to the mobile backhaul. In this case the firewall functionality obviously would need to be integrated in the BTS itself, with at least the minimum functionality of packet filtering.

9.3.9 Network Control Protocols Protection

As we have seen, IPsec can be used to protect protocols running on the top of IP. This general rule is difficult to apply in those cases where the protocol operates in broadcast or multicast mode, such as OSPF. There are solutions to make it work (see VPN Resilience section), but in some cases it is better to use the native security mechanism. In particular, OSPFv2 (for IPv4) supports authentication with a simple password, MD5 [23] and SHA1 hashes [31] (note: for OSPFv3, IPv6, only IPsec is standardized).

Another of the protocols used in mobile backhaul that features own protection mechanism is BFD [32]. BFD supports as well, simple password, MD5 and SHA1 protection.

9.4 IP Sec VPN Deployment

9.4.1 Cell and Hub Site Solutions

When designing the VPN and choosing suitable implementation for the cell site, the network designer should take into account the unique requirements of the BTS site.

BTS's are located in a variety of places, indoor, outdoor on rooftops, inside or outside protecting cabinets, attached to building walls or at the antenna mast. They can also be subject to rain or high humidity, extreme temperatures, etc. Most of these site solution and environmental conditions present some challenges to the BTS and how to provide a suitable VPN solution.

Another factor that needs to be considered is the topology of the site, whether the site has only one BTS, multiple BTS's or other site support equipment share the same backhaul links.

Existing standalone SEGs for cell sites are based on small routing devices, with a capacity suitable for a few BTS's. These devices are not usually meant to withstand the outdoor weather conditions so they are only suitable for sites where there is at least a cabinet to shelter the equipment. They also possess enough physical ports and routing capabilities to act as site routers. So they are mostly to be found in bigger sites.

Management of the standalone SEGs can be either done using the appliance vendor specific system or it could be integrated with the overall transport network management system.

For smaller sites with only one BTS or for sites with reduced footprint, an integrated SEG solution could be the most convenient one. An integrated SEG is provided as part of the BTS own implementation, with no extra HW required in the site other than a possible expansion module to cater for the additional processing load. Capacity of the integrated SEG will be lower than for the standalone one, but enough for serving the host BTS and possibly some other site equipment. Resilience in this case would be according to the general BTS high availability concept, as it is an integral part of the BTS.

Management would normally be integrated also to the BTS management system, so it is simpler to perform normal provisioning and maintenance tasks. However, some operator organizations might require that the management of the security appliances is not performed by a mobile network department but by the security department. In these cases the integrated SEG introduces a challenge that would call for either a change in the operator's processes, or a device with split management or alternatively the standalone device would be used.

In the hub site, requirements are completely different to the cell site, with controlled environmental conditions and no practical footprint limitation. On the other hand, capacity and high availability are the main factors to take into account. High capacity SEGs are usually deployed, providing VPN services to a number of controllers and Core Network devices. Several SEGs are also equipped to support high availability, and load sharing configurations would probably be implemented to increase both the capacity and availability. There might be some exceptions for small controller sites (with only one or two controllers) where the VPN termination is fully integrated into the mobile equipment.

Some solutions are built on a dedicated security appliances, while others are built around general purpose routers. In both cases, chassis of different sizes are equipped with application cards or cryptographic accelerators to fit the capacity requirements.

The VPN solution is an integral part of the hub site solution, and the interconnection and interoperability with the other devices, namely the site routers, needs to be taken into account. In particular, it should be ensured that the routing and high availability configuration of the

SEG is compatible with and supported by the site routers. Another aspect to consider is the separation that the SEGs introduce between private and public network. The site routers and the overall site need to be designed keeping in mind that separation, and ensuring that both networks are kept separated at all times. Separation can be performed by means of VLANs, for traffic separation, and Virtual Routing, for routing separation.

9.4.2 IPsec Profiles

3GPP specifies the following profiles for IKEv1, IKEv2 and IPsec [3]:

IKEv1:

Phase 1:

- Pre-shared keys for authentication (note that TS33.310 also specifies the use of certificates).
- Main mode (no aggressive mode).
- FQDN is supported for node authentication.
- Encryption algorithms: ENCR_AES_CBC (128bits key), ENCR_3DES.
- Authentication algorithms: AUTH_HMAC_SHA1_96.
- Diffie-Hellman group 2.
- IKE SA lifetime longer than IPsec SA lifetime.

Phase 2:

- PFS (Perfect Forward Secrecy) is optional.
- Only IP addresses or subnets shall be mandatory.
- Notifications are mandatory to support.
- Diffie-Hellman group 2 (required if PFS is used) is mandatory to support.

IKEv2:

IKE_SA_INIT

- Encryption algorithms: ENCR_AES_CBC (128bits key), ENCR_3DES [29].
- Authentication algorithms: AUTH_HMAC_SHA1_96 [29].
- Pseudo-random function: PRF_HMAC_SHA1.
- Diffie-Hellman groups 2 and 14.
- Optionally, AUTH_AES_XCBC_96 should be used for authentication and PRF_AES128_XCBC as PRF.

IKE_AUTH

- Pre-shared keys for authentication (note that TS33.310 also specifies the use of certificates).
- IP addresses and FQDN are supported for node authentication.

CREATE_CHILD_SA

- PFS (Perfect Forward Secrecy) is optional.

IPsec:

- ESP protocol in tunnel mode.
- Encryption algorithms: null, ENCR_AES-CBC (128bits key), ENCR_3DES [30].
- Authentication algorithms: AUTH_HMAC_SHA1_96 [30].
- The IV (Initialization Vector) should be random and of the same size as the block of the chosen encryption algorithm.

Obviously, this is the minimum set which ensures compatibility between different implementations. However many implementations support a much larger variety of algorithms and key lengths.

9.4.3 *VPN Resilience*

Mobile networks are so widely spread that they support a significant amount of today's voice and data communications. Some of the services offered by the network are critical by the nature of the service (emergency calls) or because of the high revenues they bring to the operator. Those services put tight requirements on the network availability so it becomes paramount for the operator.

Besides, end user quality expectations for some of the services are the same as for wireline services in terms of call breaks and download times. A voice call user will certainly not be willing to wait for tens of seconds for the network to restore the service so the call will be terminated by the user.

Additionally, long breaks can lead to higher layer protocol timers to trigger recovery actions, to unstable network behaviour or to network restarts, which delays even further the service restoration.

From the end user point of view, resilience requirements are largely independent to the radio technology in question. However, the more capacity a technology has, the more important it is to provide a reliable network.

These factors need to be taken into account when defining the availability target of the mobile backhaul and therefore of the security solution. Exact figures are up to the service availability targets defined by operator, but there is certainly a much shorter break tolerance than in traditional data communication networks.

Resilience should be considered for the backhaul link as well as for the network equipment. Backhaul link resilience is discussed in Chapter 7.

Network equipment resilience, focusing in this context on the SEG, is provided by deploying redundant devices in a farm with at least two devices. In case one of the devices fails, any other of the devices in the farm will take over. There are different approaches for the service restoration depending on the capabilities of the devices, as well as targeted service break duration. In the following resilience discussion it is considered that the BTS terminates the VPN. However the same conclusions can be drawn for cell sites with an external SEG.

One possible approach for the service restoration is that once the active SEG is down, the clients (the BTSs) will trigger a recovery action. The BTS would monitor the SEG availability using a mechanism such as DPD, and as the monitoring is done across the backhaul, also the backhaul availability is taken into account. Therefore, this approach can protect against certain

transport failures. Once the failure is detected the BTS would select a SEG from a list of backup devices and would re-establish all the SAs.

On the other hand, this approach presents the significant drawback that the BTS needs first to detect that the active SEG is down, usually by using a mechanism such as DPD. In order not to load the network with excessive monitoring traffic, the detection mechanism is rather slow so the failure detection might take any time from a few seconds up to several minutes depending on the implementation and configured timers. Additionally the re-establishment of the SAs would take additional time, in the order of a few seconds depending on the number of SAs and the performance of the SEGs. Altogether, the outage period is sufficiently long as to have all the voice calls dropped.

Furthermore, it is also possible that upper layers in the BTS would detect that the control plane and management plane connections are down, and a recovery action could be started. One of the common recovery actions is the restart of the BTS, which leads to an extended outage. This scenario is depicted in Figure 9.21.

Outage periods can be dramatically reduced if failure detection is shortened. Instead of relying on the BTS to detect the failure, the SEGs can monitor the availability of each other with a fast polling mechanism, and upon failure detection, initiate the re-establishment of all the connections. Given that the polling is locally performed and only a few devices are monitored, the amount of traffic is not relevant. If the SEGs are configured with the identities (IP addresses) of the BTSs, they will re-establish the VPNs.

However, the SEGs are probably configured without the identities of the BTS (*road warrior* configuration) so that when new BTSs are added, policies do not need to be updated, making the management of the VPN easier and scalability better. In this case, the SEGs are only able to accept incoming IKE requests from the BTSs instead of initiating the connections themselves. Also the backup SEGs do not have previous knowledge of which connections were already established in the active SEG, so they are unable to restore the connections by themselves.

Another approach to resilience is to have two redundant tunnels for each BTS towards two different SEGs in the farm. The selection of which tunnel to use would be done by the BTS based on standard routing techniques, and the monitoring of the availability of the tunnel

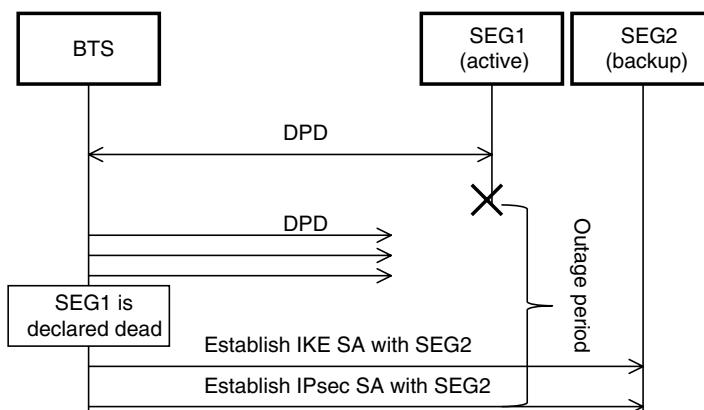


Figure 9.21 Service restoration initiated by the BTS.

would be left to the routing protocol. The restoration of the service would depend once more on how fast the BTS and the SEGs are able to detect that one of the paths is down. Typically routing protocols are not able to detect the failures very fast. However, when they are combined with fast detection protocols such as Bidirectional Forwarding Detection (BFD) [32] the failure detection can be performed in a few seconds or less.

One aspect of this approach to take into account is that many routing algorithms operate by broadcasting or multicasting the advertisement and monitoring packets. While broadcast and multicast is possible with IPsec when the SAs are established manually (via management interface), IKE does not support that possibility so only point to point connections are possible. This limitation can be overcome by using GRE encapsulation on the top of IPsec. In this way IKE only needs to handle the GRE tunnel (which is point to point) while the routing advertisements and monitoring packets travel transparently within the GRE tunnel (see Figure 9.22).

An additional aspect to consider in this approach is that addressing of the BTS becomes a bit more complicated. While in other approaches the traffic endpoint address can be the same as the tunnel endpoint address, in this case they need to be different, so that routing is possible at the BTS. One possible configuration for the BTS addressing is to use network interface addresses as tunnel addresses, and loopback addresses for the traffic.

An additional and convenient approach from the BTS point of view is to rely completely on the SEG to restore the service without any action from the BTS, and with a minimum effect to the end user. We saw earlier that re-establishing the connections by the backup SEG might not be feasible if it lacks the knowledge of which connections were already established. In the stateful failover case, a synchronization connection exists between the SEGs so that the backup SEGs are updated continuously with the state information required to maintain the IKE SAs and IPsec SAs up. They also share virtual IP addresses for tunnel termination. Therefore when the failure is detected, the SAs are shifted to one of the backup SEG and the BTSs are not aware of the failover (see Figure 9.23). The effect on the end service could also be small if the failure detection and the failover are fast enough. The performance of this approach should be expected to be in the range of a few seconds.

To benefit from the stateful failover, the two SEGs should also synchronize their routing state so that they behave as a virtual router. This could be achieved by using HSRP/VRRP. Both SEGs will share the same virtual IP addresses, but only one of them is forwarding traffic. When a failure is detected, both VPN and routing functionalities are transferred to the backup SEG, which will advertise to the network neighbours that the IP address has been transferred.

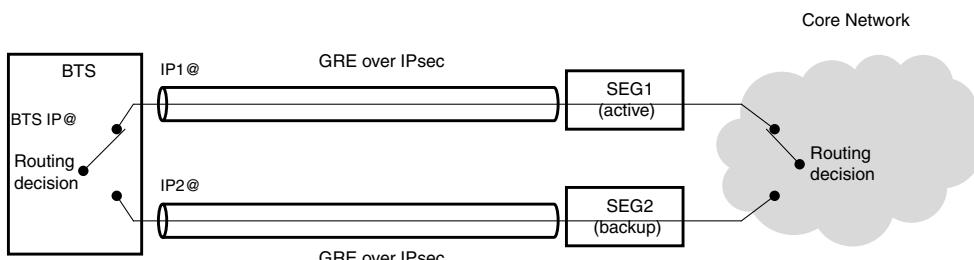


Figure 9.22 Service restoration by means of routing.

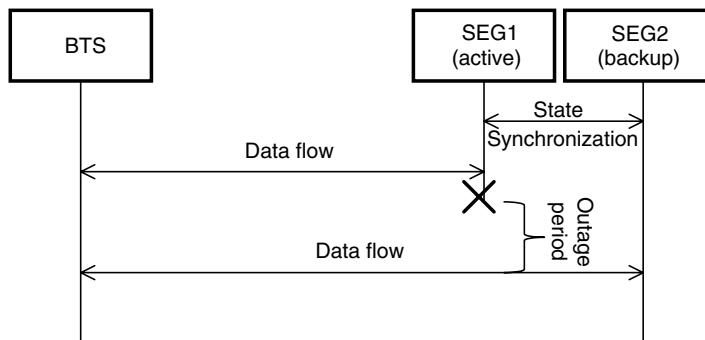


Figure 9.23 Service restoration by using stateful failover.

An altogether completely different approach is not to have a fully redundant system but to mitigate the impact of a SEG failure by sharing the load among multiple devices. The failure of one SEG will put out of service all the BTS's connected to it, but the service could still be offered by neighbouring BTS's. The network capacity would be reduced but it could be acceptable depending on the area to be served. This approach can be combined with any of the other approaches, either for the benefits of load sharing, or to reduce the impact of the failover (see Figure 9.24).

9.4.4 Fragmentation

Applying IPsec in tunnel model implies that clear text IP packets are encapsulated into another IP packet, typically using ESP encapsulation. If GRE is also used, two encapsulations take place. The encapsulation overhead varies depending not only on the protocols but also on the security services (encryption vs. integrity protection), the selected algorithms and original packet size. In any case it can cause the encapsulated packet to exceed the egress interface MTU which would require IP fragmentation.

In general, IP fragmentation should be avoided because of the reassembly effort required at the receiving node, increased network load, packet delay and delay variation, etc. In order to

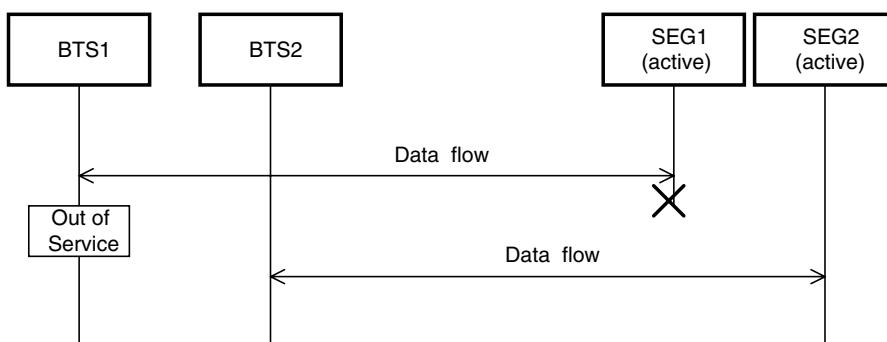


Figure 9.24 Service restoration by using load sharing.

do so PMTUD (Path MTU Discovery) [21][22] can be used, if supported by the BTS's, SEGs and the end nodes. By using PMTUD, the nodes will discover the smallest of the MTUs along the path to the destination and they will be able to adjust the size of the data to be sent to the IP protocol, so that the packet will not experience any further fragmentation along the path.

Unfortunately, PMTUD is not supported by all the implementations or by all the protocols. While PMTUD is an integral part of TCP and SCTP, there is no support in UDP itself, but rather in the applications using UDP.

The MTU of the sources could also be manually configured to ensure that IP fragmentation happens only there, and nowhere else in the path. While this approach is possible in most of the cases, some implementations might not support configurable MTU. It also requires a good knowledge of other MTUs in the path to avoid additional fragmentations.

As can be seen, there will be cases when the IPsec of the SEG or the BTS stack needs to perform fragmentation. When the clear text at the IPsec stack is expected to exceed the interface MTU after encapsulation, the IPsec stack could decide to fragment the packet before the encapsulation (pre-fragmentation) based on pre-defined tunnel MTU. Alternatively, the packet can be encapsulated and fragmented by the IP layer before forwarding (post-fragmentation). Each approach has its benefits and drawbacks which are analyzed next.

- Pre-fragmentation

It has the distinctive advantage that the VPN termination does not need to reassemble the packets before decryption. Only the final destination will perform the reassembly. In this way the VPN termination is offloaded of this resource consuming task. This is mostly important for the SEG if there are multiple devices behind. For the BTS it is less important since typically all the traffic is consumed by the BTS itself, so it still needs to do the final reassembly.

On the other hand, if further fragmentations happen in the public link, the benefits of the pre-fragmentation are void as both types would take place at the same time. To avoid this additional fragmentation, a careful planning is needed, or PMTUD (Path MTU Discovery) should be used.

Pre-fragmentation by the IPsec stack is not compliant with IPv6 as the packets can be fragmented only at the source. In this case, if the IPsec stack needs to fragment, post-fragmentation would be the only possibility.

- Post-fragmentation

As noted above, when the public link MTU is not known or there is routing change, the packet would need to be fragmented by the transit routers. In this case it is better to perform only post-fragmentation since there is no benefit in pre-fragmentation.

Post-fragmentation would also be the only possibility for IPv6.

9.4.5 *IPsec and Quality of Service*

As discussed in Chapter 8, Quality of Service (QoS) is a fundamental concept in today's mobile networks, with multiple services of different types, different customer expectations and a variety of service levels offered by the operators. End-to-end QoS is based on a collection of mechanisms and tools which are deployed in network elements and subnets, and it is

therefore paramount for the overall QoS that each of the components works as planned, together with other network aspects, in particular security.

Packets are classified according to the QoS class and they are marked in the IP header with a DSCP value. This marking is meant to be inspected by the network elements in order to apply a suitable QoS mechanism to the packet. When applying encryption with a protocol such as ESP in tunnel mode, all the information carried inside the tunnel is hidden, including the DSCP. So if the public network is meant to apply a differentiated service to the packet, this value should remain visible. The approach to follow is that the IPsec implementation in SEG and the BTS populates the outer IP header DSCP with a proper value.

The most straightforward way to generate the DSCP is just to copy the DSCP value received in the clear text packet, which is a suitable approach in many cases. However, the encapsulated packet will transit a different network, which could also be a different QoS domain, managed by a different organization or a service provider. This service provider might have a different QoS policy and the packet markings can be different. Therefore the IPsec implementation, to be compatible with the new QoS domain, would need to have a flexible mapping between the inner DSCP values and the outer ones.

It should be noticed that the outer DSCP is not authenticated (for both ESP and AH). This creates the risk that the DSCP is changed by an attacker in order to create a DoS, which is difficult to mitigate.

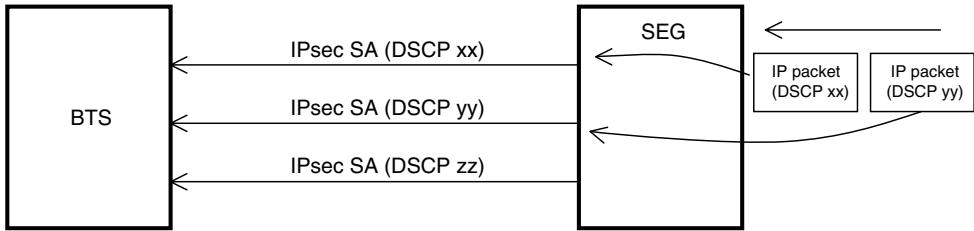
At the receiver side, when IPsec is terminated, the implementation has the option to maintain the inner DSCP as received. This has the advantage that the DSCP can be trusted as it has been authenticated and not altered by any node during transit in the VPN. According to RFC4301, implementations might also choose to use the outer DSCP in those cases when the DSCP space at the receiver and the sender sides are different, and this inner DSCP is not meaningful anymore.

When defining the security policies in the BTS and the SEG, usually they apply to a traffic aggregate, defined by the IP addresses, protocol types and sometimes also port numbers. This traffic aggregate might contain flows with different QoS classes and accordingly they are marked with different DSCP. This is typically the case for the user plane of the mobile network, where real time calls will be assigned to higher QoS class than non real time calls. For other traffic types, such as control plane, all the packets usually belong to the same QoS class.

The traffic aggregate with multiple traffic classes will be carried by a single IPsec SA, and therefore a single running sequence number is used for all the packets in that SA. When the packets are transmitted across the network and arrive at the routes, they can be assigned to different queues as the packets have different priorities, and if congestion happens, the packet order will change within packets of the same IPsec SA. When the packets arrive at the receiver, it will check if they fit within the anti-replay window. Packets with higher priority will probably be at the beginning of the window as they arrive first, and packets with lower priority will be towards the end. If the congestion is sufficiently high for a given window size, low priority packets will not fit in the window anymore and they will be dropped.

In order to avoid the packets being dropped, the first possible solution is to disable the window but it would open the system to possible anti-replay attacks. A more sensible solution would be to use large window sizes. This requires more memory and there could be a performance impact for large windows.

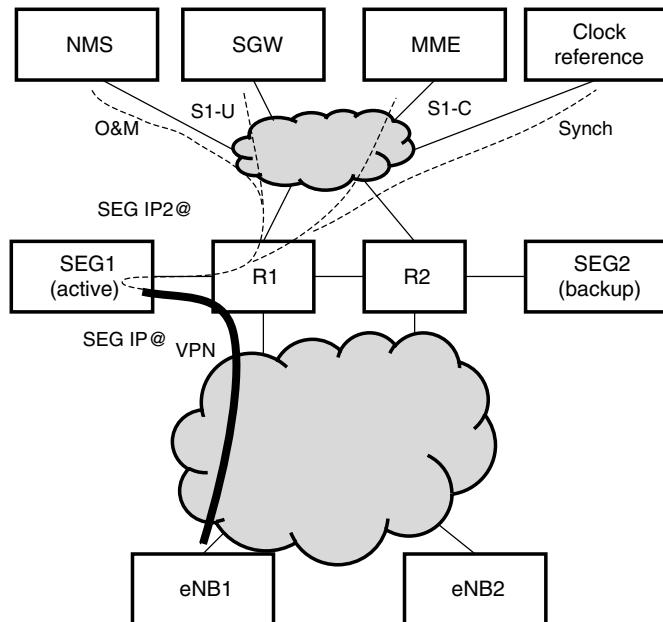
Another solution supported by newer IPsec specification [25] is the use of multiple IPsec SAs with the same traffic selectors for packets with different QoS class. In the egress direction



Note: only one direction is shown. The other direction would be similar

Figure 9.25 IPsec SA selection based on the DSCP.

the IPsec implementation inspects the DSCP of the clear text packet along with the other relevant fields in the header, and maps it to the correct SA (see Figure 9.25). In this way, all the packets within the SA have the same DSCP and re-ordering should not happen. It should be noticed that the DSCP is not negotiated by IKE as it is a local matter in the sender to do the mapping between the DSCPs and the SAs, and therefore all the SAs established for the same traffic aggregate will have the same traffic selectors. On the other hand, these parallel SAs are only supported by IKEv2, not by IKEv1 as it requires the established SAs to have unique traffic selectors (see RFC5996, Section 2.8 for further details).



Note: only the connections for eNB1 are depicted

Figure 9.26 S1 interface, management and synchronization connections.

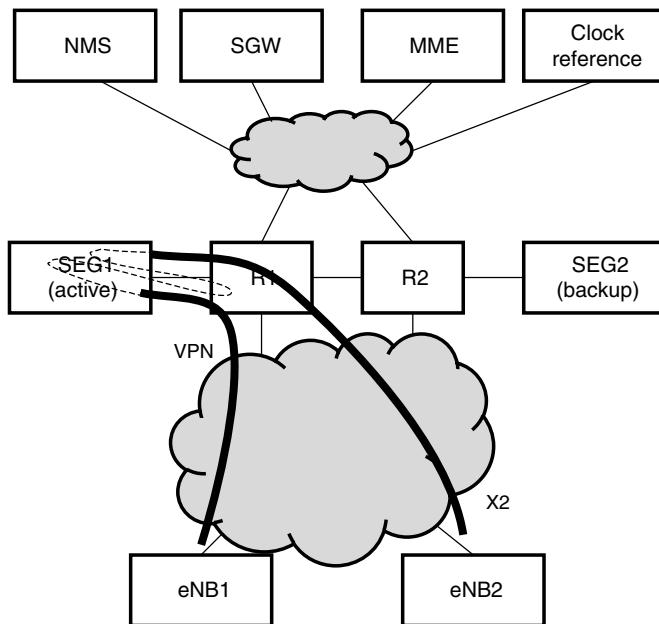


Figure 9.27 X2 interface connections.

9.4.6 LTE S1 and X2 Study Case

In this section, we will look at how the concepts explained previously apply to a practical case. For this case we will consider the LTE system, with two eNodeBs, eNB1 and eNB2, connected through a leased line service to the Core Network site.

In the Core Network site there will be an MME, a SGW, a clock reference (for packet based synchronization) and a gateway towards the management system.

A pair of redundant routers provides the site connectivity, and two SEGs are connected to them in redundant configuration, terminating the VPN tunnels from the eNodeBs. A single tunnel is configured between each eNB and the pair of SEGs.

Figure 9.26 shows how all the S1 traffic from eNB1 is carried over the tunnel, and routed in clear text from the SEG1 to the destination.

The X2 interface, between eNB1 and eNB2, is implemented through the Core Network site routers (star topology). Therefore the packets are forwarded first to the Core Network site, decrypted there, routed, encrypted again and forwarded towards the destination eNB (see Figure 9.27).

A direct connection between the eNodeB's (mesh topology) would also be possible, but it would require additional complexity in VPN configuration as distinct policies for each X2 neighbour would need to be defined. This would be difficult to manage unless an automatic mechanism is used, such as ANR.

For simplicity, a single VPN is created between each eNodeB and the SEG, and traffic is separated for each plane by assigning it to a different IPsec SA:

- User plane (S1-U, X2-U).
- Control plane (S1-C, X2-C).

- Management plane.
- Synchronization plane.

For this case, it is assumed that there is no application level protection for the Management plane, such as TLS. If TLS were used, there is no security reason to protect the Management plane with IPsec as well.

The routers provide a separate virtual routing instance for the VPN interfaces, and others for the Core Network interfaces, effectively separating the routing domains and reducing the risk of misconfiguration.

The device authentication, for both the eNodeBs and the SEGs, is based on digital certificates, which are signed by a CA owned by the mobile operator and located at the management site. Certificate Management Protocol (CMP) is used for the certificate management of the eNodeBs and the SEGs. A simple trust model is used, where the root CA issues directly the certificates of the end entities. Given that all the devices belong to the same operator, there is no need to have cross-certification.

Therefore, each eNodeB will be provisioned with its own device certificate and the root CA certificate, and the SEGs will be provisioned with the SEG device certificate and the root CA certificate. Additionally each device requires the private key associated to its own device certificate. Figure 9.28 illustrates the certificate management architecture.

All the traffic exchanged between the eNodeBs and the SEGs receive the same kind of protection, and it is encrypted, authenticated and protected against replay. Even if not all the traffic types might require confidentiality protection, it makes configuration easier as there are less policies to handle. As indicated earlier, each traffic type is in its own IPsec SA and one security policy is defined for each SA. This allows a dedicated running Sequence Number

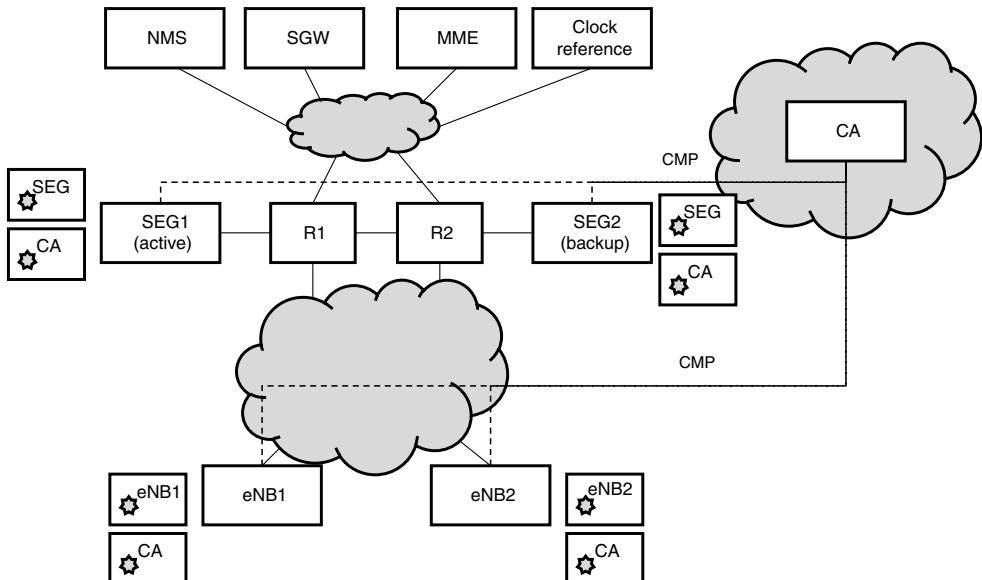


Figure 9.28 Certificate management interfaces.

Table 9.1 eNB1 security policies.

Policy no.	Local	Remote	Local GW	Remote GW	Action
1	eNB1	NMS	eNB1	SEG	Protect
2	eNB1	Clock ref.	eNB1	SEG	Protect
3	eNB1	SGW	eNB1	SEG	Protect
4	eNB1	MME	eNB1	SEG	Protect

Table 9.2 eNB2 security policies.

Policy no.	Local	Remote	Local GW	Remote GW	Action
1	eNB2	NMS	eNB2	SEG	Protect
2	eNB2	Clock ref.	eNB2	SEG	Protect
3	eNB2	SGW	eNB2	SEG	Protect
4	eNB2	MME	eNB2	SEG	Protect

and anti-replay window to be had, which mitigates the effect of packet re-ordering, in case it happens. The security policies are defined based on the IP addresses only. Tables 9.1 and 9.2 show the policies for the eNBs.

Table 9.3 shows the SEG policies. Note that the SEG policies do not contain the eNB addresses in order to simplify the commissioning (additional eNB can be added without changing the SEG policies).

IKEv2 is selected for the key management. The selected algorithms are AES-CBC with 128bits for confidentiality, HMAC-SHA1 for the hash and Diffie-Hellman group 2. Same algorithms are used for IKE. Other parameters are also according to the 3GPP IKEv2 profile.

Resilience is provided by a redundant pair of SEGs, with stateful failover. The two SEGs are seen by the eNBs as one single node, as they have the same addressing IP@ (virtual) and the same credentials (certificate). The SEG resilience relies also on the fact that the site routers are also redundant. The SEGs also present a virtual address IP2@ towards the Core Network. Figure 9.29 shows the addressing used by the SEGs.

In case SEG1 fails, the SEG2 will take over and advertise to the R1 and R2 routers that it now has the addresses IP@ and IP2@. This advertising will happen usually by using Gratuitous ARP. After that, the traffic from the eNBs and from the Core Network devices is rerouted by R1 to the SEG2. No disruption would normally be visible to the eNBs or to the Core Network elements except a few lost packets. Figure 9.30 depicts the traffic path after the failover, showing how the addresses IP@ and IP2@ are now active in SEG2.

Table 9.3 SEG security policies.

Policy no.	Local	Remote	Local GW	Remote GW	Action
1	NMS	Any	SEG	Any	Protect
2	Clock ref.	Any	SEG	Any	Protect
3	SGW	Any	SEG	Any	Protect
4	MME	Any	SEG	Any	Protect

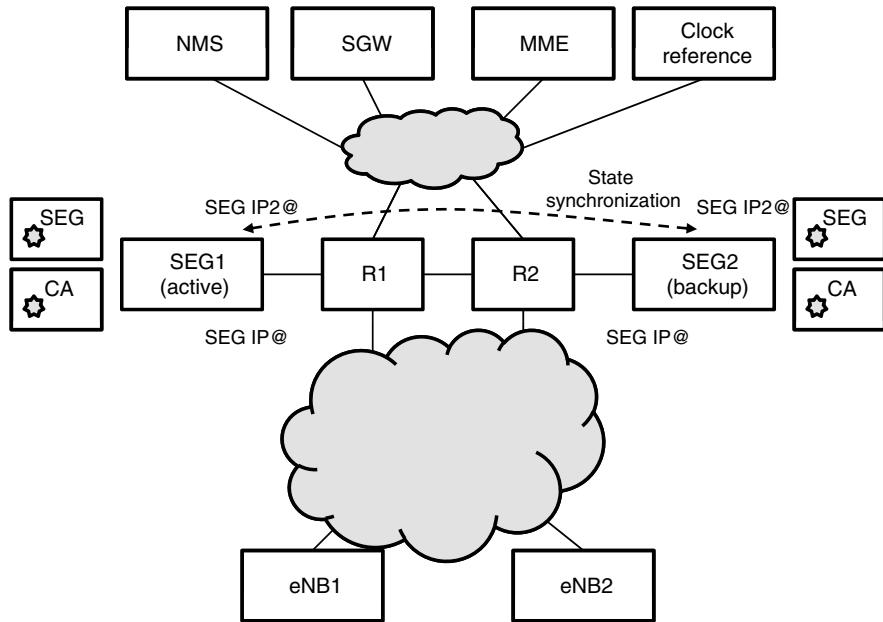


Figure 9.29 Stateful failover.

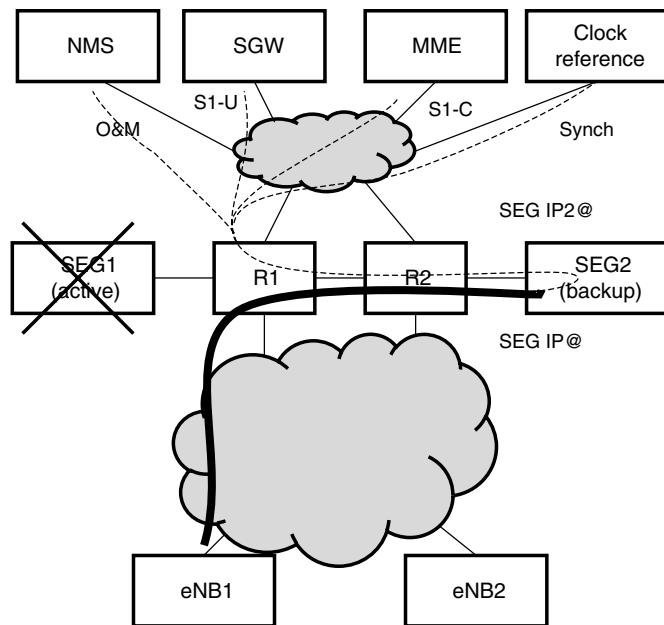


Figure 9.30 Stateful failover after the failure.

9.5 Summary

A packet mobile backhaul is based on an open and well-known protocol, IP. Even though the backhaul is a separate, private network (not directly connected to the public internet), it is still vulnerable to many threats that do not exist in a TDM or ATM network. Introducing IP based logical interfaces and IP/MPLS/Ethernet technologies for the backhaul necessitates that these threats are identified and properly addressed.

In many cases, 3GPP requires an implementation of cryptographic protection using IPsec. For LTE, 3GPP specifications for the network domain are clearly written and explicit. For 2G and 3G, guidance from the more recent LTE specification work can be used.

The backhaul security is not only about IP layer and IP layer protection with IPsec. Other layers and other types of threats need to be considered as well. Many of the topics of this kind are addressed by sound design guidelines and operating practices.

IPsec VPNs are deployed for a cryptographic protection for the traffic carried in the mobile backhaul. IPsec supports encryption, authentication and confidentiality. Typically IPsec is implemented between a BTS (as BTS integrated IPsec function, or as a separate cell site gateway), and a central site IPsec GW. With the IPsec VPN, high availability of the IPsec GW is necessary, since a number of BTSSs depend on it. Different models for the resilience exist that depend on the IPsec GW implementation.

ESP in tunnel mode is the selected IPsec protocol to provide the protection, which is able to support all of the required security services. IKEv2 (and IKEv1 for interworking with legacy equipment) is used to handle key management and control the establishment and release of IPsec SAs. As part of the IKE exchanges, the peers must authenticate each other. The authentication can be based on PSK, or preferably on X.509 digital certificates as they provide a better scalability. IPsec affects other system areas as packet fragmentation and Quality of Service, and the system designer needs to take those effects into account when planning an IPsec VPN for the mobile backhaul.

References

- [1] 3GPP TS33.120 Security Objectives and Principles, v4.0.0.
- [2] 3GPP TS33.102 3G Security; Security architecture, v10.0.0.
- [3] 3GPP TS33.210 3G security; Network Domain Security (NDS); IP network layer security, v11.2.0.
- [4] 3GPP TS21.133 3G security; Security threats and requirements, v4.1.0.
- [5] 3GPP TS33.401 3GPP System Architecture Evolution (SAE); Security architecture, v10.2.0.
- [6] 3GPP TS36.300 Evolved Universal Terrestrial Radio Access (E-UTRA), Overall description, v10.5.0
- [7] 3GPP TS43.051 GSM/EDGE Radio Access Network (GERAN), Overall Description, v10.0.0.
- [8] 3GPP TS25.401 UTRAN overall description (Release 10), v10.2.0.
- [9] Niemi, Nyberg: UMTS security. Wiley, 2004.
- [10] Forsberg, Horn, Moeller, Niemi: LTE Security. Wiley, 2010.
- [11] 3GPP TS33.310 Network Domain Security (NDS); Authentication Framework, v10.1.0.
- [12] Vyncke, Paggen: LAN Switch Security: What Hackers Know About Your Switches. Cisco Press, 2007.
- [13] IEEE 802.1X-2010, IEEE Standard for Local and metropolitan area networks. Port-Based Network Access Control
- [14] IEEE 802.1AE-2006, IEEE Standard for Local and metropolitan area networks. Media Access Control (MAC) Security.
- [15] IEEE 802.1AR-2009, IEEE Standard for Local and metropolitan area networks. Secure Device Identity.
- [16] MEF 6.1 Ethernet Services Definitions Phase 2
- [17] MEF 10.2 Ethernet Services Attributes Phase 2.

- [18] MEF 22 Mobile Backhaul Implementation Agreement (2/09).
- [19] Adams, Lloyd: Understanding PKI. Second Edition, Addison Wesley, 2003.
- [20] IETF RFC 1191 Path MTU Discovery.
- [21] IETF RFC 1981 Path MTU Discovery for IP version 6.
- [22] IETF RFC 2328 OSPF Version 2.
- [23] IETF RFC 2409 The Internet Key Exchange (IKE).
- [24] IETF RFC 3948 UDP Encapsulation of IPsec ESP Packets.
- [25] IETF RFC 4301 Security Architecture for the Internet Protocol.
- [26] IETF RFC 4302 IP Authentication Header.
- [27] IETF RFC 4303 IP Encapsulating Security Payload (ESP)
- [28] IETF RFC 4306 The Internet Key Exchange (IKEv2) Protocol
- [29] IETF RFC 4307 Cryptographic Algorithms for Use in the Internet Key Exchange Version 2 (IKEv2)
- [30] IETF RFC 4835 Cryptographic Algorithm Implementation Requirements for Encapsulating Security Payload (ESP) and Authentication Header (AH)
- [31] IETF RFC 5709 OSPFv2 HMAC-SHA Cryptographic Authentication
- [32] IETF RFC 5880 Bidirectional Forwarding Detection (BFD)
- [33] IETF RFC 5996 Internet Key Exchange Protocol Version 2 (IKEv2)
- [34] ITU-T, 'Information technology – Open Systems Interconnection – The Directory: Public-key and attribute certificate frameworks', X.509, August 2005.

10

Packet Backhaul Solutions

Erik Salo and Juha Salmelin

Mobile traffic growth, especially the data traffic growth, makes transition to packet based MBH networks indispensable, as neither TDM nor ATM technologies can in the longer term offer economic solutions for the high transport capacities needed. However, there is no single packet based MBH solution that fits all, and there are many possible evolution paths to get to a desired network solution.

Many technical alternatives and different requirements, often contradicting requirements, need to be taken into account in creating and optimizing a packet based MBH solution, and in formulating its evolution steps. Most of the attention shall be directed to the MBH network lower tiers, to the MBH access network and even to its lower parts, as most of the backhaul costs lie in the lower access networks.

10.1 Creating a Packet Based MBH Solution

Finding the most suitable MBH solution for each mobile network and for all its coverage area is a multidimensional optimization task. When going over to packet based solutions, optimization of the MBH network needs to be considered from many different view points, including:

- economic optimization (investments and their timing, network opex, cash flow etc);
- technical optimization (enough capacity, proper QoS, reliability and security etc);
- optimization for a particular operator (goals and strategy, resources, cost structure etc), or for a shared use of two or more operators;
- optimization for a certain region (infrastructure, cost level, outsourcing options etc);
- optimization for flexibility (network plans or service targets or other goals change);
- project optimization (practicality of plans, implementation, time tables etc).

The best or the most practical MBH solution is hardly ever exactly the same for any two cases when all the differences in the starting points and goals are taken into account. In different environments (e.g. dense urban vs. rural, or ample existing transport infrastructure

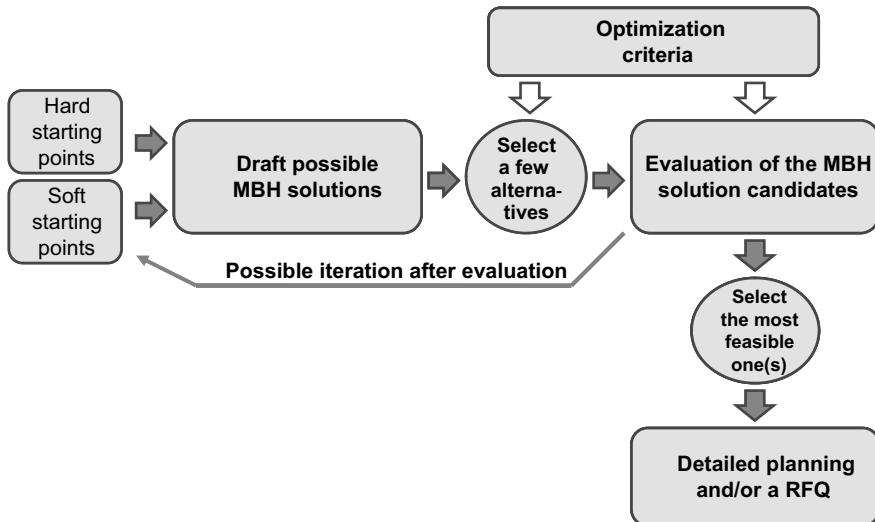


Figure 10.1 Creating a packet based MBH solution.

vs. no infrastructure at all) quite different MBH solutions are needed. Optimal solutions are also clearly quite different for very different mobile networks and different traffic volumes or profiles (e.g. 2G + 3G or advanced 3G or 3G + LTE).

However, *in similar environments and in similar mobile network development phases* optimal MBH solutions have a lot of commonalities. Often these solutions can be based on the same technologies and on the same equipment families; then the differences are for example in the MBH roll-out speed, in equipment capacities and deployment quantities and phasing, in overlaying or replacing existing networks and division between using in-house MBH facilities or outsourcing.

Creation and definition of a packet based MBH solution starts with a careful study of the starting points, e.g. existing MBH networks, existing mobile networks, (data) traffic growth forecasts and network expansion plans as well as targets for mobile services and mobile operator's other goals (e.g. due to competition or license obligations). Starting points also include existing in-house resources for MBH planning, building and operation and available external services, e.g. for planning or operation MBH networks.

When a good picture about the starting points is available different technical solutions can be drafted. After that, first comparisons can be made between the possible solution types and directions, and a few MBH solution types and network evolution approaches can be selected for closer study for a particular MBH case. Then a few possible MBH solutions are defined more exactly and these are compared for technical, economic, strategic and implementation related merits and challenges. Finally one (or two) is selected for the detailed planning and/or to draw up a request for quotation (RFQ) for this solution; an overview of the solution creation process is shown in Figure 10.1.

10.2 MBH Solution Starting Points

'Hard' starting points consist of things that cannot be changed in the short term, including things already existing, such as the existing networks and their environment, but usually also

include the legal/regulatory environment. ‘Soft’ starting points are things that are given for the MBH planning but can still be discussed, and even changed if there are very good reasons; examples are mobile network (expansion) plans and MBH network targets and goals. Many estimates about future conditions and traffic and other forecasts can also be considered to be ‘soft’ starting points, as they can only represent a view (a consensus view or an expert view) about the future.

10.2.1 Hard Starting Points

For a new MBH network or for a MBH extension, the most obvious starting points are the mobile operator’s existing networks (existing mobile, MBH and other transport networks) which are usually well known and documented. Also the present mobile customer base (type of customers and their present network usage) and present measured traffic in the existing network(s) are important starting points.

But also the other operators’ networks in the same area (especially their MBH and other transport networks) may have a significant influence on what kinds of MBH solutions are feasible. These other networks are usually significant from a competition point of view – they are competitors’ starting points for their network expansions – but often they can also mean good opportunities for network sharing and/or outsourcing.

Also many things related to network environment and service area are fixed, at least in short term, often also in medium term. Such factors include geography and demography (e.g. size, distribution, density and growth of population) of the area, as well as the general infrastructure available in the network area (e.g. buildings, roads and availability of electricity).

In addition these ‘physical’ hard facts, other things limiting solution alternatives may be unchangeable in the short term, especially all kinds of regulations and license requirements and conditions. For example, the possibility of laying cables for new transport routes and the cost of these routes often depend heavily on local (authority) permits needed and the conditions set in those permits. Mobile operators’ own license requirements usually relate only to the mobile network, more rarely to the MBH network; but the mobile network requirements obviously can influence the MBH network a lot (e.g. coverage requirements in low-density areas means longer transport links with not-so-high capacity requirements but with very high cost pressure). On the other hand, frequency band regulations, channel allocations and license requirements and fees may directly affect feasibility of microwave radio based MBH solutions.

10.2.2 Soft Starting Points

The ‘soft’ starting points consist mainly of various kinds of forecasts and plans and operator longer term goals (strategies). Most obvious starting points for the MBH solution work are the plans concerning mobile network expansions and phasing of mobile network building in those plans.

Forecasts about future mobile traffic, which may or may not be included in those plans, are very important for the proper planning and dimensioning of the MBH network. As discussed earlier in Chapter 2, forecasts for individual services are not so important for the MBH network design, but instead characterization of the mobile network traffic as a whole. Forecast for the busy hour total traffic (per base station site) is, as in earlier networks, still very important.

The share of real time streaming of the total traffic is also important, particularly if it is big, as statistical multiplexing does not influence this part of the traffic as much as other traffic. In the high bandwidth packet networks additionally forecasts (or the target values set by the operator) about real single user peak rates are very significant, especially for MBH lower access parts, where these peak rates often determine the needed transport capacity.

General goals for new MBH networks may also be seen as part of ‘soft’ starting points; for example, there may be targets to reduce the power consumption of the network, or make it more remotely manageable so that site visits will become increasingly rare events.

In the ‘soft’ starting points one may also include forecasts about development of the other networks and infrastructure in the area. Future development of other (transport) networks influences possibilities for sharing and/or outsourcing of connections, and the abundance or scarcity of transport options strongly influences expected price levels.

In addition to the operator network plans mentioned above, some points from the mobile operator’s strategies may be worth noting in the MBH solution work. Quite obvious things to take into account are the general longer term targets for in-house vs. outsourced transport (with related targets for in-house resources and skill sets). In addition it may be very useful to pick up some points from mobile operator’s competition strategy: how this mobile network is intended to be better than the competition in the area? If the target is, for example, to offer better end-user service quality (e.g. higher peak rates and/or lower latencies), the MBH network also needs to be designed to support realization of these targets. If the target instead is to be significantly cheaper than competition, quite different priorities may be applied in the design.

10.3 MBH Optimization Considerations

10.3.1 Economic Optimization

Economic optimization of the packet based MBH network is always a very important point of view. In general it does not mean optimization for the lowest possible cost (sometimes this may be the case) but for a high cost-efficiency and making timing of expenditures suitable for the mobile operator’s economy. Therefore phasing of the MBH network building is also an important consideration. In addition, network running costs or OPEX also needs to be addressed – with new power-efficient and remotely manageable equipment significant savings are possible both in electricity bills and in network operations.

Comparable cost-efficiency values for different packet based MBH solutions are not necessarily simple things to define, as they should include – in addition to the cost of a certain network solution – also the achieved performance, e.g. network capacity, throughput, latency and reliability.

Possible use of leasing or network out-sourcing for the MBH implementation is also an important economic consideration – it will change some of the shorter term capital expenditures (capex) to medium or longer term more continuous operational expenditure (opex). This makes cost comparisons of MBH solutions more complex: net present values (NPV) or total costs of ownership (TCO) over several years for different solutions need to be compared. However, these types of calculations are dependent of many forecasts and assumptions (e.g. future leasing fees, interest and discount rates) and thus necessarily contain more uncertainties than just comparing capex for different network implementations.

10.3.2 Technical Optimization

Optimization of a packet based MBH network for technical performance requires consideration of several technical parameters simultaneously – some of them can be the most important in one MBH case, some others in another MBH case, depending how the MBH network goals are set.

Much of this optimization goes as for any packet network (available capacity/throughput and expected delay/latency for different traffic classes, network reliability/resilience etc), taking into account the different goals and weighting of these factors in access, aggregation and backbone tiers of the MBH network. (Minimum capacities or dimensioning of the packet based MBH network links is briefly discussed in Section 2.3.3 ‘Backhaul efficiency improvements’).

In addition, there are some MBH specific requirements that need to be fulfilled. The most important special requirements are related to the transfer of synchronization, when it is carried over the packet network – such requirements generally do not exist in other packet networks, and therefore they need special attention, especially when a multipurpose or multi-service packet network is used also for MBH traffic. Also the delay requirements set by the mobile network (round trip delay, RTT, from a controller to the base station and back) may be shorter than otherwise required in a multiservice network.

10.3.3 Optimization for a Particular Operator

This is already partially covered above, but it is important to remember that part of the MBH requirements are operator specific. They depend on mobile operators’ business strategy and competition strategy, operators’ organization and cost structure and are related to what other businesses the same operator is possibly running (e.g. a transport business).

Operator organizational structure and responsibility allocations are worth considering here as well: very strict division of responsibilities into different departments may mean putting e.g. on base station sites seemingly unnecessary parallel functionality and equipment – it may be needed by those departments for doing their jobs independently from each other, but shared equipment would be a much better solution from the total cost optimization point of view (and sometimes just sharing of data, e.g. network quality monitoring data, could be enough for avoiding parallel functionality).

Also, some very important technical parameters, in spite of various guidance values and recommendations, are finally based on the mobile operator’s own business decision. Perhaps the best example of this is the guaranteed single user peak rate value within an area of the mobile network (operator decision obviously within the technical capabilities of the mobile network) that has a significant effect on the dimensioning of the MBH access links, and thus can have a very significant impact on the MBH access network cost in that area.

Another important point is the possible use of the same packet network for other purposes – if the same operator is also carrying (legacy) fixed services over the same packet network, their traffic matrix can be much more mesh like (not a logical star as in MBH), and the shared packet network may require a different kind of optimization to a pure MBH network.

10.3.4 Optimization for a Certain Region

Optimization for a region means that the existing local infrastructures or lack of them, and local cost levels are taken into account early in the solution work. Sometimes even the local security conditions play a role in the solution selection (e.g. if unguarded equipment or more expensive material on the cell-sites or cables tend to disappear).

Local conditions may, for example, dictate building of one's own transport network, if there is no suitable network available in the area and thus no leasing or outsourcing possibilities. Or local conditions may strongly favour wireless transport solutions for the MBH network if there are no fibre cables and installation of new ones is difficult and/or very much time consuming. On the other hand, in some other areas low local digging cost may make new cable routes very feasible.

10.3.5 Optimization for Flexibility

Optimization for flexibility is here understood to mean ability of the network design to cope with the changing requirements in the middle of the implementation or immediately after that. In the fast developing mobile telecommunication market it is not so uncommon that the operator targets and plans change quickly and new unexpected requirements or goals are put on the MBH network and its design – an example can be a new co-operation agreement between two operators, or merger of the companies. Also, changes on a smaller scale, e.g. very much intensified competition in an area, may require fast changes in the targets and design of a mobile network and thus also those of the MBH network.

In network design this kind of flexibility means that from a certain network implementation phase there should be several different ways to continue, not just one predetermined way, so that in cases of changing requirements the design of the next steps can also be relatively easily changed. In practice this kind of network flexibility is often neglected, as it may be difficult to include in the plans and also because it can be partly contradictory to a very tight techno-economic optimization for the given targets. However, if and when this kind of requirement and target changes happen, all flexibility included in the network design can be very valuable.

10.3.6 Optimization of Implementation

Implementation considerations are an important final step in solution optimization. Phasing of the building of the new capabilities and capacities, connecting new equipment and links into service and possibly taking some others out of service need careful consideration, as well as all project work practicalities; making a draft project plan relatively early is useful for checking that the planned solution can also be smoothly implemented.

For example, if the same packet network will be used for fixed (legacy) services and for the MBH traffic, the total traffic matrix may be very complex; then it may be tedious and very labor-intensive to move traffic in small steps from the legacy network into the new packet network and to take the legacy network out-of service piece by piece, as could be required in a gradual replacement strategy – the operative costs during a relatively long transition phase could be excessive. In such a case, it may be more cost-effective to leave the existing legacy running as it is until all traffic has been moved, and then close it down in a single step.

10.4 MBH Solution Alternatives

When the hard facts about the environment and existing networks are known, as well as the goals of the new networks or network expansion, it is time to look at available solution options. A MBH network for a given mobile network can usually be designed in many different ways, even using quite different technologies and network strategies.

Therefore the first task is to limit the number of options by defining the main lines of the MBH solution – for example, whether an existing network is expanded with some functionality enhancements (minimizing disruption), or whether it is time for a wholly new network based on new technologies and equipment (minimizing ‘burden of past’).

Often the new packet based MBH network needs to be planned for significantly higher capacity than the existing MBH network, due to high mobile traffic forecasts already for the life-time of the first phase of network. For example, mobile traffic forecasts for an area for the next 2...5 years may mean that the MBH capacity should be 4...10 times higher than the present network. In such cases the present network offers mostly the infrastructure (sites and routes and possibly cables/fibers) for the new network, but otherwise the new packet based MBH network needs to be built more or less independently of the equipment presently used in the MBH network.

Another very important consideration is whether the new MBH network will be wholly based on the mobile operator’s own facilities and equipment, or smaller or bigger parts of the network will be based on leased connections/outsourced transport services. Whether or not to use outsourced transport obviously strongly depends on what the offerings are in the area in question, what their present and forecasted future price levels are and how reliable those services are estimated to be. Selections between in-house facilities and outsourced transport services are among the most important decisions to be made in the MBH solution creation, as these decisions have not only immediately big economic impact (e.g. capex or opex weighted cost structure), but also very significant long term influence on operator organization, and on strategic options available for the following network building phases.

In the next subsections different approaches for building packet based MBH based on in-house equipment are discussed, with the main division between enhancing existing MBH to better handle packet traffic (Section 10.4.1 and 10.4.2) or taking a more disruptive approach and starting to build fully packet based and packet optimized network parts (Sections 10.4.3...10.4.6); in practice a good solution may also be something between these cases. Then the following section 10.5 discusses the important question about the role of leasing and outsourcing as a part of the packet based MBH solution.

10.4.1 Enhancing SDH/Sonet Networks with NG-SDH/MSPP Equipment

Perhaps the least disruptive and easiest way to improve packet traffic carrying capability of the existing MBH network is to add new nodes and replace some existing ones based on existing technology and product family, but with added packet switching capabilities. Then in the first phase for example only some of the existing SDH nodes need to be replaced or upgraded to be NG-SDH/MSPP nodes – this type of equipment is described in Chapter 5 (Section 5.2.5).

The MBH network efficiency for packet traffic can be greatly improved in this way, especially in the network parts where statistical multiplexing can work effectively, i.e. in network parts where traffic flows from several base stations are already combined; thus the

main application area could be the MBH aggregation networks and upper parts of the MBH access networks.

In this solution the SDH network management (with upgrades for NG-SDH/MSPP nodes) can still be used for the upgraded network, and thus it can remain under a single NMS system. Also one of the benefits of SDH network, easy distribution of synchronization towards the base stations, is kept in this phase and consideration of the packet based schemes can be postponed to a later phase.

The main questions to be considered here are whether this improvement is big enough (compared to mobile traffic forecasts), and after this first step, how packet traffic capacity extension continues in the future. When new transport nodes with packet switching capabilities are added (or present ones upgraded with new units) and this is done in the majority of the nodes, it may also be possible to increase the trunk line bit rates and in that way increase the overall capacity, and so significantly extend the expected life-time of this solution. However, this kind of solution is always more rigid in the longer term than a pure packet network, and this disadvantage needs to be weighted against short term benefits.

Another limitation, compared to the next example, is that the packet switching technology and solutions are limited to those alternatives available in the NG-SDH/MSPP nodes belonging to the same family as the existing SDH nodes – for example, only some of the Layer 2 packet switching solutions may be possible.

The solution example shown in Figure 10.2 is for a case where new high capacity base stations generating a lot of packet based traffic are added first to one area; so in this area the packet transport efficiency needs now improvement. Here it is done by upgrading part of the SDH nodes to MSPP nodes with packet switching capability, so that the statistical multiplexing benefits of the packet traffic can be utilized. First upgrades are where the packet traffic enters the SDH domain and where it leaves it; the SDH nodes in the middle (shown by an asterisk) can be upgraded to MSPP when there is packet traffic coming from several nodes below it (e.g. when another area is also equipment with new high capacity base stations).

The new MWR links shown with dotted lines can be packet based from day one. The situation for the two MWR links marked with two asterisks depends on the capacity and type of the existing links: they can be used, if their capacity is high enough and if they can be provided with packet (Ethernet) interfaces to work in hybrid mode – Ethernet interfaces are needed, as carrying the packet traffic over a number of (bundled) E1 interfaces would be cumbersome and expensive. If this upgrade is not possible with the used equipment or if their capacity is not high enough, new high capacity radio links need to be installed – either packet based radio links in parallel to existing ones, or new links replacing the existing ones. In the latter case the new links also need to support the existing base stations, i.e. their TDM (or ATM) traffic and satisfy their synchronization requirements.

10.4.2 Enhancing SDH/Sonet Networks with a Packet Overlay

This approach is similar to the previous one except that the packet switching capability is added using separate nodes which are then connected to the existing transport equipment. Thus starting packet based MBH network in this way is also relatively easy and flexible if the new

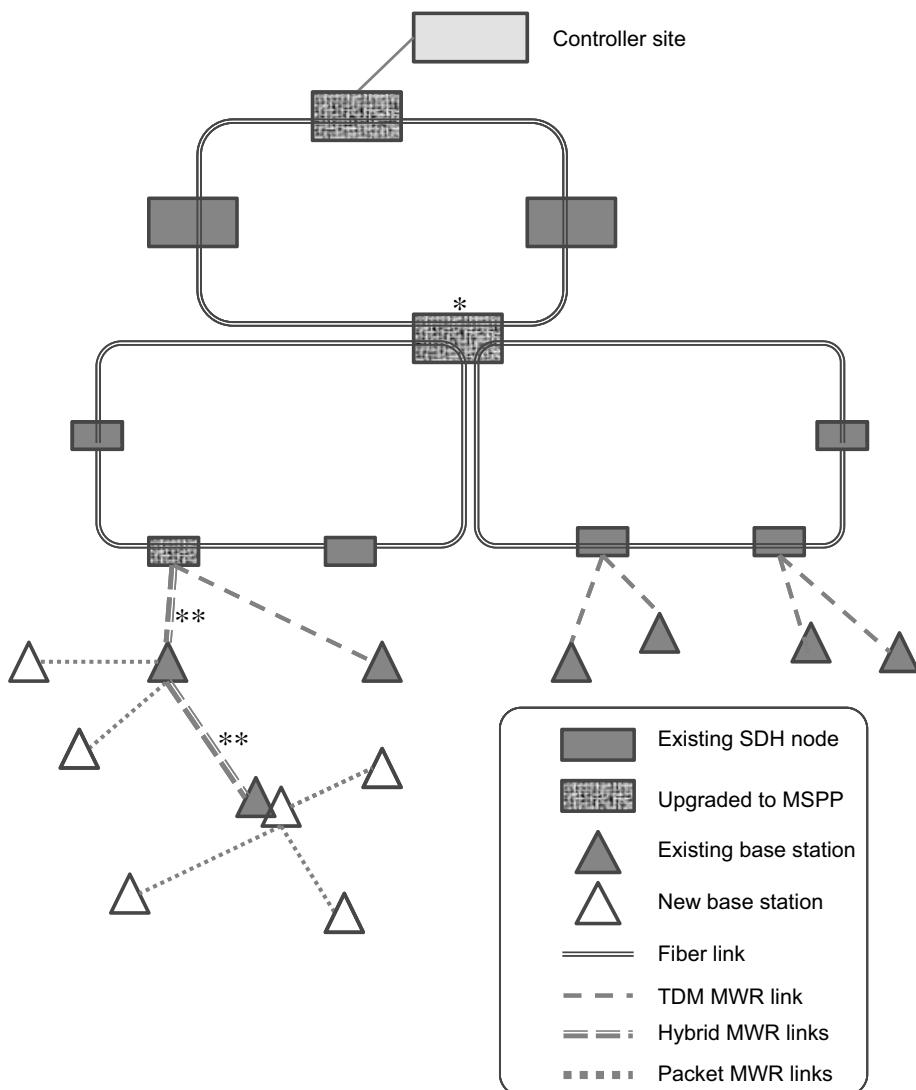


Figure 10.2 MBH network upgraded for packet traffic using MSPP nodes.

packet nodes can be accommodated in the existing sites. Existing synchronization schemes can be kept to a large extent; in addition here the packet switching technology can be selected independently of the existing nodes (e.g. Ethernet Layer 2 solution or partly IP/MPLS solution). A limitation can be (or some extra costs may be caused) because of the packet nodes may in this solution need SDH type interfaces for the interconnection to existing nodes.

The existing trunk capacity is thus also shared in this approach, which can be the main limitation of the solution; if the existing links still have a significant amount of unused capacity or if the mobile packet traffic growth in this area is moderate, this approach can provide a solution for several years.

This solution has the benefit, compared to the previous one, that the new packet nodes can be immediately of significantly higher capacity (even much higher than present trunks can handle), to be ready for the future network evolution steps when the trunk lines will be replaced with new higher capacity ones.

The solution example shown in Figure 10.3 is also for a case where new high capacity base stations generating a lot of packet based traffic are added first to one area which then needs packet transport efficiency improvement. Here it is done by adding packet switching capability

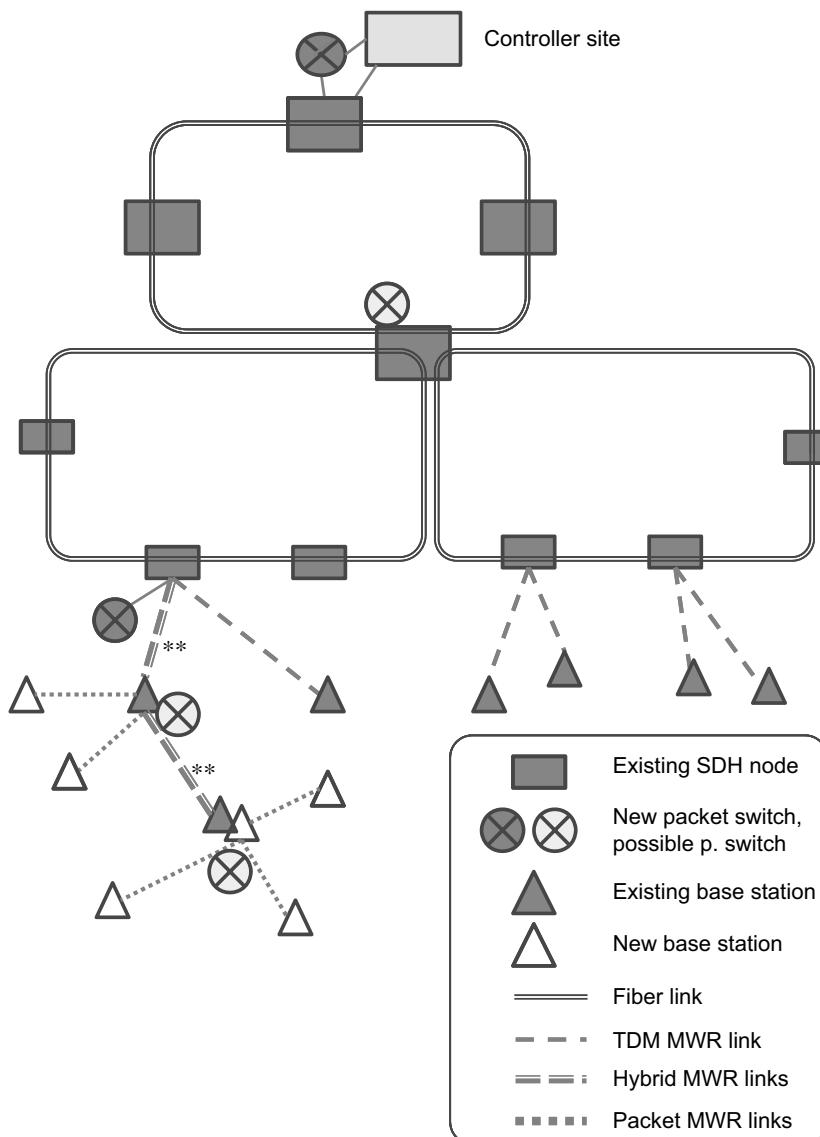


Figure 10.3 MBH network upgraded for packet traffic using packet switch nodes.

using new nodes but still utilizing the existing trunk line capacity. These new nodes are first added only to the sites where the statistical multiplexing benefits of the packet traffic can be best utilized, i.e. where the packet traffic enters the shared transport domain and where it leaves it (dark nodes in the figure), and possibly also to the sites in the middle where several packet based connections merge (light nodes in the figure). These nodes can also be added later when the packet traffic starts to grow faster.

The new MWR links shown with dotted lines can be packet based from day one. The situation for the two MWR links marked with two asterisks is similar to that in the previous solution: the existing links can be used, if their capacity is high enough and if they can be provided with packet (Ethernet) interfaces to work in hybrid mode. Here the need for Ethernet interfaces is even more important, if the packet switches are to be connected to these links. If Ethernet interfaces cannot be added or the capacities are not high enough, new high capacity radio links need to be installed – either packet based radio links in parallel to existing ones, or new links wholly replacing the existing ones. In the latter case the new links also need to support the existing base stations, i.e. their TDM (or ATM) traffic and satisfy their synchronization requirements.

10.4.3 Fully Packet Based Networks for MBH Backbone and Aggregation

A common way to start building a fully packet based backhaul network is to build a packet based transport for backbone and aggregation networks first, and to extend packet networks at a later phase to MBH access networks.

The backbone network used for MBH traffic – also in those cases when connections are not leased but based on in-house facilities – is usually shared with other traffic types (fixed network), and therefore the packet network in this tier is a multiservice network.

In general there is a lot of fiber pairs between the major sites, and then the packet network can use its own fiber pair; in other cases wavelength division multiplexing is used and the packet network can have own wavelengths. In both cases the packet network is thus parallel to the ‘legacy’ transport network and mainly independent from that. The backbone packet networks are optimized for total traffic carried over those, and therefore some care needs to be taken that all mobile specific requirements are also fulfilled. Most importantly, when at a later phase of the evolution the legacy backbone is taken out of service, synchronization of the mobile networks must then be good enough over the packet network (unless external means like GPS is used; see closer in Chapter 6).

Aggregation networks are more local, and building of packet based aggregation networks can be done area by area. Often also here the packet network is first built to be in parallel (as an overlay) to the existing legacy transport network. When the packet based MBH aggregation network requires very high capacities from the early phases on, sharing of the infrastructure with the existing transport is best done on the basis of own fibers or own wavelengths.

Aggregation networks can be mobile specific (for MBH traffic only) or shared with other traffic; in the latter case mobile specific requirements again need special attention. And when the legacy transport network is planned to be taken out of service in an aggregation area, mobile network synchronization also needs a new solution in the aggregation tier.

10.4.4 Building Fully Packet Based MBH Access Network for New Base Stations

One pragmatic way to move towards a packet based MBH network in the access tier is to build a packet transport infrastructure first for new high-capacity base stations – these base stations are anyway likely to carry a lot of packet based traffic which could heavily load the existing (smaller capacity) MBH network.

10.4.4.1 Green-Field Case

If the new high-capacity base stations are coming to an area where there is no MBH network yet, this is a pure ‘green-field’ case, and the MBH solution starts from finding the most cost-efficient physical (‘layer 0’) connections. The fastest and also most cost-efficient solution in this case may be a wireless one, i.e. transport from base station sites based on microwave radios. Installing fiber cabling to the new sites requires time and permissions, may be very expensive but provides for the highest capacity, especially in the longer term. Leasing or outsourcing of connections is an option if another operator (mobile or transport operator) has already built a transport network in the area, or is ready to quickly build one at a reasonable cost.

When the feasible ‘layer 0’ solution alternatives are known, design of the packet networks itself can be made. Newest available technologies and equipment can be considered for such a new network, obviously with the optimization criteria of Section 10.2 in mind. Usually there are many technical alternatives, and just as an example a possible solution is described below; in practice other options may be considered as well.

An example of the mobile network coverage area extension is shown in Figure 10.4. Here the existing part of the network is left as it is for the time being; new coverage area is built with new higher capacity base stations. The new base stations are assumed to have native packet interfaces (Ethernet at the lower levels), and thus the MBH solution can be fully packet based; also it is assumed that the new base stations have integrated packet switches (typically Ethernet switches) so that the intermediate base station sites do not need external switching equipment.

Here the MBH access network is built with high capacity packet based microwave radios – in other areas fiber based solutions may be considered as well. In this example the new packet switch (at the site B) is using different fiber pairs (or different wavelengths of a WDM system) than the existing SDH nodes, making the new packet network an overlay network and thus independent of the existing TDM transport – benefits include independent network design and e.g. free selection of trunk capacities, but on the other hand a packet based synchronization scheme must be in use from the day one.

10.4.4.2 Overlay Case

When the new high-capacity base stations are coming to an area where there are already base stations and sites, there is also an existing MBH network, and the new packet based network will form some kind of an overlay structure. Existing physical links may be utilized for the packet connections by sharing the capacity; however, if the new base stations require much higher capacity per site than the existing ones, it is likely that many existing transport links do

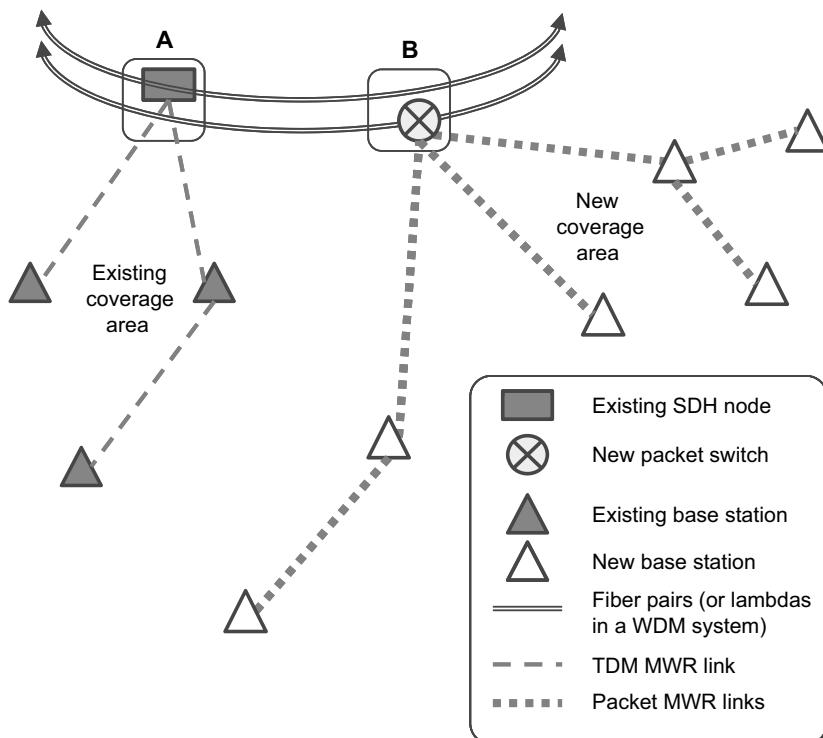


Figure 10.4 A fully packet based MBH network for a new coverage area.

not have sufficient capacity for this and new links need to be considered. Then again there are two options: either build new links in parallel to existing ones, or to replace them with new high-capacity links and then share capacity from the new links for the existing system. In the latter case the new links need to be interoperable with the existing transport (e.g. support the existing connection types and provide suitable interfaces).

In both cases the packet network itself will most likely be based on its own new nodes, and its design can be similar to the case above. However, in this case, the possibility of carrying all the traffic of these base station sites over the packet network at a later date needs to be considered, and thus options for supporting ‘legacy’ traffic can be important here.

An example of mobile network capacity expansion is shown in Figure 10.5. Here the existing base stations are left operational as they are for a significant period of time (for several years) and more capacity is provided by newer generation base stations installed mainly on the same sites as the existing ones, and some new sites added where capacity targets cannot be met otherwise (in the example the site F).

In the site A the same basic solution is used here as above, i.e. different fiber pairs (or different wavelengths in a WDM system) are used for the new packet based network, making it independent of the existing one. The MBH access solution is also here based on microwave radios, with two different approaches shown in the picture. For the sites B and C at left the existing microwave radios are replaced with new hybrid MWR systems, providing TDM capacity for the existing base stations and packet connections for the new base stations (only

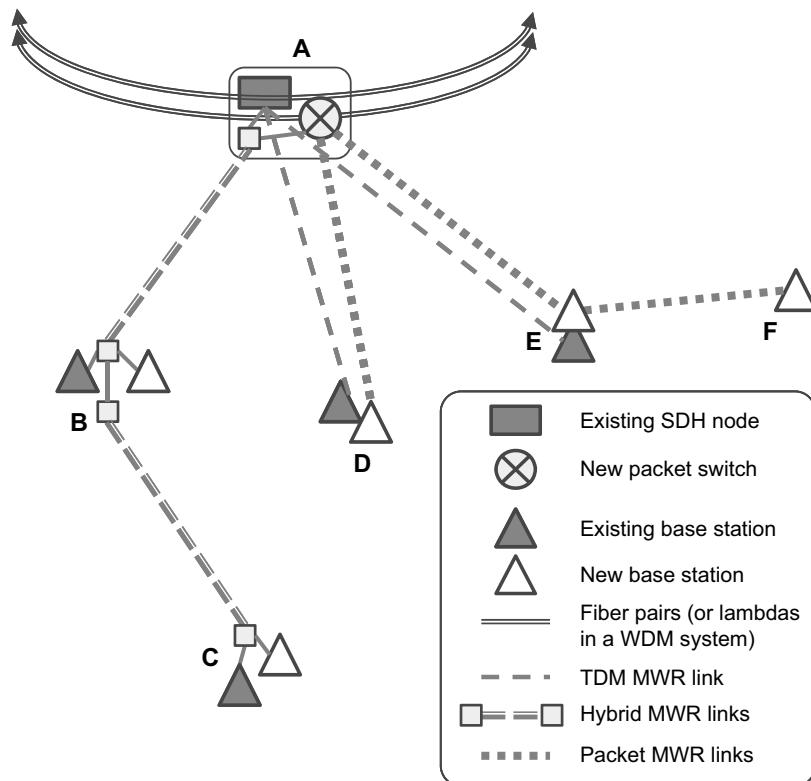


Figure 10.5 An example of the packet based overlay MBH network.

for these hybrid links the MWR terminals are shown, just to make the connections clearer). On sites D and E the existing MWR are left as they are, and in parallel to them new high-capacity packet based radios are built. And the new site F obviously only needs a packet MWR connection (if the packet based synchronization is ready in use in this network – if not, a hybrid MWR could also be considered to this site).

10.4.4.3 ‘Fill-in’ Case

A combination of the cases above is a case – likely to be very common in the near future – where new high capacity base stations are coming to an area with existing base station sites and with an existing MBH network, but also a significant number of new base stations are located between the existing ones. Thus the existing MBH network needs to be strengthened significantly (as in the overlay case above), and in addition a significant number of new links needs to be added to connect the new base station locations (number of F type sites is much higher than in Figure 10.5).

A special case of this is when the new base stations are very small physically and lower cost location solutions can be used – for example, outdoor base stations connected on building walls or even to street lamps. In such cases the cost pressure for the backhaul connection is very

strong, and new ‘light-weight’ MBH solutions may be needed, as the cost of a conventional connection link may be un-proportional to the cost of the base station itself (see also the example later in Section 10.5.1).

10.4.5 Building Fully Packet Based MBH Access Networks Area by Area

Another way to start moving to the packet based MBH network (in the access tier and possibly also in the aggregation tier) is to build a packet based transport area by area, simultaneously for both new and existing base stations. Then the legacy MBH network(s) in those areas will be quickly and wholly taken out of service, and the cost of continuous operation and maintenance of parallel networks can be avoided.

The first planning questions in this approach are how big areas are taken in each step, in which order those network areas are handled and how the moving boundary between the new packet based MBH network and the legacy network is taken care of. If the aggregation tier has already been built earlier for packet traffic, then interfacing problems are greatly reduced, as the access areas can work without direct interconnections, at least temporarily.

The technical solution for the packet based access MBH here is similar to the overlay case in Section 10.4.4 ‘Building packet MBH network for new base stations’, except that the support for all the base stations, including the oldest ones, is more important: all base stations must be fully supported by the packet based MBH from day one, as the legacy MBH network is going to be dismantled very soon after the traffic change-over. And this full support obviously needs to include synchronization transfer for all different base station with their different requirements, if synchronization is based on transport network (i.e. not based on GPS or other network external means).

10.4.6 Other Possible Approaches/Strategies

In practice transition to a (high-capacity) packet based MBH network is often some combination of the above cases, as conditions and targets within a large network area can vary considerably. Then different approaches may be applied within the same mobile network for the transition to the MBH packet network; however, still a clear change-over strategy is usually preferable (if needed, separately for each major network area).

Various combinations of the solutions are also possible when part of the packet based MBH network is leased or out-sourced in certain areas or network tiers – those cases are discussed in the next section.

10.5 Outsourcing the MBH Network or Parts of it

Use of leased connections or outsourced services as a part of the MBH network on a longer term basis is both a big economic question and a major strategic decision, and it is also related to a number of technical considerations, especially in the case of packet networks. But such services can also be used on a more temporary basis, e.g. for filling a gap between two implementation phases of the in-house MBH network, and in such cases this has little or no influence on strategic or organizational issues.

The size of leasing or out-sourcing, i.e. how a big part of the MBH network is in question, affects greatly all the following considerations and comparisons – for a small local use of transport services some rough calculations may be enough (still the technical issues must be clarified, see Section 10.5.3) while for a regional or network wide MBH out-sourcing clearly very extensive calculations and comparisons of different aspects are necessary.

In all cases availability of transport services in the mobile network area is the first question to be answered; in addition to the present situation, some view of the future development of the service offerings is required. It is also necessary to evaluate whether reliability and quality of these services will be of a proper level for the MBH usage (detailed technical clarifications may be left to a later phase). Then next tariffs and their forecasted development is to be taken into account and, if the situation looks good or at least reasonable at first view, one can start more detailed economic evaluations.

10.5.1 Economic Considerations

Use of leasing or out-sourcing as a part of the MBH solution is obviously an important economic question. Use of services instead of in-house network resources reduces the need for network investments and thus decreases mobile operator's capital expenditures (capex), but at the same time it significantly increases MBH network operational expenditures (opex) – permanently or for the period when these services are used.

As the expenditures are of a different type and especially as the timing of expenditures is very different, reliable economic comparison of the in-house network solution and service-based is not easy. Methods for the economic comparisons for such cases are well known, one can use for example net present value (NPV) or total costs of ownership (TCO) calculations for comparing the in-house and outsourced solutions. The difficulties or inaccuracies lie in the data – all these methods require cost and tariff as well as interest forecast for several years into the future; depending on the area and environment, such forecasts contain smaller or bigger uncertainties. However, in a number of cases the results are so clear that small data inaccuracies do not influence the conclusions. When the first calculations do not give clear results, different scenarios for the costs and prices can be played to get a better view on the economic aspects of the decision.

When estimating future development of tariffs for services, one should also consider expected development of competition in such services in that area. If there are several service providers (e.g. several operators having transport networks in the area) tariffs are likely to be better foreseeable than in the case of only a single provider. In the latter case possibility for long term contracts may partly alleviate the forecasting problem.

One should also note that the total cost of the solution (e.g. lowest NPV) is not the only economic measure of significance; in some situations for example operator cash flow considerations can be more important. Therefore in certain cases outsourcing may be justified even if it does not result in lower NPV than the in-house network.

10.5.2 Strategic and Organizational Considerations

When a bigger part of the MBH network is planned for out-sourcing, strategic considerations become as important as the economic comparisons. Strategic things include for example

dependency of other parties (especially on companies close to competitors), flexibility of the planned solution in cases of changing one's in-house targets, influence on one's own organization and the resources and skill sets needed, and expected (longer term) development of reliability and quality of the MBH network.

Dependency on other (telecommunication) companies is as such not any reason to avoid outsourcing solutions, but it is good to understand well how these dependencies are likely to affect future competition situations, and how much of negotiation power each party is likely to have. If the (transport) service providers are not in the mobile service business and not likely to enter into it, evaluation of service provider dependencies means in the first place evaluation of how well the MBH requirements are likely to be fulfilled. If the transport service provider (or its mother/daughter company) is also in the mobile business, significance and possible longer term influence of this needs to be evaluated.

A related issue is flexibility of the outsourced solution in the case of when one's own needs are changing. This is dependent on the (transport) service providers' flexibility and response times, how willing and how fast those organizations are to change earlier agreed or planned network services, and finally also on how the outsourcing contracts are written – how the flexibility aspects are taken into account in the agreement terms and conditions.

Another significant business issue is the influence of outsourcing on the mobile operator's own organization; outsourcing of a significant part of MBH network means reduction in necessary in-house resources, and also some changes in the required skills. In outsourcing cases the general network understanding and very exact requirement specification skills are important, as well as skills needed to make reasonable contracts with service providers. Contracts shall have wide enough coverage of both economic and technical issues, including (packet network) service quality and reliability specifications. Skilled resources are also needed for the follow-up and technical monitoring so that the mobile operator can be sure that received service according to the agreements. Resource savings are then possible in detailed network planning and especially in the operation and maintenance of the MBH network. Changes in resources and skills have, in addition to the economic side, strategic aspects, as once activated changes cannot be instantly reverted.

10.5.3 Technical Issues

Technical issues related to MBH outsourcing are discussed here last but it does not mean that they are less important – in the case of packet based transport, the situation is definitely the opposite: technical requirements, interfaces and interworking issues as well as monitoring and network management require quite extensive considerations and specifications.

First technical issues relate to the network interfaces between the mobile operator network and the service provider network. In a packet based leasing and network services the physical interface is often an Ethernet interface; in the case of the MBH access networks it is typically 100M or 1G interface, on the backbone side also higher rates (10GE) are applicable. Network interfaces need to be well defined also for higher layers (logical network interfaces), e.g. termination points of connections, traffic classes and their priorities and addressing used in the host and serving networks.

The next thing to specify is the available capacity for the MBH connections over the network interfaces (which can be significantly smaller than the interface speeds). Capacity

may be specified in terms of ‘guaranteed bandwidth’ (or throughput) which is always available for the MBH traffic, and in terms such as ‘excess capacity’ – this traffic is carried over the network when there is no congestion but it is not guaranteed. And when traffic classes are used over the network interfaces, it may be necessary to specify capacities for various traffic classes separately.

It is more complex to define network performance, often described for packet based services by a Service Level Agreement (SLA). It covers in addition to the throughput also connection quality measures, for example expected latency or delay (possibly with some measure about its distribution) and packet loss probability; and again, when traffic classification is used, performance values will be different for various traffic classes.

A related area is the service or network reliability, or dependability, described in terms of probability of outage, downtime (per month), (maximum) duration of outages and recovery time. Reliability may be specified for end-to-end connections (base station site to controller/server site), or separately for different network tiers (as upper MBH tiers influence a higher number of base station sites than the lower tiers).

A new very important area is related to base station synchronization, when this is carried out over the packet based connections (and not done locally based e.g. on GPS). Providing good enough reference for the base station synchronization over the packet network is not trivial but insists on a number of requirements for the packet connections, as discussed in Chapter 6. In the case of leased connections, these requirements need to be included in quality specifications.

One very important area still to be agreed upon is monitoring and management of the connections. Monitoring and supervision of connections can be made separately by the mobile operator and the service provider operator, on both sides of the network interfaces, but it can be done more efficiently when information is shared. If the connections within the service provider network are to be managed by the mobile operator (e.g. capacity reservations), a network management level interconnection needs to be agreed and specified. In cases where different parts of the operator’s network are outsourced to different external organizations, complexity obviously increases significantly, as there are more parties needing access to the (partially) same data; early agreeing of proper sharing of data as well as of responsibility for overall network supervision and end-to-end quality of connections is very important in such cases.

However, in spite of many technical areas requiring considerations and specifications when using external packet based network services, leasing and outsourcing shall not be seen only as creating a number of complex interconnection and specification issues. A packet based MBH network run by a specialized packet network operator, with a staff specialized in packet transport issues, can provide a cost-efficient MBH network solution; it is also possible to get very good performance and high quality that may not be as easily achieved if in-house resources or organization skills or investments in the packet network technology are limited.

10.6 Selecting MBH Access Solution for a Particular Case

The question of selecting a packet based MBH solution for a particular practical case is discussed here based on a few examples. The process begins from the ‘hard’ starting points of a particular case, then adds the ‘soft’ ones (including the MBH targets and goals), continues with

surveying possible technical options and studying these alternatives against the optimization criteria (with proper weighting for the actual case), and finally finds the most suitable solution(s) for a more detailed analysis, detailed planning and/or for formulating an RFQ.

10.6.1 MBH Solution for LTE in a Dense Urban Area (in a Developed Environment)

Building of a high-capacity mobile network layer based on LTE technology will be a common way to increase the capacity of an existing 3G mobile network in the central areas of big developed city areas ('dense urban environment').

The new LTE base stations are in the first place located on the existing sites, but often this will not provide high enough capacity over the whole area, and new fill-in sites are needed; how many, depends on the density of the existing mobile network and on the decided capacity targets. If we assume for example that the number of fill-in sites will be roughly equal to the existing sites, then one 'hard' starting point for the new packet based MBH network is that about half of the last-mile links exists (but the link capacities are based on the 3G network needs) and the other half of the last-mile links will be new-builds.

MBH capacity requirements or dimensioning of MBH for LTE network is now, as discussed earlier in this book, based partly on the LTE traffic forecasts and partly on the mobile operator's decision on the guaranteed single user peak rate (to be available throughout this area). The latter defines for all last-mile links the minimum capacity (it is the peak rate + small reservation for other traffic, including signaling + the transport overheads) while traffic forecasts determine capacities needed upper in the MBH networks, for links serving several cells. Capacity targets for each link in the new packet MBH is now an important 'soft' starting point.

In this situation MBH capacity is one of the biggest issues, and the other one is the cost of the significant number of new (high-capacity) links needed; construction time may be a third big issue, especially if the new sites are to be joined to the existing network with cable (fiber) connections.

In many cases capacities of the existing MBH links are simply too small for the LTE traffic, possibly too small even for the guaranteed single user peak rate. If we assume that this is the case (at least for the majority of the links), the conclusion is that new equipment with higher capacity will also be needed for existing links in this area. So, the selection here is between a high-capacity packet overlay to the present MBH network or replacement of the existing MBH links with a high-capacity packet network, in the latter case also supporting all the existing base station traffic.

In both alternatives new links are needed for a significant number of the new sites. These links shall carry high capacity traffic to the new packet based LTE base stations, and thus naturally will be packet based. In addition, links need to be of low cost, cost pressure being the stronger the smaller (and cheaper) the base stations themselves are. Low cost high capacity packet based microwave radios are one solution; another one (at least in some cases) is optical fibers to the LTE base stations, if and where the cable installation costs are not prohibitive.

Construction time of the new links may also be an issue limiting choices. If the new network shall be taken quickly into service, all installations need to be done in a strictly limited time. This may exclude elaborate cable laying projects, especially in dense city centers, and especially in old (protected) city centers.

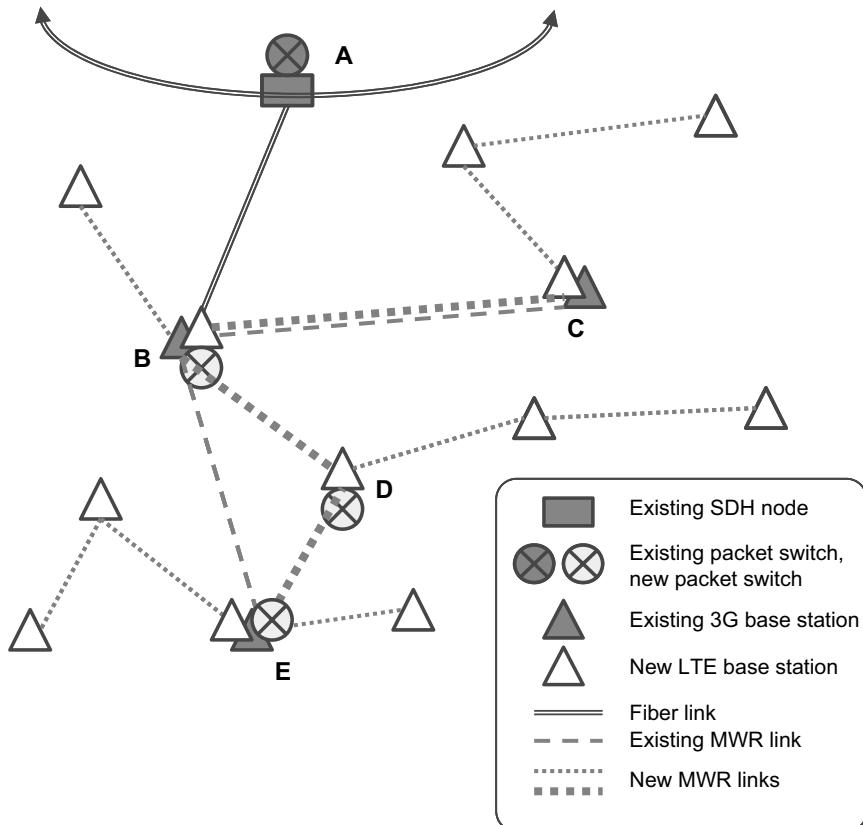


Figure 10.6 Example of a MBH network for new fill-in LTE base stations.

Next we consider here an example, remembering that it can only represent one particular case, and in other kind of environments different solutions may result in lower costs.

An example case is shown in Figure 10.6; in this area there are three existing 3G base stations and thus three existing base station sites, one of which is connected with fiber to the aggregation network and the two other sites are connected to that site with microwave radios of 16...34 Mbit/s capacity. For the mobile network LTE upgrade, new LTE base stations are first installed on the existing sites, but here we assume that in addition to those, nine small LTE base stations are needed for having full coverage and enough capacity (a dense city center area assumed). Here we assume that those LTE base stations have an integrated Layer 2 switch for the transport, so that all sites do not need external packet switches for the MBH network. Further, we assume that mobile operator wants to guarantee a 50 Mbit/s data peak rate to all locations in the network; that determines the minimum capacity of all the MBH links (need to be $> 50 \text{ Mbit/s} + \text{overhead} + \text{signaling} + \text{O\&M}$, in practice $> 60 \text{ Mbit/s}$). The example MBH solution shown in Figure 10.6 consists of new low cost short-range microwave links: smaller capacity links (thin lines, e.g. 100 Mbit/s) for connecting tail sites and for next connection, and higher capacity links for carrying LTE traffic aggregated from three or more base stations (thick lines, e.g. 150...300 Mbit/s, depending on the availability of various

capacity options). Line-of-sight limitations mean that not always the shortest paths can be used; on the contrary, in a dense urban environment MWR links need to follow streets and other ‘low’ areas – thus the topology can be less straightforward, as shown in the example. Another thing is that, as here the minimum capacity for the LTE MBH links is around 60 Mbit/s, the existing MWR links cannot be used, and shall be upgraded or replaced; in this case, however, a lower cost solution is to use new low-cost packet MWR in parallel to the existing one (between sites B and C) and to use a new base station site (site D) to create new shorter links with packet MWR – later in the future fiber may be extended to this site. Finally, in this example external packet switches (with their own power supplies) are added to the sites where traffic of four or more LTE base stations is handled.

Leasing or out-sourcing transport in the dense urban area is in principle easier than in many other environments, as often there are several transport networks owned by different operators in the area, and therefore it is possible to get several offerings for the needed packet based connections. The challenge lies in the smaller base station sites – the smaller the sites are, the more challenging it can be to get a physical connection there from another operator. Thus it is easier to lease capacity for the main routes (and the upper tiers, e.g. aggregation network) while the tail links to the smallest sites remain easily within the mobile operator’s own responsibility domain (suitable leased connections are not at all available, or are tailor-made and too expensive).

10.6.2 MBH Solution for Suburban Area for 3G + LTE (in a Developed Environment)

Another common case in mobile network developments is expansion of high capacity 3G networks from city central parts into suburban areas, to cope with increasing data needs also in residential areas. This expansion may bring LTE base stations at the same time in high traffic density spots in these areas, or then just a reservation is made to enable adding them later to a part of base station sites.

The mobile operator’s own GSM or GSM/EDGE network (or other mobile network) may exist already in the suburban area, in which case at least a significant number of sites is already there and there are MBH links for connecting them. However, these MBH links are likely to be of very small capacity compared to the requirements of new data oriented 3G (HSDPA and HSPA) and 3G + LTE networks. Therefore most of links need either to be immediately replaced or higher capacity links need to be constructed in parallel to them. In both cases the new MBH network shall be future-proof, accommodate further data increases smoothly, and therefore it is built from the beginning using packet network technologies. And as earlier, in the replacement scenario the requirements of the existing mobile network once again need special attention in the packet network implementation (especially the synchronization) so that service quality of the existing mobile networks is not compromised. Depending on local circumstances, the MBH solution may be similar to one presented in Figure 10.4 in Section 10.4.4, but possibly with smaller capacities.

If there are no older mobile networks in the suburban areas in question, then we are back in the green field case, and the solution is a wholly new packet based MBH network. In this case the MBH network costs will be higher (all routes are new) but from the planning perspective the network can be easier to do and optimize.

The packet based MBH network solution in suburban areas is often based on extensive use of Ethernet microwave radios – especially in the case of fast growing networks. This enables fast building of the MBH connections and also re-structuring of the MBH topology, if that is needed when new sites are added. Critical things are availability of suitable frequency bands and wide enough radio channels to accommodate the high capacity Ethernet MW radios; in some countries also the fees to be paid for these channels can be a significant factor.

A suburban area may also have an existing relatively dense fiber cable network installed, and in such cases it may be feasible to use optical transport for part of the MBH connections, especially for higher capacity links serving several base stations. The challenge is often the last tens or hundred meters of the connections – to get the base station site (e.g. on a high hill or on roof of a house) connected to the fiber network.

Leasing or out-sourcing transport is in these cases often a valid option to be considered, as in this kind of areas it is often possible to get offerings from other operators having transport or access networks there. The question is whether those networks are fully packet based and can they provide high enough capacity, throughput and latency, with reasonable guarantees. This needs to be considered also for a few years to come, comparing network capabilities (and their expected developments) with the mobile traffic forecasts for that area.

10.6.3 MBH Solution in a Rural Area for a New 3G Network

This case is not as common as the two previous ones, but will occur where high data capacity is required to be available on highways or railway lines or wider in a rural area. The latter can be the case when living in rural areas is supported by provision for remote working and fast access to internet content (wired connections may be not at all available or very expensive). Other possible cases include a new network built to cover a rural area where high data capacity is wanted to be available from the beginning – to enable various data intensive IT applications – for example in a developing area with little or no fixed telecommunication infrastructure.

In a rural area network the main challenge is often distances more than capacities – base stations are located relatively far from each other and depending on local topography, free line-of-sight between all nearby sites may not be available. Cable laying between the sites may be straightforward in certain (flat) terrains, even if the costs increase with the distance, but cable laying can be very difficult and expensive in a hilly or mountainous area, or in an area with a lot of swamps, lakes or rivers. Thus the MBH network topology needs to take into account the possibilities of building physical links between the sites.

In rural network MBH capacities needed per site are smaller, especially during the first phase of the mobile network. Thus the last-mile links can be of smaller capacity (than in dense urban areas). However, the topology of the MBH access network may be much deeper in a rural area, and then the links closer to the aggregation network serve quite a few base stations, and therefore need higher capacity (and may need higher reliability as well). In selecting the equipment for the MBH links also the forecasted traffic growth or traffic scenarios need to be taken into account, so that the links are adequate also for the next network phase(s) and do not need replacement too soon.

A microwave based MBH solution is often the first consideration in rural areas along with a packet based network. Microwave links can be fast set-up, and the MBH topology can be changed later relatively easily if a need arises (e.g. due to new fill-in base station sites).

Also the cost of cable laying over longer distances often makes a wireless transport solution more attractive. In a green-field network Ethernet microwave radios can be used, together with small packet switches (may be integrated into the base stations), and thus a wholly packet based MBH network created from day one. Protection of the most important MWR links against equipment failures and adverse propagation conditions needs to be considered.

In other rural environments cable laying is feasible and costs are not too high, and thus optical transport solution can be applied, at least for a part of the MBH access network. In this case the new MBH network is based on fiber links and small packet switches attached to them; those switches may also be integrated into the base stations. Such a network can be built from day one for very high capacity, high enough for several mobile network development phases.

In practice the rural MBH network is often a combination of microwave and fiber link connections, with fiber links in the first place laid for the highest capacity connections. Also the (road) infrastructure may play a role – sometimes it is relatively easy to install fiber cables along the main roads, or along railway tracks, and thus those road-side or rail-side base station sites may be fiber connected while the rest of the rural network is microwave based.

Leasing or out-sourcing transport is quite rarely a valid option in rural areas, as often there are no existing transport networks suitable for the MBH needs. Still there may be cases where due to business reasons the new (to-be-built) MBH network is out-sourced from the beginning, i.e. it is built by another company and the mobile operator leases packet based connections over it, or out-sources the whole MBH function.

10.7 From the Selected MBH Solution to Detailed Network Plans

When a MBH solution (type) has been selected, basic deployment strategy defined and equipment families defined, a detailed planning of the MBH network can start.

The selected solution needs first to be broken into a number of implementation steps each of which is built ‘off-line’ in parallel to the working network and then taken into use. Depending on the selected basic approach (see Section 10.3), there may be a large number of smaller steps or just a few major ones; the first is likely in replacement scenarios while in the overlay scenarios quite big network areas can be built before taking them in to service.

In both cases each step needs careful detailed planning during which the exact equipment types and units (cards) and exact physical connectivity (including cabling) are defined and, above all, the exact logical network structure and related equipment configurations (including naming and addresses) are defined. In particular, when a scenario with high number of implementation steps is planned in detail level, careful checking of the proper operation of the network after each step is important; for example, there shall be no undefined end-points nor missing or double domains or addresses.

A detailed plan for packet based MBH networks needs to include at least the following elements:

- network (physical) connections plan within the transport network and interfaces towards the mobile network equipment (base stations and various core elements);
- transport network capacity plan (initial dimensioning and upgrade options);

- connectivity plan or logical connections between the mobile network elements for each logical network (e.g. user plane, control plane and network management plane) and for the transport network, including use of tunnels and/or VLANs;
- naming and addressing (and possible capacity reservations) for all logical networks (e.g. separately for user plane, control plane and network management planes);
- related to the above, termination points and addressing plans for all network layers;
- plan for network performance implementation, use of traffic classes and other QoS methods;
- a plan for monitoring of quality and a general MBH network management plan;
- protection of connections, network resilience plan;
- network synchronization plan;
and based on above plans, for example:
- equipment and unit lists for each site;
- power supply for equipment and possible back-up arrangements, including power cabling;
- physical interfaces used and connectivity, cabling between the (telecom) interfaces;
- configuration of interfaces (in networking layers L1/L2/L3) and of network nodes;
- addresses to be used in each individual equipment for all termination points in all used network layers and in various VLANs.

This kind of detailed level planning is a very significant part of the total project work and enough time and skilled resources need to be reserved for this phase. Detailed planning can also be outsourced, especially when the network expansions are extensive or in-house resources are tied up for other tasks, or when a new technology and equipment types are used for the first time for the MBH network development.

10.8 Summary

In this chapter we have discussed the process of selecting a proper MBH solution, when mobile traffic growth and especially data traffic growth makes transition to the packet based MBH network necessary, and the various things to be taken into account in formulating and optimizing the MBH network solution.

Typically, the optimization of the MBH is a complex task and each network case has its own characteristics and requirements. The final target is in all cases a fully packet based MBH solution but the way there can be very different in varied physical and network environments.

Also the wholly packet based MBH networks can be quite different in different environments, e.g. in very dense urban area, lower density suburban areas and finally in low-density rural areas. Thus a lot of evaluation and planning work is needed to find a proper and most suitable MBH solution for each case, based on the technical requirements, available technologies and network options described in the earlier chapters of this book.

11

Summary

Esa Metsälä and Juha Salmelin

Mobile backhaul technology is fast moving from deterministic TDM networks to packet based technologies. This book described this change step by step and function by function. There are many obstacles on the way as little remains the way it used to be in the ‘good old’ TDM. With packet technology the specifications are often not precise and are more dependent on the implementation of vendors.

Explosion of traffic in mobile networks, and mobile broadband specifically requires a high capacity backhaul. At the same time, strict control of backhaul costs are mandated, as the mobile broadband business case is typically based on a fixed monthly charge (‘flat rate’) instead of a per-megabit – charging. For the backhaul this implies that high data amounts need to be carried cost efficiently. Voice and other real-time services then need a premium service over the bulk data transfer, so the packet network needs both low cost bulk data transfer (which drives the capacity) as well as premium bits for voice (which drives the strict QoS).

The quality of service with packet based networks is also not as clearly defined as with TDM, but still typically quite good. If one has a background in TDM networks, it might be a challenge to start thinking packets. Just reading all the standards does not give the big picture of how packet networks work. There are many different specifications of different features, most of which are not used or not implemented into the products, or not into all products in the market. On the other hand, there are many features needed and they will be different in each network.

In particular, when the backhaul connection is leased from a service provider, the service level agreements might be very difficult to understand and agree on for the backhaul deployment. From the service provider’s point of view, the mobile backhaul sales will be only less than 20% of the fixed broadband sales. When the driver is fixed broadband no big effort to understand mobile specific needs exists. This book helps this change by describing the packet network specifications which are relevant to mobile systems in in-house backhaul as well as in leased ones.

In the future backhaul will be totally packet based. Other technologies are too expensive. But when the traffic continues to expand even the packed backhaul cost is too much. New backhaul optimized technologies are needed. The slogan ‘Ten times more capacity with ten

times lower cost' will be used more and more. On the other hand, new 3GPP radio technologies and features are also asking for tight requirements for synchronization, resilience, QoS and security.

It is useful to consider the backhaul as a service that is provided to the radio network layers. The radio network layer then supports a service to the end users. This view reveals interactions between the radio network and transport layers and also allows technology used in the implementation of the mobile backhaul to be discussed separately from the services it provides.

Often, the different radio network technologies are supported by the same backhaul network, even though not all of these radio technologies are based on the use of IP in the backhaul. Initial 2G is TDM, and initial 3G is ATM. When all traffic is converged to a single backhaul, emulation services are needed for the non-IP base stations. MPLS was presented as an example of a technology that supports both native IP and these emulation services.

Frequency synchronization technologies over packet networks are becoming mature, allowing mass deployment today. IEEE 1588 and Synchronous Ethernet are the mainstream solutions. This book describes both, but the main emphasis is in IEEE 1588 and the challenges of packet based synchronization. Synchronous Ethernet resembles closely the well-established SDH technology so the references to Synchronous Ethernet standards documents satisfy most needs. Accurate time synchronization, on the other hand, is needed in LTE TDD systems, which are expected to become quite popular. Even more accurate timing may be needed in some of the new LTE-A features. Time synchronization is based currently in satellite systems. This technique is described briefly. IEEE 1588 will be suitable for time transport. However, the standardization work of time synchronisation for telecom networks is still in an early phase. The work will cover, for example, how the network nodes throughout the transmission path participate in timing, which was not necessary for frequency synchronization.

For resilience, native Ethernet, carrier Ethernet, MPLS and IP were discussed. Packet networks fail differently from the TDM networks and new types of anomalies exist. Also, recovery of the node and link failures in the packet network is different from how traffic is recovered within TDM networks. Packet networks tend to be less deterministic than TDM networks due to the re-routing capabilities, although Sonet/SDH-like protection behaviour is also possible to achieve.

With the base station access tier typically only a single path exists to the aggregation network. In these cases resilience to link failures does not exist. In the aggregation tier the situation is different. Due to the amount of traffic carried, and to the amount of sites dependent on the service of the aggregation network, resilience to both link and node failures is important. Arranging resilience in the aggregation network depends on the technology used.

Quality of service is an end-to-end topic, and alignment of the radio network QoS with the mobile backhaul QoS is essential. For this purpose, mapping of the radio network layer bearers into transport QoS classes was discussed – including IP layer Differentiated Services, and a further mapping into Ethernet and MPLS layers. For the backhaul, all traffic types existing were addressed; not only user plane traffic, but also control, management and synchronization planes.

Security is a new issue to be addressed with the IP based protocol options, and with the packet backhaul. With TDM networks, security was typically not perceived as a concern. This has changed with IP networks. Even though the IP network used for the mobile backhaul is a dedicated, closed network separate from the public Internet, it is based on the same protocols,

and thus the attacks that are a reality in enterprise and service provider networks, are also risks for the mobile backhaul. At the IP layer, IPsec protocols provide protection, and are also in many cases mandated by the 3GPP. Especially with LTE, an unprotected network is a clear threat, as each base station supports IP layer connectivity to the core network.

There are no two similar mobile networks with similar backhauls. All have differences when compared to each other in some of the features or cost points. The backhaul solutions in a couple of practical cases were described in this book to give examples of how to help create individual and different solutions. In the future, when even more capacity and tighter requirements are needed, new ways of building and controlling the backhaul becomes a must. Very high capacities call optics, but there are also new wireless technologies coming which might ease the most costly first hops implementation in particular.

The mobile backhaul challenges will continue year after year. Moore's law will ensure that more and more bits will be consumed in mobile systems. More and more throughput is needed and less and less latencies are allowed.

New trends to concentrate base station processing units in one location and integrate the radio heads with antennas will change part of the backhaul to 'fronthauls'. Fronthaul connects base station processing units (Baseband Hotels) and RF-heads. Today's fronthauls need several gigabits/s with very low latencies and so the only feasible technology is point to point fibres. If those are not available the digging cost might inhibit wider use.

In cases where there are many users and simultaneously a lot of traffic per each user, macrosites will not provide enough capacity. Cell sizes must be decreased, which means a lot more small base stations. For backhaul it means a significantly higher amount of connections. Future small cell backhaul requirements drive backhaul connections towards high capacities, short delays, short hops and very low costs with as many self organizing features as possible. That kind of backhaul technology does not yet exist.

The last free IPv4 address pool was allocated by IANA in February 2011. This speeds up IPv6 deployment in general, although in the mobile backhaul private IPv4 addresses are used and thus the IPv4 address pool exhaustion mentioned is not similarly a driver for IPv6. In addition, as the end user IP is tunneled with GTP-U and other protocols, the mobile backhaul can be IPv4 while the end user IP protocol would be IPv6. Anyway, both IPv4 and IPv6 are already included in 3GPP standards meaning that an 'IPv6 readiness' exists from a standardization viewpoint. As well in practice, many router and switch platforms today support both IPv4 and IPv6 (a dual stack).

Backhaul is becoming the most costly part of the mobile network unless new innovations emerge. Operators start to look for all possible ways of reducing costs, and in particular, the total cost of ownership. Sharing the backhaul and some kind of backhaul virtualization is also waiting on the horizon.

Index

- access network **8**, 21–2, 357–60, 364–8
access tier **8**–10, 22, 25, 27, 70–1
adaptive clock recovery (ACR) 173, 187
aggregation network **8**, 356, 365
aggregation tier **8**–11, 22, 25, 2, 82, 88, 93–4, 100, 114, 123, **155**, 208–9, 224, 230, 237–242
announce message 176–9
anti-replay (protection, services, window) 311–13, 318, 322–3, 342
ARP, address resolution protocol 86, 98, 102, 108, 210, 217, 236, 283–4, 315, 320, 342
ARP, allocation and retention priority 283–5, 288, 296–7
ATM 40, 47–51, 56, 65, 70–71, 78, 93–4, 109–10, 117–18, 123, 137, 172, 305, 310, 344
authentication 81, 221, 223, 304–10, 313, 316–22, 324–6, 329–30, 332–3, 341, 344
backbone network **8**, 10, 13, 17, 20–1, 26, 155, 356
backbone tier **8**–10, 22, 27, 71, 100, 209, 219, 237, 239
backhaul dimensioning 14, 18–19
bandwidth profile (BWP) **160**
bandwidth profile enforcement 160
BCH, broadcast channel 43
Beidou 201
boundary clock (BC) 174, 202
bridge, bridging, ethernet bridging 73, 83–7, 91–2, 95, 118–19, 210–14, 216, 271, 313, 315
broadcast address 85, 101
broadcast domain 86, 88, 97–8, 119–21, 210, 212, 314
broadcast storm 86, 88
carrier ethernet 75, 84, 92, 94, 157, 210, 214, 314
CBC, cell broadcast centre 49
CDMA 168–70, 201, 203
CES 173, 180, 203, *see also pseudowire*
class of service (CoS) 160, 273–4
Color Aware Mode, color aware policer **160**, 273–4
Color Blind Mode, color blind policer **160**, 273–4
Color Mode (CM) 160
Committed Burst Size (CBS) 160
Committed Information Rate (CIR) 160, 273–4
congestion control 46–8, 255, 257, 260–2, 266, 285, 288–93
control plane policing 315
controller/gateway site 7–8, 10
convergence 100, 209, 213, 218, 222, 226–7, 297
core site 7–8, 16, 27
cost distribution 9
Coupling Flag (CF) 160
delay 269
delay budget 72
delay jump 196, 198, 200

- delay request, Delay_Req message 176–8, 202
 delay variation, *see* packet delay variation
 detailed plans/planning 347, 364, 368–9
 Diffie-Hellman 320–4, 332, 342
 digital subscriber line (DSL) **148**, *see also*
 SHDSL and VDSL
 domain (in PTP) 178
 DOCSIS 154
 Doppler shift 167–8
- E911 169–70
 E-LAN 84, 92, 118, 157–8, 162, 210, 214–15, 242, 314–15
 E-Line 92, 120, 157–8, 162, 214–15, 238, 242, 314–15
 end user peak rate 39, 58, 76, 78, 232
 end-to-end delivery 104
 end-to-end emulated service 117, 138
 end-to-end QoS 251, 255, 264, 337
 end-to-end service 61
 EPS bearer 56, 60–2, 104, 293–7
 E-RAB (evolved radio access bearer) 61, 294, 298
 Ethernet in the First Mile (EFM) 154
 Ethernet over Copper (EoC) 162
 Ethernet Private LAN (EP-LAN) 158–9
 Ethernet Private Line (EPL) **158**
 Ethernet Private Tree (EP-Tree) 158, **159**
 Ethernet service 157–62
 Ethernet service definition **157**
 Ethernet Virtual Connection (EVC) **157–61**, 273
 Ethernet Virtual Private LAN (EVP-LAN) 158, **159**
 Ethernet Virtual Private Line (EVPL) **158**
 Ethernet Virtual Private Tree (EVP-Tree) 158, **159**
 E-Tree 92, 157–8, 162, 215, 242, 314–15
 Excess Burst Size (EBS) 160
 Excess Information Rate (EIR) 160–1, 273–4
 Expedited Forwarding (EF) 200, 265, 268
- FDD 129, 142, 168, 170, 200, 203
 FEC (forwarding equivalence class) 110–11, 113, 229
 fill-in base stations 359, 364–5, 367
 filtered packet time error sequence 197
 firewall 99, 104, 217, 303, 329–30
 floor delay packet population 197
 follow up, Follow_up message 176–7
- forward delay 212–13
 fragmentation 54, 59, 95, 107–8, 112, 303, 336–7, 344
 Frame Delay (FD) 161–2, *see also* packet delay
 Frame Delay Variation (FDV) 161–2, *see also*
 packet delay variation
 Frame Loss Rate (FLR) 161–2
 frequency accuracy 167–8
 frequency error 171, 182–4, 186, 189–91, 194–5, 199
- general transport 13, 21
 global navigation satellite system (GNSS) 201
 GLONASS 201
 GMPLS (Generalized MPLS) 123, 230
 grandmaster (GM) 175, 177–9
 green-field case/network 357
 GTP-U, GPRS tunneling protocol 22, 41–2, 52–6, 57, 59–60, 62, 95–7, 103–4, 217, 265, 368, 297, 310, 312
- hello (LDP) 111, 219, 226
 hello (OSPF) 219, 220–4, 235–6
 hello interval (OSPF) 222
 HVPLS (horizontal VPLS) 121
- IEEE 1588 *see* PTP
 impairment emulator 198–9
 ITU telecom profile (for IEEE 1588) 177–81
 ITU test case 180, 198–200
- leased lines 13, 19, 26–7, 156, 352, 360–1, 363, 366
 load sharing 39, 87, 99, 122, 210–11, 214, 217–18, 241, 244–6, 306, 331, 336
 loopback address 98, 229, 233–5, 335
- MAFE 193–7, 199–200
 mask 184–5, 195, 197, 199–200
 master 174–81, 187–9, 192, 200
 MATIE 193–5, 197
 MEF (Metro Ethernet Forum) 157, 273–4
 microwave, link calculation **139**
 microwave, spectrum **143**
 milli-metre wave radio (mmWR) 138
 minimum delay 189–90
 mobile traffic forecasts 12, 14–15, 348, 362, 364
 Modified Allan Deviation (MDEV) 185–6
 MPLS TE (MPLS Traffic Engineering) 116, 121–3, 205, 228–9, 231, 236

- MPLS TP (MPLS Transport Profile) 110, 123–4, 205, 230
MPLS VPNs (Virtual Private Networks) 110–11, 114–17, 118–21, 227, 237
MTIE 179, 182–5, 195–8
multihoming 39, 107, 114, 227–8, 242, 245–7
multiple access **130**
MWR Microwave Radio 77, 88–9, 156, 200, 240–2, 252–3, 262, 278, 280, 288, 308–9

NAS, Non-Access Stratum signaling 35, 40, 43, 58, 62, 246, 311
NBAP, Node B Application Part 43, 47, 49–50, 97, 107, 233, 245, 270, 283–5, 287, 310
net present value (NPV) 349, 361
network costs 8, 9, 18
network interface device (NID) 162
network interfaces 131, 132, 136, 149, 153–4, 362–3
network plans 144, 368
network topology 10, 22, 144
networks tiers 7, **8**, 9, 10
NTP 173–6, 181, 187

observed time difference of arrival
(OTDOA) 169–70
OFDM 168, 170
one tunnel, direct tunnel 39, 41–2
operational costs (opex) 6, 10, 13, 18, 19, 20, 24, 25, 34, 163, 349, 361
optical transport **150**
Optical Transport Hierarchy (OTH) **135**
OSI ISO-model **128**
Optical Transport Network OTN **135**, 169–72, 179, 181, 203
overlay (network) 25, 353, 357

packet delay 223, 296, 336, *see also* frame delay
packet delay variation (PDV) 180, 182, 192, 197–9, *see also* frame delay variation
packet rate 173–5, 187, 190, 199
packet selection 190–2, 197
packet synchronization principles 187–91
packet time error sequence 192–7
path max retrans 246
PCP, priority code point 87–8, 91–2, 271–2, 276
PDH transport 10, **130**, 169–73, 184, 199, 201 PDV, *see* packet delay variation
PHB, per-hop-behaviour 264–6, 271–2, 286, 299, 300
physical network topology 10, 22
pktfilteredMTIE 196–7, 200
PON Passive Optical Network **152**, 180, 201
ppb 167, 182
ppm 182
primary reference clock (PRC) 171, 184
primary reference source (PRS) 171
primary reference time clock (PRTC) 202
priority, of PTP grandmaster 176–7
priority, of traffic class 200
ProfileIdentifier (in PTP) 178
protection switching 87, 91, 123, 204–5, 214, 216, 230, 234, 235, 244
protection, microwave radio **145**
protection, SDH **134**
proxy mobile IP 57, 61
pseudowire, pseudowire emulation 49, 110, 117–18, 120–1, 123–4, 138, 230, 270, 273, 281–2, *see also* CES
PTP (IEEE1588) 173–8, 181–2, 187, 200–3

RAB (radio access bearer) 54, 104, 282
RANAP (radio access network application part) 35, 37, 39–40, 42, 51–2, 97, 104, 107, 245–6, 282, 309–11
recovery 50, 91, 100, 116, 204, 206, 208–9, 214–15, 217–18, 222–9, 236, 248, 315, 333–4
redundancy in PTP 178
regulations 142, 348
retransmit, retransmission 30, 37, 56–7, 58, 103, 218–19, 250, 256, 260–3, 279–81, 288–90, 293–4, 298
RNSAP (radio network subsystem application part) 51, 107

S1 -AP, S1-MME 57, 59, 60, 64, 237, 246, 270, 298, 300, 312
SDH/Sonet 10, 49, 78–83, 90, 117, 122–4, 131, 169–73, 179–80, 184, 204–5, 207–8, 214–16, 228–30, 239, 244, 248, 352–5
SDH, next-generation 71, 78–9, 83, 136, 214, 216, 352–3
SEC 179–80, 184–5, 197–200
selected packet time error sequence 192–7
service level agreement (SLA) 156, **159**–60, 162, 273, 363
service performance **161**
SHDSL **149**, 201
(single user) peak rate 14, 39, 58, 76, 349, 350, 364

- single-frequency network (SFN) 168–70
slave 174, 176–81, 187–9, 192, 198, 200
SRVCC, single radio voice call continuity 63
sync message 176–8
synchronization metrics for packet timing 192–8
synchronization metrics for TDM 182–6
synchronous Ethernet 179–82, 202

TDD 129, 142, 168, 170, 201
TDEV 179, 185–6, 193–5, 197
TD-SCDMA 170
telecom profile (for IEEE 1588) 177–81
testing synchronization 172, 197–9
TICTOC 181
tiers 8, 10, 155
time accuracy 168–9
time error 169, 174, 176–7, 183, 185–6, 191–4
time stamp 173–4, 176, 187–8, 191–2
time synchronization 168, 174, 201–3
total cost of ownership (TCO) 349, 361

traffic class 255, 277, 279
traffic class, Ethernet 272
traffic class, IPv6 109, 265
traffic class, MPLS 109, 207
traffic class, UMTS 283
traffic forecasts 12–15, 348, 352–3, 364, 367
traffic peak rate 14
transport out-sourcing 26, 352, 360–3, 366–8
troubleshooting 84, 86, 216, 307
two rate, three color marker (trTCM) 160

UMTS bearer 282–3
user network interface (UNI) 157–61, 273–4

VDSL 149, 190, 201
VoLTE, voice over LTE 63

wavelength division multiplexing (WDM) 151
WiMAX 168–70
wireless backhaul 138