

Verteilte Nutzungsdatenerfassung und nachgelagerte Weiterverarbeitung der Nutzungsdaten im Mobil Internet

Von der Fakultät Informatik, Elektrotechnik und Informationstechnik
der Universität Stuttgart zur Erlangung der Würde
eines Doktor-Ingenieurs (Dr.-Ing.) genehmigte Abhandlung

vorgelegt von

Jürgen Matthias Jähnert

geb. in Stuttgart Bad-Cannstatt

Hauptberichter:	Prof. Dr.-Ing. Dr. h. c. mult. Paul J. Kühn
Mitberichter:	Prof. Dr. Burkhard Stiller, Universität Zürich
Tag der Einreichung:	17. Mai 2006
Tag der mündlichen Prüfung:	27. März 2007

Institut für Kommunikationsnetze und Rechnersysteme
der Universität Stuttgart

2007

Para Angela y Marina.

Inhaltsverzeichnis

Inhaltsverzeichnis.....	i
Zusammenfassung.....	v
Abstract	vii
Abkürzungsverzeichnis	ix
1 Einleitung	1
1.1 Ziel und Abgrenzung der Arbeit	2
1.2 Übersicht über die Arbeit	3
2 Netztechnische Grundlagen.....	5
2.1 Das Internet	5
2.1.1 Das Internet-Protokoll	6
2.1.2 Dienstgüte im Internet	9
2.1.3 Internet Transportprotokolle	13
2.1.4 Mobilität im Internet	15
2.1.5 Sitzungskonzepte im Internet.....	21
2.2 Das Mobilkommunikationsnetz	24
2.2.1 Die Evolution der Mobiltelefonie.....	25
2.2.2 UMTS.....	29
2.3 Zusammenfassung.....	35
3 Ökonomische Aspekte eines Netzbetreibers	37
3.1 Einleitung	37
3.2 Grundlagen der Preispolitik	38
3.2.1 Relevante Marktformen.....	39
3.2.2 Überblick über Preisstrategien	40
3.2.3 Preisbildungsmethoden und Preisbildungsmechanismen.....	44
3.2.4 Rolle des Preises.....	46
3.3 Marktorientierte Preisbildung	47
3.4 Produktbündelung	48
3.5 Existierende Preismodelle in der Telekommunikation	49
3.5.1 Preisstruktur im Festnetzbereich	49
3.5.2 Preisstruktur im Mobiltelefonbereich.....	50
3.5.3 Preisstruktur im Internet.....	52
3.6 Kostenstruktur eines Netzbetreibers.....	53
3.6.1 Fixkosten	54
3.6.2 Variable Kosten.....	54
3.6.3 Weitere Kosten	55
3.7 Kostenmodell	56
3.8 Zusammenfassung und Ausblick	58
4 Strategien, Konzepte und Mechanismen zur Kommerzialisierung des Mobiltenet59	
4.1 Einleitung	59
4.2 Mobilitätsaspekte	59
4.2.1 Endgerätemobilität	59
4.2.2 Benutzermobilität	60

4.2.3	Dienstmobilität	60
4.2.4	Definition und Begriffe	60
4.3	Authentisierung, Autorisierung und Accounting	63
4.3.1	Authentisierung	63
4.3.2	Autorisierung.....	65
4.3.3	Aufbereitung der Nutzungsdaten	65
4.3.4	Charging	65
4.3.5	Auditierung.....	66
4.3.6	Nutzungsdatenerfassung.....	66
4.4	IETF AAA-Architektur und angrenzende Standards	66
4.4.1	Diameter	67
4.4.2	Die AAAArch-Arbeitsgruppe der IRTF	72
4.4.3	PANA	73
4.4.4	EAP	74
4.4.5	RTFM	75
4.4.6	Policy-based Accounting.....	77
4.5	Internet Protocol Data Record.....	78
4.6	Abrechnungsinfrastruktur in der ITU.....	82
4.7	2G und 3G Charging-Protokolle	83
4.7.1	GPRS-Nutzdatenerfassung.....	83
4.7.2	3GPP-Abrechnungssystem.....	84
4.7.3	OSA/PARLAY	85
4.8	Die Moby Dick-Architektur	87
4.9	Existierende Vorschläge für Internet-Preismodelle	92
4.9.1	Flat-Rate	93
4.9.2	Prioritätsbasierte Preisbildung.....	94
4.9.3	Kapazitätsallokation	94
4.9.4	Paris Metro Pricing.....	95
4.9.5	Smart Market.....	95
4.9.6	Responsive Pricing.....	96
4.10	Zusammenfassung und Beschreibung des weiteren Vorgehens	97
5	Annahmen und Problemdiskussion zum Mobilien Internet	99
5.1	Einleitung	99
5.2	Evolution und Konvergenz.....	99
5.3	Aspekte einer kommerziellen B3G-Infrastruktur.....	100
5.3.1	Benutzerorientierte Sichtweise.....	100
5.3.2	Netzbetreiberorientierte Sichtweise	101
5.3.3	Regulatorische Sichtweise.....	101
5.4	Randbedingungen eines B3G-Preismodells	101
5.4.1	Roaming zwischen Betreibern	101
5.4.2	Kundenbeziehungen	102
5.4.3	Prepaid.....	103
5.4.4	Preispolitik	104
5.4.5	Benutzerverwaltung	105
5.4.6	Zukünftige Dienststruktur	106
5.5	Preismodell für ein B3G-Netz.....	106
5.6	Problemdiskussion	111
5.7	Einschränkungen und Schwächen der existierenden Konzepte	113
5.8	Zusammenfassung.....	114

6	Vorschläge zur Weiterentwicklung von Mechanismen im kommerzialisierten Mobilem Internet	117
6.1	Einleitung	117
6.2	Entwurf einer Signalisierungsarchitektur	117
6.2.1	Schlüsselszenarien für das zukünftige Mobile Internet	117
6.2.2	Genereller Ansatz	124
6.2.3	Prinzipielle Konfigurationssignalisierung	124
6.2.4	Prinzipielle Signalisierung der Nutzungsdaten	125
6.2.5	Sicherheitsaspekte	126
6.3	Vorschlag für ein Sitzungskonzept	126
6.3.1	Beschreibung des Sitzungskonzeptes	126
6.3.2	Definition der Sitzungsberichte	129
6.3.3	Sitzungsverwaltung	130
6.3.4	Vertraulichkeit	132
6.3.5	Benutzerklassen Management	133
6.3.6	Datensatz der Nutzungsdaten	133
6.4	Zugangsroutearchitektur	134
6.4.1	Meter und Meter-Manager	134
6.4.2	Skalierbarkeit von Meter und Meter-Manager	135
6.4.3	AAA-Manager	135
6.4.4	Skalierbarkeit der Registrierung	136
6.4.5	QoS-Manager	136
6.4.6	Skalierbarkeit des QoS-Managers	137
6.4.7	Mobility-Manager	137
6.4.8	Skalierbarkeit des Mobility-Managers	137
6.5	Integration des Gesamtsystems	137
6.5.1	Nutzungsdatenerfassung und nachgelagerte Weiterverarbeitung	137
6.5.2	Skalierbarkeitsbetrachtungen des Gesamtsystems	138
6.6	Zusammenfassung	138
7	Validierung und Verifikation des Gesamtsystems	141
7.1	Einleitung	141
7.2	SDL und MSC	141
7.2.1	MSC-Basiskonstrukte	142
7.2.2	Telelogic TAU	143
7.3	Modellierung	143
7.3.1	Ziel der Modellierung	143
7.3.2	Teilnehmer und Betreibermodell	144
7.3.3	Netzmodell	144
7.3.4	Mobilitätsmodell	145
7.3.5	Lastmodell	145
7.3.6	Vertragsmodell	146
7.3.7	Strategien zur Verwaltung der Nutzungsdaten	146
7.4	Referenzberechnung	147
7.4.1	Nutzungsdatenaufkommen	147
7.4.2	Bandbreitenabschätzung	149
7.5	Parametrisierung im Gesamtmodell	150
7.5.1	Grundsätzliche Überlegungen	150
7.5.2	Parametrisierung bei Benutzern mit Prepaid-Vertrag	150
7.6	Analyse der Ergebnisse	152
7.6.1	Fokus der Analyse	152

7.6.2	Validierung der Spezifikation und Durchführung der Leistungsbewertung	152
7.6.3	Quantitative Reduktion der Nutzungsdaten	157
8	Zusammenfassung und Ausblick.....	159
8.1	Zusammenfassung der Ergebnisse.....	160
8.2	Empfehlungen an Netzbetreiber	161
8.3	Ausblick.....	162
	Literaturverzeichnis.....	163
	Anhang A: Signalspezifikation.....	169
A.1	Nachrichtenformate zwischen Komponenten.....	169
A.2	Nachrichtenformate innerhalb des Zugangsrouters.....	170
	Anhang B: Testprotokolle.....	172
B.1	Statische Tests.....	172
B.2	Dynamische Tests.....	177
	Anhang C: SDL – Spezifikation.....	180

Zusammenfassung

Die hier vorgelegte Arbeit beschäftigt sich mit einer verteilten Nutzungsdatenerfassung und nachgelagerter Aufbereitung der erfassten Daten im Mobil Internet. Am Anfang dieser Arbeit steht logisch die These, dass der global zu beobachtende Konvergenzprozess, der internetartige Datenkommunikation und telefonorientierte Telekommunikation, und im Endstadium auch der Verteilungskommunikationssysteme schließlich zu einer kohärenten Infrastruktur zusammenführen wird. Diese Infrastruktur unterstützt jegliche Art von Mobilität und wird im weiteren Verlauf dieser Arbeit „das Mobile Internet“ genannt. Dieser Konvergenzprozess besteht im Wesentlichen aus einem „Sieg des Internets“, d.h. Telefon (Mobiltelefon) und Fernsehen „migrieren“ zur Internet-IP-Technologie.

Direkt im Anschluss an diese Ausgangsthese wird eine weitere These erstellt, welche als Konvergenztechnologie für diesen Prozess ganz wesentlich Mobile IP - insbesondere der Version 6 - erwartet. Von der Spezialisierung „Mobile IPv6“ werden dabei aber eigentlich nur die klassischen Eigenschaften von IP, vermehrt um einige geeignete Mobilitätsmechanismen und die entsprechende Integration in ein Gesamtsystemkonzept, benötigt.

Zu diesen Thesen treten als nächste Voraussetzungen dieser Arbeit eine gewissermaßen „natürliche“ Forderung und zwei Feststellungen: „Das zukünftige Mobile Internet muss aus kommerzieller Sicht erfolgreich sein“ und „das Internet ist aktuell - im Vergleich mit dem existierenden (mobilen) Telekommunikationsnetz - aus rein kommerzieller Sicht nicht ausreichend erfolgreich“, sowie „die klassischen Telefonnetze - fest und mobil - arbeiten profitabel“. Die oben erwähnte und auf den ersten Blick sehr banal lautende Forderung ist, wie im weiteren Verlauf dieser Arbeit gezeigt wird, nicht problemlos in die existierende Internet-Architektur integrierbar, da im Internet kommerzielle Elemente historisch von nachrangiger Bedeutung waren. Dies liegt vor allem an den unterschiedlichen Konzepten und Architekturen. Als Folge hiervon hat sich im Internet ein „Always-On“ Paradigma herausgebildet, welches grundsätzlich nicht mit den kommerziell erfolgreichen Konzepten der Mobiltelekommunikation zu vereinbaren ist. Weiter erfolgt im Internet keine Trennung zwischen Signalisierungsverkehr und Datenverkehr, was den gesamten Abrechnungsprozess in einem Mobil Internet erschwert.

Somit wird in dieser Arbeit als nächster Schritt eine akzeptierte Erkenntnis aus der Betriebswirtschaftslehre, genau genommen aus dem Marketing, angenommen: Betriebswirtschaftlich erfolgreiche und in vergleichbaren Sektoren bewährte Strategien und Konzepte sind „kundenbezogene“ Preisdifferenzierung und Produktbündelung, welche eine benutzerindividuelle Festsetzung des Preises ermöglichen. Weiter ist im Telekommunikationsbereich unter dem Marketing-Aspekt „Preispolitik“ auch die Prepaid-Methode ein Element, welches in diesem Zusammenhang nicht unberücksichtigt bleiben darf. Der Vorteil der Prepaid-Methode liegt neben dem Liquiditätsgewinn, welcher im Vergleich zum traditionellen Bezahlverfahren Liquidität vor Leistungserbringung für den Netzbetreiber sichert, darin, dass die Benutzerakzeptanz dieses Modells für bestimmte Benutzergruppen gesteigert wird, und dass in diesem Marktsegment Rechnungen in der Regel nicht offen bleiben können. Allerdings soll an dieser Stelle nicht verschwiegen werden, dass die Prepaid-Methode sehr viel schärfere Anforderungen an die gesamte Abrechnungsinfrastruktur mit sich bringt, da der Gesamtprozess nun zeitkritisch wird.

Es werden schließlich noch zwei Feststellungen zur Menge der Eingangssätze hinzugefügt: „Telefonnetze beherzigen die zitierten Erkenntnisse aus dem Marketing“ und „das feste Internet vernachlässigt den Lehrsatz bisher“ (was natürlich auch für das Mobile Internet gilt).

Es besitzt entsprechend auf technischer Ebene dazu insbesondere kein dynamisches, benutzerspezifisches Konzept zur Erfassung und Weiterverarbeitung der Nutzungsdaten, welches die technische Grundlage schaffen würde, eine Infrastruktur bereitzustellen, welche erforderlich wäre, um im Internet eine Profitabilität vergleichbar mit der des Telefonnetzes zu erlangen. Es soll an dieser Stelle ausdrücklich erwähnt werden, dass die geforderte Infrastruktur als notwendige, nicht jedoch als hinreichende Voraussetzung im erwähnten Kontext betrachtet wird. Mit diesen Sätzen als Voraussetzung lässt sich der Beitrag dieser Arbeit wie folgt formulieren:

Um zu einer erhöhten Profitabilität im Mobil Internet zu gelangen, wird eine Mobile Internet-Architektur so erweitert, dass Mechanismen zur Unterstützung von Preisdifferenzierung und Produktbündelung sowie für die Prepaid-Methode in einer Art unterstützt werden, dass die eingangs erwähnte kohärente Infrastruktur - dieses Mobile Internet - diesen Schritt in Richtung Kommerzialisierung grundsätzlich vollziehen kann. Diese erforderlichen Mechanismen werden technisch realisiert durch eine verteilte Erfassung der Nutzungsdaten, eine in Echtzeit operierende nachgelagerte Aufbereitung der Nutzungsdaten und durch ein benutzerspezifisches Mobilitätskonzept für Sitzungen in Multi-Betreiber-Szenarien mit heterogenen Technologien (multi-provider/multi-technology).

Das in der Arbeit verfolgte Kernkonzept beruht auf der Strategie, die rohen Nutzungsdaten nur zu erfassen, wenn sichergestellt ist, dass Sie für den Abrechnungsprozess benötigt werden. Da im Markt verschiedene Preismodelle angeboten werden, kann hier keine globale Strategie der Nutzungsdatenerfassung in Betracht gezogen werden, sondern erfordert eine benutzerindividuelle und an das jeweilige Vertragsmodell gekoppelte Konfiguration und Parametrisierung der Abrechnungsinfrastruktur. Durch die Berücksichtigung des Mobilitätsaspektes erfordert die Umsetzung dieses Kernkonzeptes eine zentral administrierte und verteilt arbeitende Infrastruktur, welche die Nutzungsdaten entsprechend benutzerspezifisch erfasst und weiterverarbeitet. Hier sind Übergänge zwischen Netzinfrastrukturen sowohl eines Netzbetreibers, als auch zwischen verschiedenen Netzbetreibern entsprechend zu berücksichtigen.

Das Gesamtsystem wurde zunächst konzeptionell erarbeitet und im Anschluss in einer MSC/SDL-Umgebung spezifiziert und simulativ validiert. Hierfür wurde ein Werkzeug verwendet, welches vor allem bei der Spezifikation des aktuellen Mobilkommunikationsnetzes von einigen Herstellern erfolgreich eingesetzt wurde, jedoch in dieser Form im Internet eher weniger verbreitet ist.

Abstract

The work presented here is concerned with the acquisition of distributed utilisation data in the Mobile Internet and the downstream processing of this acquired usage data. At the beginning of this work, it was logically assumed that the globally observed convergence process of Internet-like data communications and telephone-oriented telecommunications and also, in the final stages, the broadcast communication systems would finally merge to a coherent infrastructure. Further, it is assumed, that these coherent infrastructure is supporting any kind of mobility and therefore the term ‘The Mobile Internet’ is introduced within this thesis. This convergence process consists primarily of a “Conquest of the Internet” - i.e., telephony and television migrating to Internet IP-technology.

In direct connection with this initial assumption, a further assumption was made, in which it was expected that Mobile IP, especially version 6, would be quite important as convergence technology for this process. Despite this customised ‘Mobile IPv6’, only the classic features of IP, augmented with some suitable mobility mechanisms were actually required for a suitable integration into a total system concept.

Next to these assumptions and conditional to this work, there appear two statements and a, to some extent, ‘natural’ demand: “The future Mobile Internet must, from a commercial point of view, be effective” and “The Internet is currently, from a purely commercial point of view, not sufficiently successful”, as well as “The classic telephone network - fixed and mobile - operates profitably”. In the further course of this work, the last mentioned and, at first sight, very simple sounding demand will be shown to have problems of integration into the existing architecture, since, historically, the commercial elements in the Internet were irrelevant. Reason for these problems is based on the fundamentally different concepts and architectures of the mentioned systems. In detail, in the Internet there is the “always-on” paradigm successfully implemented which basically does not fit into the commercially successful concepts of the (mobile) telecommunication sector. Additionally, in the Internet there is no clear separation between data and control, as currently observed in the traditional telecommunication network, which will further complicate the overall billing process in a future Mobile Internet.

Consequently, as a next step in this work, an acceptable awareness of business economics taken specifically from marketing, is being adopted: “Business successes in comparable sectors come from seasoned strategies and the concepts of ‘customer friendly’ price differentials and product bundling”. Furthermore, in this context, the prepayment approach is, in the realm of telecommunications under the marketing aspect of price politics, an element that must also be considered. The bid advantage of prepayment is the gain of cash ratio, which drops liquidity into the operator before the service is delivered. A further advantage is the increased user acceptance of this conceptual model at some customers and, finally, this model conceptually avoids any outstanding bills, which might be encashed in an expensive process, or might remain unpaid at all. However, it should be mentioned, that the concept of prepayment in a Mobile Internet brings in some increasing requirements to the overall billing infrastructure since the overall process is then transferred to a real-time process.

Finally, two further statements are being added to the host of propositions: “Telephone networks heed the findings cited from marketing” and “Up till now, the fixed Internet has neglected these propositions” (which, of course is also valid for the Mobile Internet). Accordingly, it exhibits, on a technical level, no particular dynamic user-specific concept for the acquisition and subsequent processing of the utilisation data, which would create the

technical basis on which to produce an infrastructure that would be necessary for the Internet to achieve profitability comparable to the telephone network. With all these propositions as preconditions, this work allows the following formulation to be made:

In order to arrive at a higher profitability in the Mobile Internet, a Mobile Internet architecture will be extended in a way that mechanisms to support price differentiation and product/service bundling as well as some sort of prepayment method shall be supported, so that the initially mentioned coherent infrastructure - this Mobile Internet – can take this step in the direction of commercialisation in order to be able to eventually replace the traditional voice network. These necessary mechanisms were technically realized through a distributed acquisition of the utilisation data, in a real-time downstream processing of the utilisation data and through a user-specific mobility concept for sessions in multi-operator scenarios with heterogeneous technologies (multi-provider/multi-technology).

The proposed solution follows the key concept of gathering only the utilization data really required for the overall billing process. Since in a market several products are offered represented by different service bundles, which are user-centric, there is no way to deploy a global strategy how to gather utilization data at all. So, a user-individual configuration and parameterisation of the overall billing system is required, which is closely related to the contract a customer has made with the operator. In a mobile environment as represented by the Mobile Internet, the deployment of such a strategy requires a centralized administration of a completely distributed system for metering, accounting and post-processing the gathered utilization data. Within this context, both, inter-domain and intra-domain scenarios have to be treated according to their specific requirements.

The complete system has been first designed and then formally specified, and validated by a simulation using a commercial MSC/SDL environment. The used tool and thus, the adopted methodology has been widely used in the specification of the 2G mobile telecommunication network by some well known vendors. In the Internet community the adopted method is rather new, but equally valuable.

Abkürzungsverzeichnis

1G	Mobilfunktechnologie der ersten Generation
2G	Mobilfunktechnologie der zweiten Generation
3G	Mobilfunktechnologie der dritten Generation
B3G	Beyond 3G
AAA	Authentisierung, Autorisierung und Accounting
AAA.f	Foreign AAA Server
AAA.h	Home AAA Server
AF	Assured Forwarding
AR	Access Router
API	Application Programming Interface
BE	Best Effort
BSC	Base Station Controller
BTS	Base Transceiver Station
BU	Binding Update
B3G	Beyond 3G
CAMEL	Customized Applications for Mobile Enhanced Logic
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CHAP	Challenge Handshake Authentication Protocol
CL	Controlled Load
CM	Kabelmodem
CMTS	Cable Modem Termination System
CN	Correspondent Node
CoA	Care-off Address
CRM	Customer Relationship Management
DAD	Duplicate Address Detection
DiffServ	Differentiated Services
DNS	Domain Name Service
DOCSIS	Data Over Cable Service Interface Specification
DRNC	Drift Radio Network Controller
DSCP	Differentiated Services Code Point
DSL	Digital Subscriber Line
EAP	Extensible Authentication Protocol
ECSD	Enhanced Circuit Switched Data
EDGE	Enhanced Data Rates for the GSM Evolution
EF	Expedited Forwarding
EGPRS	Enhanced General Packet Radio Service
ETSI	European Telecommunication Standards Institute
EUI	Extended Unified Identifier
FA	Foreign Agent
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
FQDN	Fully Qualified Domain Name
FP	Format Prefix
GSM	Global System for Mobile Communication
GPS	Global Positioning System
GPRS	General Packet Radio Services
GGSN	GPRS Gateway Support Node
GS	Guaranteed Service

GSMK	Gaussian Minimum Shift Keying
GTP	GPRS Tunnelling Protocol
HA	Home Agent
Ha	Heimadresse
HIP	Host Identity Payload
HLR	Home Location Register
HO	Handover
HSCD	High-Speed Circuit Switched Data
HSS	Home Subscription Server
HTTP	Hyper-Text Transfer Protocol
ID	Identification
IETF	Internet Engineering Task Force
IHL	IP Header Length
IMEI	International Mobile Equipment Identity
IMSI	International Mobile Subscriber Number
IMT	International Mobile Telecommunication
IntServ	Integrated Services
IOS	Internet Operating System
IP	Internet Protocol
IPDR	Internet Protocol Data Record
IRTF	Internet Research Task Force
ISDN	Integrated Services Digital Network
ISO	International Standardisation Organisation
ISP	Internet Service Provider
Iu	Utran CN Interface
Iub	UTRAN Node B – RNC Schnittstelle
Iur	Schnittstelle zwischen 2 RNCs
JAIN	Java API for Integrated Networks
JDBC	Java Database Connectivity
KDC	Key Distribution Center
LAN	Local Area Network
LLC	Logical Link Control
MAC	Medium Access Control
MAN	Metropolitan Area Network
MD5	Message Digest Algorithm
MIB	Management Information Base
MISP	Mobile Internet Service Provider
MN	Mobile Node
MSC	Message Sequence Chart
MSC	Mobile Switching Center
MSI	Modulation Symbol Index
NDM-U	Network Data Management – Usage
NLA	Next-Level Aggregation
NPA	Network Provider Address
NTP	Network Time Protocol
OSA	Open Services Access
OSI	Open Systems Interconnection
ODMA	On Demand Multiple Access
PA	Paging Agent
PAA	PANA Authentication Agent
PaC	PANA Authentication Client

PANA	Protocol for carrying Authentication for Network Access
PAP	Password Authentication Protocol
PAM	Presence and Availability Management
PDCP	Packet Data Convergence Protocol
PDP	Packet Data Protocol
PDU	Protocol Data Unit
PHB	Per-hop Behaviour
PIN	Personal Identification Number
PMP	Paris Metro Pricing
RNC	Radio Network Controller
ROI	Return on Investment
RSVP	Resource Reservation Protocol
RTFM	Real-Time Flow Measurement
RTSP	Real-Time Streaming Protocol
SCN	Service Class Name
SDL	Specification and Description Language
SDP	Session Description Protocol
SDU	Service Data Unit
SFID	Service Flow ID
SGSN	Service Gateway Support Node
SIM	Subscriber Identity Module
SIP	Session Initiation Protocol
SLA	Site-Level Aggregation
SMS	Short Message Service
SNMP	Simple Network Management Protocol
SPI	Security Parameter Index
SRNC	Service Radio Network Controller
SS7	Signalling System 7
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TCP	Transmission Control Protocol
TIA	Telecommunications Industry Association
TLA	Top-Level Aggregation
TMSI	Temporary MSI
TSN	Transmission Sequence Number
UAC	User Agent Client
UAS	User Agent Server
UDP	User Data Protocol
UTRAN	UMTS Radio Access Network
UML	Unified Modelling Language
UMTS	Universal Mobile Telecommunication System
URL	Uniform Resource Locator
USSD	Unstructured Supplementary Services Data
Uu	UTRAN Luftschnittstelle
VHE	Virtual Home Environment
VPN	Virtual Private Network
WAN	Wide Area Network
W-CDMA	Wideband Code Division Multiple Access
WLAN	Wireless LAN
XML	Extensible Markup Language

1 Einleitung

Im Sommer 2000 führte die Regulierungsbehörde für Post und Telekommunikation in der Bundesrepublik Deutschland eine historische Versteigerung durch, die den Kassen der öffentlichen Hand ca. 50 Mrd. € zuführte [126]. Es wurden die für die nächste Generation der Mobilien Kommunikationsnetze erforderlichen Frequenzbänder versteigert. Die auf diesen Frequenzen aufbauenden UMTS (Universal Mobile Telephone System)-Netze sollen unsere Gesellschaft mit neuen Kommunikationsformen ausstatten, von denen allgemein erwartet wird, dass sie einen sehr starken Einfluss auf unser tägliches Leben in der nahen Zukunft haben werden.

Die Entwicklung, die diese Versteigerung erst ermöglichte, ist auf zwei historische Entwicklungen zurückzuführen. Zum einen ist es der Erfolg des Internets in den letzten 10 Jahren und parallel dazu erfuhr das Mobiltelefon sehr große Marktakzeptanz und hat sich innerhalb sehr kurzer Zeit zu einem für viele Menschen unverzichtbaren Begleiter im täglichen Leben entwickelt.

Schon seit Jahren gibt es nun in der Telekommunikationsindustrie Bestrebungen, die zurzeit noch getrennten Netze (Telefonnetz und Internet) auf einer einheitlichen Netzplattform zu integrieren. Der erste Versuch, das so genannte diensteintegrierte Digitalnetz (ISDN) wurde diesem Anspruch nie gerecht, und es wird nun in Form der UMTS-Technologie ein weiterer Versuch gestartet, der zusätzlich eine Mobilitätskomponente integriert.

Da Sprach- und Datennetze historisch gewachsen sind und auf völlig verschiedenen Paradigmen basieren, ist der Migrationprozess dieser beiden Netztechnologien im Netzzugangsbereich sowohl theoretisch als auch praktisch noch nicht abgeschlossen. Es darf schon an dieser Stelle prognostiziert werden, dass die aktuelle Architektur von UMTS noch nicht als Plattform für ein integrierendes Zugangsnetz betrachtet werden kann, da sie noch sehr stark von leitungsvermittelnden Architekturprinzipien dominiert wird und der Trend heutzutage von einem Architekturansatz ausgeht, in dem die paketorientierten Prinzipien dominieren werden.

Der bevorstehende Migrationsprozess [133] erfordert von den Betreibern der UMTS-Netze ein geeignetes Konzept, welches es ermöglicht, vor allem die traditionellen Internet Benutzer als Nutzer der konvergierenden Netzinfrastruktur einzubinden, damit das zukünftige Mobile Internet ein kommerziell tragfähiges und somit erfolgreiches System wird. Hier spielt die Bereitschaft, für mobile Dienste grundsätzlich bezahlen zu wollen, eine zentrale Rolle.

Somit ist es aus Sicht eines Netzbetreibers erforderlich, ein Preismodell zu implementieren und am Markt zu etablieren, das sowohl von den historischen Internet-Nutzern, als auch von den Kunden des Mobilfunknetzes akzeptiert wird. Gleichzeitig sollte die Infrastruktur in der Lage sein, neue Dienste auf die Netzplattform zu integrieren und jegliche Kosten zu minimieren. Um diese Kosten zu minimieren, muss zum einen ein entsprechend geeignetes Preismodell entwickelt werden und zum anderen ist ein geeignetes Konzept der Nutzungsdatenerfassung in einem heterogenen und verteilten Netz erforderlich.

Diese Arbeit geht von der These aus, dass die Internet Protokollarchitektur die zukünftigen Netzlandschaft sowohl im mobilen and auch festen Bereich dominieren wird. Im Anschluss an diese These wird im Rahmen dieser Arbeit die Forderung gestellt, dass auch dieses konvergierende Netz aus kommerzieller Sicht ähnlich profitabel sein muss, wie z.B. das Sprachnetz – vor allem vor dem Hintergrund der UMTS-Auktionen. Es ist in diesem Zusammenhang festzustellen, dass diese kommerziellen Aspekte im Internet bislang nicht aus-

reichend berücksichtigt wurden, was vor allem historische Gründe hat. Diese kommerziellen Aspekte sind jedoch in den Telefonnetzen ausreichend implementiert und die Umsatzzahlen (vor allem die Umsatzzahlen pro absoluter Netzlast) der traditionellen Sprachkommunikation können aktuell nicht mit denen aus reinem Datenverkehr verglichen werden.

Um jedoch eine kommerzielle Umgebung aufbauen zu können, müssen grundlegende ökonomische Mechanismen und Gegebenheiten berücksichtigt werden, die Benutzer gemäß ihrer Zahlungsbereitschaft behandeln [117]. Derartige Mechanismen und Konzepte sind im Telefonnetz implementiert, sind jedoch im Internet und vor allem im Mobilien Internet nicht ausreichend vertreten. Weiter muss in diesem Zusammenhang festgehalten werden, dass die Unterstützung von Mobilität sehr harte zeitkritische Anforderungen mit sich bringt und durch die Dynamik das Aufrechterhalten von Sitzungen und die zugehörigen technischen Mechanismen, welche eine Kommerzialisierung unterstützen, sehr komplex werden. An exakt dieser Stelle ist der Kernbeitrag dieser Arbeit zu sehen, in der in Form einer Spezifikation ein wesentlicher Beitrag zur Kommerzialisierung einer Mobile IP basierten Internet-Infrastruktur geleistet wird. So wird im Rahmen dieser Arbeit eine Architektur spezifiziert und validiert, welche die Erfassung der Nutzungsdaten in einem solchen Kontext unter Berücksichtigung verschiedener Bezahlformen in einer heterogenen Netzbetreiber-Umgebung effizient bewerkstelligt.

1.1 Ziel und Abgrenzung der Arbeit

Die vorliegende Arbeit verfolgt das Ziel, die Einführung eines Mobilien Internets zu fördern. Hierbei wird eine IP-basierte Netzarchitektur angenommen, die als Konvergenzschicht unterhalb der Vermittlungsschicht mehrere Technologien und deren dazugehörige Protokolle, als auch in den darüber liegenden Protokollschichten mehrere Protokolle für Sprach und Datenverkehr unterstützt. Wegen der bekannten Einschränkungen kann dies aktuell nur auf einer IPv6-basierter Infrastruktur ermöglicht werden, so dass die Vorzüge von IPv6 ebenfalls vorausgesetzt werden.

Kernproblem einer kommerziellen Einführung eines integrierten Mobilien Internets ist ein geeignetes Konzept, wie die dann paketbasierten Dienste, welche sowohl Sprach als auch Datendienste umfassen, entsprechend erfasst und abgerechnet werden. Hierzu sind zwei wesentliche Voraussetzungen erforderlich: Zunächst muss ein entsprechendes Geschäftsmodell erarbeitet werden, dass sowohl Anforderungen des traditionellen, aus der verbindungsorientierten Telefonie bekannten Sprachverkehr, als auch dem paketorientierten Internetverkehr genügt.

Die Bewertung eines solchen Konzeptes kann auch mehrere Blickwinkel erfordern. Zunächst gibt es den Blickwinkel einer Regulierungsbehörde, die gesamtwirtschaftliche Ziele verfolgt, weiter gibt es den Blickwinkel des Netzbetreibers, der marktwirtschaftliche Ziele wie Marktanteil und Gewinnerorientierung verfolgt. Hier ist zusätzlich aus Sicht des Marketings die Frage der Benutzerakzeptanz relevant. Als weiteren Blickwinkel sind aus Sicht der Endverbraucher die Faktoren Preis, Benutzerfreundlichkeit und Transparenz von Bedeutung.

Bricht man das generische Ziel eines Netzbetreibers der Gewinnmaximierung in detailliertere Komponenten auf, so ist es speziell bei dem Produkt der Telekommunikationsdienste wegen der charakteristischen Kostenstruktur erforderlich, neben einer effektiven Ressourcenauslastung, die Benutzerverwaltungskosten zu minimieren und zusätzlich das System derart offen zu halten, dass Mechanismen und Methoden des Marketings effektiv, flexibel und kostengünstig in ein Netz integriert werden können.

Benutzerverwaltungskosten sind unter anderem von der Komplexität eines Preismodells abhängig. Dieses Preismodell wird in der Regel von Marketing-Experten ohne Rücksicht auf technische Randbedingungen erstellt und ist nicht zentrales Thema dieser Arbeit. Nachdem dann ein solches Preismodell gebildet wurde, müssen die im Preismodell definierten Parameter auf technische Messgrößen abgebildet werden und die entsprechenden Messdaten, die aus der Nutzdatenerfassung generiert wurden, einem Benutzerverwaltungsprozess zugeführt werden, welcher durchaus nennenswerte Betriebskosten verursacht.

Die vorliegende Arbeit soll nun dazu beitragen, dass diese Messgrößen flexibel, effektiv und aufwandsminimiert in einer Mobilien Internet-Infrastruktur erfasst und weiterverarbeitet werden können. In dieser Infrastruktur dient das IP-Protokoll als technologische Konvergenzschicht. Mehrere Betreiber bieten auf Basis von ggf. unterschiedlichen Netztechnologien am Markt Netzdienstleistungen konkurrierend an.

Aus dieser Zielvorstellung werden dann in dieser Arbeit drei Kernanforderungen abgeleitet, die letztendlich in drei Vorschlägen enden, welche die Erfassung der Nutzungsdaten zum Zwecke einer Abrechnung für die Benutzer, also die Kommerzialisierung des Mobilien Internets, unterstützen.

Kernziel dieser Arbeit ist in diesem Zusammenhang, dass ein Netzbetreiber eines zukünftigen IP dominierten Mobilien Internets in die Lage versetzt wird, flexibel und kostenminimiert eine heterogene Netzinfrastruktur kommerziell anbieten zu können. Hierzu wird eine verteilte Architektur vorgestellt, die Nutzungsdaten in einer mobilen und heterogenen Internet-Infrastruktur, technologieunabhängig erfasst und effizient an die Orte weiterverteilt, an denen die jeweilige Aufbereitung dieser Daten zur Rechnungserstellung erfolgt.

Da sich der Fokus dieser Arbeit auf IP-basierte mobile Zugangsnetze beschränkt und es aktuell zwei verschiedene Versionen des IP-Protokolls gibt IPv4 und IPv6, soll im weiteren Verlauf dieser Arbeit IPv4 nicht weiter betrachtet werden, sondern ausschließlich die neuere Version 6 des Internet-Protokolls [35].

1.2 Übersicht über die Arbeit

In Kapitel 2 werden die Grundlagen der mobilen Internet-Architektur vorgestellt. Hier wird, ausgehend von den zwei unterschiedlichen Netztechniken paketvermittelnder Datennetze und leitungsvermittelnder Sprachnetze, eine Übersicht über die verschiedenen Technologien gegeben. Dieser Teil liefert somit die technologischen Grundlagen, um den Konvergenzprozess und die sich veränderten Anforderungen an die Abrechnungsinfrastruktur bzw. den Paradigmenwechsel, den der im Sinne dieser Arbeit sich vollziehende Prozess mit sich bringt, besser verstehen zu können.

Um eine Kommerzialisierung technisch zu unterstützen ist es notwendig, Kernaspekte eines kommerziellen Systems zu verstehen, um diese dann geeignet auf die technologische Plattform abbilden zu können. Somit werden in Kapitel 3 existierende ökonomische Aspekte betrachtet, welche das kommerzielle Internet unter einem betriebswirtschaftlichen Blickwinkel beleuchten. Dies beinhaltet neben einer Übersicht über elementare Erkenntnisse aus dem Marketing eine Analyse der Preismodelle aus dem Telekommunikationssektor und eine Betrachtung ausgewählter innerbetrieblicher Abläufe eines Betreibers eines solchen Mobilien Internets.

In Kapitel 4 wird eine Übersicht über existierende Mechanismen, Strategien und Konzepte gegeben, welche zu einer Kommerzialisierung dieses Mobilien Internet beitragen können. Teil-

weise sind diese Strategien, Mechanismen und Konzepte in anderen Umgebungen bereits implementiert oder sie sind als theoretische Modelle verfügbar.

In Kapitel 5 wird aufgezeigt, dass die in Kapitel 4 beschriebenen Mechanismen nicht ausreichend sind, eine mobile Internet-Architektur geeignet zu unterstützen.

Vorwegnehmend kann an dieser Stelle schon festgehalten werden, dass nach einer Problemdiskussion drei elementare Verbesserungsvorschläge unterbreitet werden, welche in Kapitel 6 detailliert vorgestellt werden.

Kapitel 7 geht auf die auf MSC/SDL-basierte Validierung näher ein und endet mit konkreten Empfehlungen an einen potentiellen Betreiber eines zukünftigen Mobilens Internets.

Kapitel 8 fasst die wichtigsten Ergebnisse der Arbeit zusammen und öffnet einen Blick für zukünftige Erweiterungen.

2 Netztechnische Grundlagen

In diesem Kapitel werden die für den weiteren Verlauf relevanten technischen Grundlagen behandelt. Dieses sind die fundamentalen Eigenschaften des Internet und des Telefonnetzes und die jeweils mobilen Ausführungen dieser Netze. Der Beitrag dieses Kapitels ist das Bereitstellen des technischen Basiswissens, auf welches im weiteren Verlauf dieser Arbeit zurückgegriffen wird. Dies ist vor allem vor dem Hintergrund des aktuellen Konvergenzprozesses von Internet und Telefonie bei mobilen Zugangnetzen von besonderer Bedeutung.

Schwerpunkt dieses Kapitels ist zunächst die Vorstellung der Basiskonzepte und der aktuell diskutierten Mechanismen zur Mobilitätsunterstützung im Internet, welche entworfen wurden, um aus dem existierenden Internet ein integriertes Netz zu formen, das in der Lage ist, sowohl die Mobilitätsbedürfnisse der Endkunden zu befriedigen, als auch das existierende leitungsvermittelnde Sprachnetz sukzessive zu ersetzen. Im Rahmen dieser Arbeit wird dazu das Mobile Internet Protocol (Mobile IP) ausschließlich in der Version 6 behandelt.

Den zweiten Schwerpunkt dieses Kapitels bilden die Basiskonzepte der aus der klassischen Telefonie weiterentwickelten Architektur der Mobiltelefonie. Hier wird hauptsächlich die Architektur der Dritten Generation (3G) betrachtet, da davon ausgegangen wird, dass ältere Konzepte bei der zu erwartenden Evolution und Konvergenz von Internet und mobilem Telekommunikationsnetz keine Rolle spielen. Dieses Kapitel endet mit einer Zusammenfassung der wesentlichen Punkte.

2.1 Das Internet

Das Internet in seiner heutigen Form ist ein weltumspannendes, dezentral organisiertes, offenes Rechnernetz das kein zentrales Management besitzt und aus einer großen Anzahl heterogener Netze zusammengesetzt ist. Damit diese Netzwerke langfristig gesichert miteinander kommunizieren können, vereinbart die Internet Society [61] Standards. Hier ist zwischen einer operativen bzw. administrativen Koordination und einer forschungsbasierten Koordination zu unterscheiden. Aus operativer Sicht organisiert die Internet Assigned Numbers Authority [55] die globale Funktionsweise des Internets. Als Standardisierungsgremium, in dem zukünftige Standards bzw. Modifikationen existierender Standards beschlossen werden, hat sich die Internet Engineering Task Force (IETF) etabliert [57].

Das Internet ist ein paketvermittelndes Netz und stellt dem Benutzer für eine Übertragung keine vorhersagbare und verlässliche Dienstgüte bereit. Der Dienst, der vom Internet-Protokoll (IP) bereitgestellt wird, ist unzuverlässig. Im aktuellen Internet hat jeder Teilnehmer die Möglichkeit, ohne Vorankündigung, jederzeit und permanent auf Netzressourcen zuzugreifen und konkurriert mit allen anderen Teilnehmern um die verfügbaren Ressourcen. Das Internet verfügt über keine benutzerspezifische Zugangskontrolle und lehnt somit einen Benutzer nie ab, stellt ihm jedoch ggf. einen unzureichenden Zugang zum Netz bereit. Dieses Paradigma wird auch als „Always-On“ Paradigma bezeichnet.

Diese Methode der Kommunikation ist auf der einen Seite sehr flexibel, da freie Ressourcen automatisch von jeder beliebigen Nachfrage kurzfristig und ohne Signalisierung konsumiert werden können, was einen Kostenvorteil darstellt. Kritiker des Internets verweisen gerne auf den Nachteil der ungenügenden Dienstgüte, der vom Nichtvorhandensein eines Reservierungsmechanismus herrührt.

2.1.1 Das Internet-Protokoll

Das Internet-Protokoll wurde in [34] in seiner ursprünglichen Form spezifiziert. Die Hauptfunktion des Internet-Protokolls (IP) ist die Übermittlung von Paketen (Datagrammen) zwischen Endsystemen. Für einen Sendeauftrag werden von den höheren Schichten ein Datenblock und eine Zieladresse an die IP-Instanz übergeben, die daraus ein Paket zusammensetzt und dieses über mehrere Netzknoten an die Zieladresse sendet. Dazu übernimmt das Internet-Protokoll die Aufgaben der Segmentierung von längeren Datenblöcken zu IP-Paketen vor. IP arbeitet verbindungslos mit unquittierten Datagrammen und ohne Fehlersicherung. Ein IP-Paket hat ein festes Format für die Steuerinformation (Header) und einen variablen Teil für die Information (Payload). Eine weitere wesentliche Eigenschaft des Internets ist, dass auf der Vermittlungsschicht nicht zwischen Signalisierungspaketen und Datenpaketen unterschieden werden kann.

2.1.1.1 Paketaufbau

Wie bereits erwähnt, besteht ein IPv6-Paket aus einem Kopf (Header) und einem Rumpf (Payload). Der Kopf kann optional mit Erweiterungen versehen werden. Ein IPv6-Kopf hat eine Länge von 40 Byte, die sich auf 8 verschiedene Funktionsfelder aufteilen. Der IPv6-Kopf steuert mit seinen Feldern die Funktion des Protokolls, ist in [37] definiert und in Abbildung 1 skizziert.

Version (4 bit)	Priority (8 bit)	Flow Label (20 bit)	
Payload Length (16 bit)		Next Hader (8 bit)	Hop Limit (8 bit)
IP Source Address (128 bit)			
IP Destination Address (128 bit)			
Payload			

Abbildung 1: IPv6-Paketkopf

Zu erwähnen ist hier, dass der Paketkopf bei IPv6 im Vergleich zu IPv4 nicht über eine Checksumme verfügt, um den Netzknoten (Router) die aufwendige Überprüfung und damit Rechenzeit zu ersparen. Ein Übertragungsfehler muss deshalb in den darüber liegenden Schichten erkannt werden. Die Bedeutung der Felder ist im Einzelnen:

- *Version* ist ein 4 Bit-Feld, das die IP-Versionsnummer angibt.
- *Priority* gibt an, mit welcher Priorität die Pakete auf dem Weg zum Empfänger behandelt werden. Hierbei haben die Router des Transportweges die Möglichkeit, je nach Netzauslastung und Belegung des Prioritätsfeldes, Pakete auf unterschiedlichen Wegen zum Empfänger zu senden, zu verwerfen oder bei Bedarf die Werte des Feldes zu verändern. Dieses Feld wurde in IPv4 auch als DiffServ Codepoint bezeichnet und ist voll zu [97] kompatibel.
- *Flow Label* wird benutzt, um Pakete zu kennzeichnen, für die eine besondere Behandlung durch IPv6-Router gewünscht wird. In diesem Fall wird eine eindeutige Kennzeichnung gesetzt.

- *Payload Length* enthält die Anzahl der Nutzdaten-Bytes, die in diesem IP-Paket transportiert werden.
- *Next Header* enthält den Code des diesem Paket folgenden Kopfes.
- *Hop Limit* gibt die maximale Anzahl von Netzknoten an, die das IP-Paket während der Verkehrslenkungsphase zwischen Sender und Empfänger durchlaufen darf. Dabei dekrementiert jeder Router, der das Paket weiterleitet, diesen Wert um eine festgelegte Schwelle. Steht Hop Limit auf dem Wert 0, wird das Paket verworfen. Der Netzknoten, der das Paket verworfen hat, sendet an den Absender eine Fehlermeldung. Dieses Konzept verhindert, dass nichtzustellbare Pakete beliebig lange im Netz kreisen.
- *Destination Address* gibt die 128 bit lange IPv6-Quelladresse an.
- *Source Address* gibt die 128 bit lange IPv6-Zieladresse an.

Durch die optionalen Erweiterungen des Kopfes, die bei IPv6 erstmalig verwendet werden, wird Rechenzeit gespart, da bei der Verkehrslenkung (engl. Routing) der IP-Pakete nur der Basis-Kopf von den Routern überprüft werden muss. Bei Verwendung mehrerer optionaler Köpfe werden diese einfach aneinandergesetzt. Allerdings ist hierbei auf die Reihenfolge zu achten, da jeder Kopf die Identifikation des nachfolgenden Kopfes enthält. Optionale Köpfe stehen zwischen dem Basis-Kopf und dem Rumpf. Im Folgenden werden ausgewählte optionale Köpfe vorgestellt:

- *Hop-by-Hop Header* ist der einzige Kopf, welcher von den Routern überprüft werden muss. Er wird verwendet, um Optionen zu transportieren, welche bei jedem Transportschritt ausgewertet werden müssen. Die Länge dieses Kopfes beträgt 8 Byte.
- *Routing Header*, mit dessen Hilfe bestimmt werden kann, welche Router das IP-Paket auf dem Weg von Quelle zur Senke das Paket passieren soll. Dieser Kopf hat eine Länge von 8 Byte.
- *Fragment Header* wird zur Fragmentierung von IP-Paketen verwendet. Falls die Nutzdaten für ein IPv6-Paket zu groß sind, wird mit Hilfe dieser Information das Paket in mehrere kleinere Einheiten geteilt, bzw. nach der Übermittlung wieder zusammengefügt.
- *Extension Header* fungiert als so genannter Erweiterungs-Kopf, d. h. bei Erweiterung von IPv6 können neue Optionen hinzugefügt werden.

2.1.1.2 IPv6-Adressierung

Abbildung 2 zeigt exemplarisch den Aufbau einer IPv6-Adresse. Das Format und die Aufteilung der 128-bit langen IPv6-Adressen ist in [37], [53] und [54] definiert. Die Adressen sind in einen 64-bit großen Verkehrslenkungs-Bereich und in einen 64-bit großen Teil für die Endknotenadressierung aufgeteilt.

48 bit	16 bit	64 bit = 24 bit + FFFE (16bit) + 24 bit
2001:0638:0202		Ethernet.....Address
Network provider address	SLA	EUI 64 bit

2001:0638:0202: 0 :0250:daff:fecf:c6ad

Abbildung 2: IPv6-Adressstruktur

Durch diese strikte Unterscheidung kann man in der IPv6-Netzstruktur hierarchische Verkehrslenkungsstrukturen aufbauen und damit die Effizienz der Verkehrslenkung erhöhen.

2.1.1.3 Verkehrslenkung

Die Verkehrslenkung, also die Bestimmung des Pfades für den Transport von Datenpaketen vom Sender zum Empfänger über mehrere Netzknoten hinweg, erfolgt mittels einer Verkehrslenkungstabelle. Hier wird im Router für jedes eintreffende Paket ermittelt, über welchen Anschluss die Pakete weitergesendet werden, wie in Abbildung 3 gezeigt.

Destination	Next Hop	Flags	Metric	Ref	Use	Iface
::1/128	::	U	0	0	0	lo
2001:638:202:1:204:76ff:fe9b:ed77/128	2001:638:202:1:204:76ff:fe9b:ed77	UAC	0	384	192	eth0
2001:638:202:1:204:76ff:fe9b:ede2/128	::	U	0	7247	2	lo
2001:638:202:1:800:20ff:fe88:fc3/128	2001:638:202:1:a00:20ff:fe88:fc3	UG	1	246	0	eth0
2001:638:202:1:a00:20ff:fe88:fc1/128	2001:638:202:1:a00:20ff:fe88:fc1	UAC	0	26481	1708	eth0
2001:638:202:1::/64	::	UA	256	3	0	eth0
2001:638:202:15:204:e2ff:fe3a:d24d/128	2001:638:202:15:204:e2ff:fe3a:d24d	UAC	0	7775	52	wlan0
2001:638:202:15:204:e2ff:fe3a:d258/128	::	U	0	311	2	lo
2001:638:202:15::/64	::	UA	256	0	0	wlan0
2001:638:202:16:201:3ff:fe8c:8d3f/128	2001:638:202:16:201:3ff:fe8c:8d3f	UAC	0	1785	6	eth2

Flags: U Netzschnittstelle ist aktiv; A Acknowledgement (Mobile IP); G nur globale Adressen werden unterstützt; C Kompressions-Flag

Metric: Spezifiziert die Kosten eines Weges; der kleinste Weg indiziert den günstigsten Weg.

Ref: Referenzzähler, welcher angibt wie viele „Verbindungen“ gleichzeitig erlaubt werden.

Use: Anzahl der Pakete, welche über die Schnittstelle versendet wurden.

Iface: Netzschnittstelle, welche benutzt werden soll. Hier steht eth für Ethernet, lo für local und WLAN für Wireless LAN. Die folgende Nummerierung erfolgt fortlaufend bei mehreren Schnittstellen einer Technologie.

Abbildung 3: IPv6-Verkehrslenkungstabelle

Ein wichtiges Konzept von IPv6 ist das Aggregierungsmodell [53]. Hier sollen durch ein hierarchisches Verkehrslenkungssystem IPv6-Adressen auf einzelnen Ebenen so zusammengefasst werden, dass nur innerhalb einer Ebene alle Verkehrswege bekannt sind. Auf diese Weise sollen die notwendigen Verkehrslenkungs-Tabellen klein gehalten werden, um die Rechenleistung, die für die Verkehrslenkung erforderlich ist, zu begrenzen. Der Verkehrslenkungs-Bereich wurde daher in drei Bereiche eingeteilt, welche die angestrebte hierarchische Struktur des IPv6-Adressraumes und des Aggregierungsmodells widerspiegeln sollen. Diese sind:

- Top-Level Aggregation (TLA),
- Next-Level Aggregation (NLA),
- Site-Level Aggregation (SLA).

Es wird ein Adresspräfix aus dem TLA-Bereich nur solchen Netzwerksystemen und Betreibern zugeordnet, die über globale Konnektivität verfügen. Der NLA-Bereich soll dazu benutzt werden, Routing-Hierarchien innerhalb des eigenen Betreibernetzwerkes aufzubauen und um Endeinrichtungen zu identifizieren. Der Bereich hierfür kann vom TLA-Inhaber in mehrere Teile mit frei definierbaren Größen aufgeteilt werden. Der SLA-Bereich schließlich bleibt den Endeinrichtungen vorbehalten, um in Ihrem eigenen Netzwerk Routing-Strukturen aufzubauen. Somit ergibt sich der in Abbildung 4 dargestellte Aufbau des Netzpräfixes eines IPv6-Paketes.

Netzknotenadressierung					Adressierung des Endsystems
48 bit + 16 Bit SLA					64 Bit
3	13	8	24	16	64
FP	TLA	Res	NLA	SLA	

2001:0638:0202: 0 :0250:daff:feef:c6ad

Res: Reserviert

Abbildung 4: IPv6-Routingkopf

2.1.2 Dienstgüte im Internet

Die Internet-Architektur arbeitet nach dem Bestmöglichkeitsprinzip. Hier ist die höchste vom Netz zur Verfügung gestellte Dienstgüte eine zuverlässige Datenanlieferung mit Protokollen der Transportschicht wie TCP. Dieses ist für traditionelle Datenanwendungen wie *ftp* und *telnet*, bei denen korrekte Datenanlieferung wichtiger als Rechtzeitigkeit ist, ausreichend. Geht man allerdings davon aus, dass das Internet zu einem Konvergenznetz ausgebaut wird, das sowohl sprachbasierte Dienste, die sehr viel konkretere Anforderungen an die Dienstgüte mit sich bringen, und traditionelle paketorientierte Dienste transportiert werden, wird schnell klar, dass eine korrekte und angemessene Anlieferung dieser Daten nicht annehmbar ist. Somit wurden Mechanismen spezifiziert, um Echtzeit-Dienstgüte zu unterstützen.

2.1.2.1 Das IntServ-Modell der IETF

Das Integrated Services Modell, in der Literatur und im Rahmen dieser Arbeit mit IntServ abgekürzt, wurde in [11] definiert. Bei dem Entwurf dieses Modells wurde vorausgesetzt, dass Ressourcen, wie z.B. Bandbreite, explizit von den Anwendungen verwaltet werden müssen, um geeignet auf die Anforderungen der Anwendungen eingehen zu können. Weiter werden die einzelnen Verkehrsflüsse, als die Sequenz logisch zueinander gehörender Datenpakete, unabhängig voneinander behandelt und können die Behandlung anderer Datenflüsse nicht beeinflussen. Beim IntServ-Modell sind zwei verschiedene Verkehrsklassen definiert:

- Guaranteed Service (GS) [118] und
- Controlled Load Service (CL) [131]

Diese Dienstklassen bilden zusammen mit einem Signalisierungsprotokoll die Basiskomponenten des IntServ-Modells. GS und CL spezifizieren hierbei die Behandlung der verschiedenen Verkehrsflüsse innerhalb eines Netzknotens auf dem Datenpfad zwischen Quelle und Senke. In Verbindung mit dem Ressourcen Reservierungsprotokoll (Resource Reservation Protocol kurz RSVP) [12] wird somit der Anwendung eine Ende-zu-Ende Dienstgüte garantiert zur Verfügung gestellt.

Beim IntServ-Modell werden die Netzressourcen reserviert, die einem zusammenhängenden Datenfluss eindeutig zugeordnet werden können. Diese Reservierung ist unidirektional. Prinzipiell werden Sitzungen von der Senke aus definiert. Die Parameter hierfür stellen die IP-Zieladresse, das verwendete Transportprotokoll und die Transportadresse, auch Port genannt, dar. Um eine bidirektionale Kommunikation mit einer garantierten Dienstgüte einzurichten, ist es notwendig, die Ressourcen in den beiden Richtungen unabhängig voneinander zu reservieren. Es ist bei diesem Modell erforderlich, dass alle Mechanismen in jedem auf dem Pfad zwischen Quelle und Senke liegenden Knoten implementiert sind.

Guaranteed Service (GS)

Der GS-Dienst ist für Anwendungen konzipiert worden, die eine strenge Garantie erfordern, dass ein Paket innerhalb einer festgeschriebenen Zeit, der maximal zulässigen Gesamtverzögerung, beim Empfänger ankommt. Hier sind besonders Audio und Videodaten sehr zeitkritisch, weil hier neben einer sicheren Ankunft die Rechtzeitigkeit des Eintreffens eines Paketes von wesentlicher Bedeutung ist, da ein zu spät eingetroffenes Paket nutzlos ist. Der Guaranteed Service basiert auf einer Abschätzung eines Flussmodells und hat folgende Eigenschaften:

- Maximal zulässige Verzögerung bei einer Ende-zu-Ende Übertragung
- Garantierte Übertragungsbandbreite
- Kein Verlust für alle Pakete, die innerhalb der ausgehandelten Parameter liegen.

Dass Flussmodell mit einer Dienstrate R beschreibt den Dienst, der von einem spezifischen Pfad mit der Bandbreite R zwischen Sender und Empfänger bereitgestellt wird. Hierbei ist der Dienst, der einem bestimmten Fluss zugeordnet wird, unabhängig von dem Dienst, der anderen Flüssen zugeordnet wird. GS kontrolliert nicht die Untergrenzen von Gesamtverzögerung, Durchschnittsverzögerung oder Verzögerungsvarianz, sondern es werden stets nur die Obergrenzen dieser Parameter kontrolliert.

Controlled Load Service CL)

Ein Netzelement, das den CL-Service unterstützt, stellt einen Dienst zur Verfügung, der dem so genannten „Best Effort-Dienst“ in einem relativ schwach ausgelasteten Netz ähnelt. Nach [131] erwarten Anwendungen, die diesen Dienst beziehen, dass

- ein großer Anteil der übertragenen Pakete fehlerfrei beim Empfänger ankommt. Der genaue Anteil sollte sich hier an die Fehlerrate des Übertragungsmediums anlehnen und
- die Verzögerung, welche der überwiegende Teil der zugestellten Pakete erfährt, die minimale Verzögerung eines erfolgreich zugestellten Pakets nicht signifikant überschreitet.

Eine Anwendung, welche den CL-Dienst bezieht, kann die Ende-zu-Ende Verzögerung nicht beeinflussen. Die Anwendung hat jedoch die Möglichkeit, aus der Verzögerung, welche bei der Signalisierung festgestellt wurde, geeignet Rückschlüsse auf die aktuelle qualitative Situation im Netz zu ziehen.

2.1.2.2 RSVP

Beim IntServ-Modell definiert eine Anwendung einen erforderlichen Dienst auf Vermittlungsebene durch Auswahl der Dienstklasse samt dazugehörigen Parametern. Hierzu wird das RSVP-Signalisierungsprotokoll verwendet. Bei RSVP handelt es sich um ein Ende-zu-Ende Protokoll, das die Ressourcenreservierung entlang eines Datenpfades zwischen Sender und Empfänger signalisiert. Hier initiiert der Sender durch Vorschlagen der Reservierungsparameter den Reservierungsvorgang. Der Empfänger entscheidet dann autonom über den Vorschlag und legt die Parameter endgültig fest. Bei RSVP sind Reservierungen immer an eine befristete Zeitdauer gekoppelt und müssen periodisch aufgefrischt werden. Eine Reservierung wird entweder durch eine explizite Signalisierungsnachricht oder aber das Ausbleiben einer Auffrischungsnachricht beendet. Die RSVP-Signalisierung ist vom Datenverkehr unabhängig und transportiert keine Anwendungsdaten. Eine RSVP-Sitzung besteht aus folgendem 5-Tupel: IP-Zieladresse, Transport-Protokoll des Empfängers, Portnummer des Empfängers, IP-Quelladresse und Portnummer des Senders.

RSVP wird von Anwendungen verwendet, um eine spezifische Dienstgüte vom Netz zu erbitten. Der zentrale Bestandteil des RSVP ist eine Flussspezifikation, auch *flowspec* genannt. Diese beschreibt sowohl den Verkehrsstrom, der durch die Datenquelle erzeugt wird, als auch die Serviceanforderungen der Anwendung. Somit legen die Flußspezifikationen die gewünschte Dienstgüte fest. Obwohl RSVP-Reservierungsanträge beim Sender entstehen, wird die tatsächliche Steuerung der Dienstgüte beim Empfänger durchgeführt. Diese Anträge werden auf dem Rückpfad vom Empfänger zum Absender geführt. An jedem dazwischen liegenden Knoten werden folgende zwei Entscheidungen getroffen:

- Einrichten einer Reservierung
- Ablehnen der Reservierung

Ein Empfänger kann um eine Bestätigung für gewünschte Dienstgüte bitten. Diese wandert sozusagen durch das Netz und wird von den folgenden Knoten entweder bestätigt oder zurückgewiesen/modifiziert. Am Empfänger angelangt wird eine Bestätigung zurück zum Absender gesendet. RSVP arbeitet nach folgendem Schema: Jeder Sender sendet periodisch so genannte *PATH*-Nachrichten ins Netz. Diese werden von den Netzknoten entsprechend der Verkehrslenkungstabellen weitergeleitet und beinhalten Informationen über die Verkehrscharakteristik des Datenflusses, den der Sender erzeugen möchte. Zusätzlich sendet der Empfänger seine RSVP-Reservierungsanforderung, die so genannte *RESV*-Nachricht zurück zum Sender. Diese werden dann entgegen der Richtung der *PATH*-Nachrichten zurück zum Sender geschickt. Jeder dazwischen liegende Knoten speichert dann die relevanten Reservierungsinformationen. Sowohl *PATH*- als auch *RESV*-Nachricht haben als einen Parameter die Gültigkeitsdauer einer Reservierung, die von den auf dem Pfad zwischen Quelle und Senke liegenden Netzknoten verwendet wird, um die labilen Zustände (soft state) zu konfigurieren.

2.1.2.3 Das DiffServ-Modell der IETF

Das IntServ-Modell wurde sehr schnell in der akademischen Diskussion wegen Skalierungsproblemen kritisiert und es wurde unter dem Akronym DiffServ eine Arbeitsgruppe mit dem Namen Differentiated Services mit dem Ziel gegründet, ein alternatives QoS-Modell zu entwerfen [9], welches in [44] verfeinert wurde. Der grundlegende Ansatz bei diesem Modell ist ein skalierendes Weitverkehrsnetz welches eine begrenzte Anzahl von Vermittlungsdienste anbietet und eine vereinfachte Architektur der Netzknoten.

Die DiffServ-Architektur basiert auf einem einfachen Modell, bei dem der Verkehr beim Eintreten in einen Netzbereich klassifiziert und entsprechend der Klassifikation weiterverarbeitet wird. Alle Pakete mit gleicher Klassifizierungskennung, dem so genannten DS-Codepoint (DSCP), welche einen Netzabschnitt in gleicher Richtung durchlaufen, werden einer Verkehrsklasse zugeordnet und auf dem jeweiligen Netzabschnitt gleich behandelt. Das DSCP-Feld ist in [97] spezifiziert. Die absolute Anzahl an Klassen, die theoretisch einen übergeordneten Datenfluss beschreiben, ist theoretisch auf 64 begrenzt. Aktuell sind jedoch nur insgesamt 14 Klassen vorgesehen. Die übergeordneten Datenflüsse werden innerhalb eines Netzgebietes mit einem einheitlichen DSCP signalisiert. Die Verarbeitung, welche ein Paket in einem Netzknoten erfährt, wird PHB (per-hop behaviour) genannt. Momentan sind zwei verschiedene PHBs spezifiziert:

- Expedited Forwarding PHB [66],
- Assured Forwarding PHB [51]

Die Gültigkeit eines PHBs beschränkt sich auf einen administrativen Netzabschnitt, der in der Regel von einem Anbieter betrieben wird. Solch eine DS-Domäne besteht also aus einer Anzahl von Netzknoten, welche die identischen PHB-Mechanismen implementiert haben und

mit einem einheitlichen Betriebskonzept operieren. Somit ist sichergestellt, dass alle Netzknoten innerhalb einer DS-Domäne die Datenströme einheitlich bedienen können. Eine DS-Domäne hat fest definierte Domänengrenzen. Die Zugangsknoten zu einer Domäne sind für die Klassifizierung des in die Domäne eintretenden Verkehrs verantwortlich. Dieses Markieren erfolgt durch das Setzen des entsprechenden DSCPs. Im DiffServ-Modell ist keine über die Markierung der Pakete hinausgehende Signalisierung vorgesehen. Das Festlegen der Regeln, nach denen die Pakete klassifiziert und markiert werden, wird von einer Management-Instanz durchgeführt. Solche Klassifizierungsregeln basieren auf einem so genannten SLA (Service Level Agreement), welches zwischen Benutzer und Netzbetreiber festgelegt wird. Dieses SLA kann als semi-statisch betrachtet werden. D.h. ein SLA wird im Jahres- bzw. Monatszyklus, in der Regel durch einen neuen Vertrag zwischen Betreiber und Kunden, angepasst.

Expedited Forwarding PHB

Das Expedited Forwarding (EF) PHB ist für Verkehr vorgesehen, der eine niedrige Verlustrate, niedrige Latenz, niedrige Varianz und gesicherte Bandbreite Ende-zu-Ende über Domänengrenzen hinweg fordert. Aus Sicht der Endsysteme handelt es sich hierbei um eine exklusive Punkt-zu-Punkt Verbindung, die in der engl. Literatur oft auch *virtual leased line* bezeichnet wird. Es handelt sich bei diesem Dienst um einen Premiumdienst. Generell ist festzuhalten, dass dieser Dienst relativ statisch ist und sich konzeptionell an leitungsvermittelnde Architekturprinzipien anlehnt. Das EF PHB ist somit für Verkehr geeignet, für den prognostiziert werden kann, dass er eine gewisse Senderate nicht übersteigt. Hierfür garantiert das EF PHB eine vordefinierte Dienstgüte. Abschließend soll erwähnt werden, dass das EF PHB für Netzverkehr konzipiert wurde, welches zwar einfach, vorhersagbar und somit planbar ist, jedoch nicht wirklich der Verkehrscharakteristik von typischem Internet-Verkehr entspricht. Typischer Internetverkehr kann als weniger statisch und somit als nicht so einfach vorhersagbar charakterisiert werden,

Assured Forwarding PHB

Bei dem Assured Forwarding (AF) PHB wurden verschiedene Niveaus eines prioritätsbasierten Vermittlungsdienstes festgelegt. Hierbei wird das Weiterleiten der Pakete in einem Netzknoten klassenbasiert vorgenommen. Hierzu werden die in verschiedene Klassen eingeteilten Pakete mit Prioritäten versehen, welche relativ zueinander statisch sind. Es kann also einem Paketstrom keine absolute Garantie vergeben werden, sondern es wird eine prioritäts-basierte Weiterleitung der in einem Netzknoten relativ zueinander Pakete versichert. Diese Versicherung gilt zunächst nur für einen Netzknoten. Es ist jedoch vorgesehen, dass innerhalb eines administrativen Bereiches die Pakete einheitlich behandelt werden. Generell ist an jedem Netzknoten eine Überprüfung der Klassifizierung und ggf. eine Neuklassifizierung möglich.

Aktuell sind vier AF-Klassen mit je drei Verwerfungswahrscheinlichkeiten definiert, welche in [51] spezifiziert sind. Dies führt insgesamt zu 12 AF untergeordneten Klassen, welche jeweils einem Paket eine entsprechende Verwerfungswahrscheinlichkeit zuweisen. In Konsequenz wird ein Paket, das der AF-Klasse i und der Verwerfungswahrscheinlichkeit j entspricht, mit dem DSCP AF_{ij} markiert. Zusätzlich können zu lokal begrenzten Zwecken weitere AF-Klassen definiert werden, da der gesamte Adressierungsraum nicht ausgenutzt wurde. Die Dienstgüte der AF-Klasse ist somit dynamisch den jeweiligen Paketen zuordenbar. Generell ist die Gültigkeit der DSCPs auf den Bereich innerhalb einer DS-Domäne begrenzt.

Wenn Netzstauung auftritt, legt die Verwerfungswahrscheinlichkeit fest, welches Paket innerhalb einer AF-Klasse zuerst verworfen wird. Ein überlasteter Netzknoten versucht die Pakete vorzugsweise zu verwerfen, welche eine höhere Verwerfungswahrscheinlichkeit besitzen. Ein

Netzknoten kann jedoch Pakete einer AF-Klasse im Wartesystem nicht neu ordnen. Weiter sind keine quantitativen Anforderungen bezüglich Verzögerung oder Varianz spezifiziert.

Am Zugang zu einer Domäne wird der Verkehr, der in eine DS-Domäne eintritt, kontrolliert und bei Bedarf erfolgt, je nach Betriebskonzept des Betreibers, eine Verkehrsformung, Verzögerung oder Verwerfung von Paketen. Weiter können die Pakete bei Eintritt in eine neue Domäne neu klassifiziert werden.

Best Effort

Der Vollständigkeit halber wird hier noch die Best Effort-Klasse vorgestellt. Hierbei handelt es sich um eine Klasse, die keinerlei Mechanismen bzgl. der Dienstgüte vorsieht. Diese Klasse entspricht dem aus dem historischen Internet bekannten Verfahren, kein Paket entsprechend bevorzugt zu behandeln. Best Effort-Verkehr zeichnet sich dadurch aus, dass im Prinzip zwei verschiedene Klassen gebildet werden, welche sich durch die entsprechenden Transportprotokolle (TCP und UDP) auszeichnen. Implizit ist hier UDP mächtiger und dominiert gegenüber TCP wegen des nicht vorhandenen Rückkopplungsmechanismus. TCP-Verkehre untereinander bzw. UDP-Verkehre untereinander können jedoch bei Best Effort nicht dediziert behandelt werden.

2.1.3 Internet Transportprotokolle

Im Internet werden über der Vermittlungsschicht Transportprotokolle verwendet. Hier dominieren TCP und UDP. Weiter kommt als Kontrollmechanismus das ICMP-Protokoll zum Einsatz. ICMP spielt in dieser Arbeit jedoch keine Rolle und wird nicht weiter berücksichtigt. Da sowohl TCP als auch UDP nicht alle Anforderungen erfüllen, wurde jüngst das SCTP-Protokoll standardisiert, welches für den weiteren Verlauf dieser Arbeit von Bedeutung ist, und auf das somit in diesem Kapitel näher eingegangen wird.

2.1.3.1 TCP

TCP ist ein verbindungsorientiertes Protokoll, das durch Fehlersicherung einen verlässlichen Dienst bereitstellt [96]. Über einen Rückkopplungsmechanismus pendelt sich die Bandbreite immer auf die aktuelle Situation im Netz ein, wobei hier Laufzeiteffekte zu berücksichtigen sind [27]. Eine TCP-Sitzung besteht jeweils aus einem bidirektionalen, asymmetrischen Datenstrom zwischen den beiden Kommunikationsendpunkten. Bekannte Anwendungen, die über TCP transportiert werden, sind File Transfer (FTP), E-Mail (SMTP) oder auch Webzugriff (HTTP). Für weiterführende Informationen wird auf [96] verwiesen.

2.1.3.2 UDP

UDP stellt einen verbindungslosen Dienst zur Verfügung, der sich vor allem für den Transport von Echtzeitverkehr eignet. UDP verfügt über keinerlei Rückkopplungsmechanismen und ist somit nicht in der Lage, auf sich ändernde Netzverhältnisse zu reagieren. Somit besteht UDP stets aus einem unidirektionalen Datenstrom. UDP wird in der Regel für Echtzeit Audio/Videokommunikation eingesetzt (z.B. IP Telefonie). Für weitere Informationen wird auf [108] verwiesen.

2.1.3.3 SCTP

Das Stream Control Transport Protocol (SCTP) ist ein relativ neues Transportprotokoll und wurde in [123] spezifiziert. SCTP wurde von der IETF für den Transport von Signalisierungsnachrichten entwickelt. SCTP ist vielseitig einsetzbar und stellt eine im Vergleich zu TCP flexiblere Möglichkeit dar, Daten zwischen zwei Endsystemen auszutauschen. Diese Flexibilität wird durch die Trennung von Daten, welche Reihenfolgesicherung benötigen, von Daten ohne diese Anforderung, erreicht. Weiter wurde bei SCTP Fehlertoleranz, welche durch die

Verwendung von Redundanz auf Netzebene und durch die gleichzeitige Unterhaltung mehrerer paralleler Datenströme (Multihoming) erreicht wird, vorgesehen. SCTP ist für den Transport von administrativen Signalisierungsnachrichten wie z.B. Authentisierung, Autorisierung und auch für den Transport von Nutzungsdaten vorgesehen.

SCTP-Protokollübersicht

SCTP-Verbindungen, in der SCTP-Terminologie auch Assoziationen genannt, werden mittels eines 4 Wege-Handshake-Verfahrens zwischen zwei Endsystemen, in der Regel zwischen einem Client und einem Server, aufgebaut. Nachdem der Server die Assoziierungsanfrage vom Client bekommen hat, bestätigt er deren Erhalt. Diese Bestätigung enthält eine Datenstruktur, auch Cookie genannt, welche mittels einer sicheren Authentisierung geschützt ist. Nur wenn dieser Cookie vom Client unverändert zurückgesendet wird, allokiert der Server Ressourcen und die neue Verbindung ist aufgebaut. Somit ist sichergestellt, dass ein Server nicht „naiv“ auf beliebige Anfragen reagiert und Ressourcen reserviert, was das Protokoll vor sogenannten Denial-of-Service-Angriffen schützt.

Der SCTP-Assoziationsaufbau beinhaltet eine Liste gültiger Client/Server-Transportadressen (z.B. eine Kombination aus IP-Adressen und Portnummern). Die Anzahl aller möglichen Kombinationen der Server-IP-Adresse und der Server-Portnummer mit der Client-IP-Adresse und Client-Portnummer bilden den maximalen Parameterraum einer Assoziierung. Dieser Adressraum ist groß genug, um gleichzeitig mehrere Netzschnittstellen, also mehrere IP-Adressen, zu unterstützen.

SCTP ist ein verlässliches, nachrichtenorientiertes Transportprotokoll, welches es erlaubt, mehrere Kurznachrichten in ein SCTP-Paket packen zu können. In diesem Fall erfolgt keine Fragmentierung auf der Netzschicht. Dies wird durch einen MTU-Erkennungsmechanismus gewährleistet.

SCTP-Paketformat

SCTP-Pakete bestehen aus einem Header und einer bestimmten Anzahl von Nachrichtefeldern für entweder Datenpakete oder Signalisierungsnachrichten, welche den Rumpf (Payload) des SCTP-Paketes bilden (Abbildung 5).

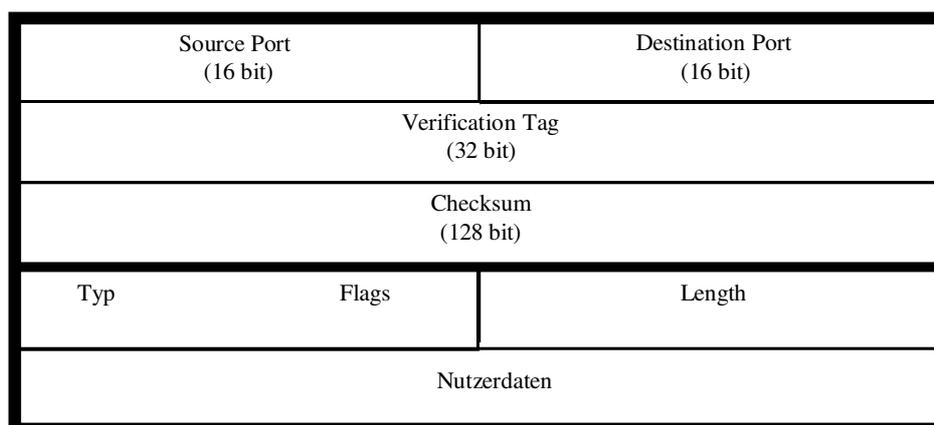


Abbildung 5: SCTP-Paket mit einem Nachrichtefeld

Ähnlich wie bei TCP und UDP enthalten die Header der Nachricht Quell- und Zieladressen auf Transportschicht, die sogenannten Portnummern. Weiter enthält der Header ein 32-bit langes Kontrollfeld und eine 32 bit Prüfsumme. Das Kontrollfeld ist eine Zufallszahl, welche mit der Kommunikationsgegenstelle in der Verbindungsaufbauphase ausgetauscht wird. Dies

schützt die einzelnen Assoziationen vor Angriffen, bei denen der Angreifer versucht, nicht zur Kommunikation gehörende SCTP-Pakete in die Assoziation einzuführen.

Nach Erhalt eines SCTP-Paketes wird zunächst die Prüfsumme verifiziert. Sollte sich diese als fehlerhaft erweisen, wird das Paket verworfen. Im Anschluss wird mittels IP-Adressen und Portnummern das Paket einer Assoziation zugewiesen. Danach wird das Kontrollfeld überprüft. Sollte sich dies als nicht übereinstimmend erweisen, wird das Paket ebenfalls verworfen.

Im Rumpf sind die verschiedenen Nachrichten enthalten, welche, wie bereits erwähnt, entweder Datennachrichten oder Kontrollnachrichten sein können. Kontrollnachrichten können sein: Bestätigungen, Überwachungsmeldungen, Start/Stop-Nachrichten für Assoziationen, Fehlermeldungen und ggf. Protokollerweiterungen.

SCTP-Nachrichten

Mit SCTP können Daten flexibel zwischen zwei Endpunkten ausgetauscht werden. Reihenfolgesicherung erfolgt durch den geeigneten Einsatz einer 32-bit langen Reihenfolgennummer (Transmission Sequence Number TSN), Empfangsbestätigungen und zeitgesteuerten Regelungen, die im Fehlerfall eine Neuübertragung der Nachricht veranlassen. Diese Funktionalität wird dadurch erbracht, in dem ein interner Multiplexmechanismus in die Protokollarchitektur implementiert wurde. Diese so genannten Streams werden mittels einer 16-bit langen Kennzahl identifiziert. Eine weitere 16 bit Kennzahl (ID) ist für die Reihenfolgesicherung vorgesehen. Sobald eine Assoziation aufgebaut wird, verhandeln die zwei Gegenstellen die Anzahl der eingehenden und ausgehenden Ströme, die während der Sitzung unterstützt werden. Prinzipiell ist es möglich, dass sich die Anzahl ausgehender Ströme von der Anzahl eingehender Ströme unterscheidet. In der Regel sind SCTP-Ströme unidirektional mit Reihenfolgesicherung definiert, jedoch kann der Benutzer explizit eine ungeordnete Übertragung der Nachrichten veranlassen.

2.1.4 Mobilität im Internet

Obwohl das Internet weltweiten Zugang zu den angeschlossenen Endsystemen bietet, ist es einem Benutzer nicht möglich, von dieser weltweiten Verfügbarkeit zu profitieren, solange er sich nicht an vordefinierten, fixen Punkten Zugang zum Internet verschafft. Durch die zunehmende Vielzahl der drahtlosen Endgeräte, die über entsprechende Schnittstellen zum Internet verfügen (Notebook, PDAs usw.), ist diese allzu statische Architektur der Internets nicht mehr zeitgemäß. Diese Entwicklung wurde von der traditionellen Telefonie in den letzten Jahren bereits vollzogen, so dass nun davon ausgegangen werden kann, dass das Internet nun denselben Weg bestreitet.

Mobilität im Internet erscheint in der aktuellen Diskussion in verschiedenen Ausprägungen. In diesem Zusammenhang kann man ausgehend von dem statischen Internet Nomadizität als quasi Mobilität erster Art bezeichnen. Hier sind die Benutzer eines mobilen Endgerätes prinzipiell mobil, bleiben aber während einer Sitzung, die theoretisch beliebig lange dauern kann, aus netztopologischer Sicht an einen festen Ort. Zu erwähnen ist hierbei, dass sich dies orthogonal zur Netztechnologie verhält, welche sowohl drahtlos als auch drahtgebunden Nomadizität unterstützen kann. Als weitere Stufe der Mobilität, also quasi Mobilität zweiter Art, kann nahtlose Mobilität bezeichnet werden. Diese kann sich in einem einfacheren Fall auf Mobilität innerhalb eines administrativen Netzbereiches beschränken, wie aktuell im 2G-Netz angeboten, oder aber auch nahtlose Mobilität zwischen administrativen Netzen. In diesem Fall bedeutet nahtlose Mobilität, dass bei Wechsel des fixen Aufpunkts, welcher die

Verbindung zum Internet herstellt, aktive Anwendungen unterbrechungsfrei weiter arbeiten, ohne dass der Benutzer davon Notiz nimmt [106].

Mittels Mobile IP [105] wurde von einer Arbeitsgruppe innerhalb der IETF das IP-Protokoll dahingehend erweitert, dass ein mobiler Benutzer einen flexiblen Zugangspunkt zum Internet unter Berücksichtigung allgemeiner Kompatibilitätsaspekte erhält. In den folgenden Abschnitten soll nun auf Mobile IP näher eingegangen werden. Mobile IP unterstützt die Mobilität erster Art, kann aber die Mobilität zweiter Art nur ansatzweise unterstützen.

2.1.4.1 Mobile IP-Basisfunktionen

Mobile IPv6 ermöglicht einem IPv6-Endsystem einen Grad an Mobilität, der sowohl nomadische Anwendungen unterstützt, als auch mit Einschränkung nahtlose Anwendungen. Ein IPv6-basiertes Endsystem kann wahlweise und ohne Vorankündigung den Ort innerhalb des IPv6-basierten Internets wechseln. Während eines solchen Netzwechsels/Ortswechsels, ist jedoch ein Paketverlust nicht vermeidbar, sollte dieser Wechsel während einer aktiven Sitzung vollzogen werden. Ein Ortswechsel wird in diesem Zusammenhang mit dem Wechsel des Netzzugangspunktes auf Vermittlungsschicht gleichgesetzt.

Generell wird innerhalb der globalen Internet-Architektur Mobilität durch Wechsel der IP-Adresse ermöglicht. Innerhalb von Mobile IPv6 gibt es nun Mechanismen, die den Wechsel einer IP-Adresse während einer aktiven Sitzung derart unterstützen, indem sie den Wechsel der IP-Adresse den verbindungsorientierten Protokollen der höheren Schichten verbergen.

Mobile IPv6 wurde so konstruiert, dass keine Eingriffe in die Protokollarchitektur nichtbeteiligter Rechner notwendig sind [71]. Es wird eine Verbindung mit Unterstützung einer statischen Adresse aufgebaut, die einem mobilen Endsystem (Mobile Node – MN) weltweit eindeutig zugewiesen wird. Jedes Endsystem ist also weltweit unter dieser Adresse erreichbar. Diese Adresse wird in Mobile IP als *Heimadresse* (Home Address Ha) bezeichnet. Eine Kommunikation erfolgt immer zwischen einem *MN* und einer Kommunikationsgegenstelle, die in der Literatur Correspondent Node (CN) genannt wird. Ein *CN* kann ebenfalls ein *MN* sein. Abbildung 6 zeigt die wesentlichen Komponenten des IPv6-basierten Mobilitätsmanagement-Konzeptes.

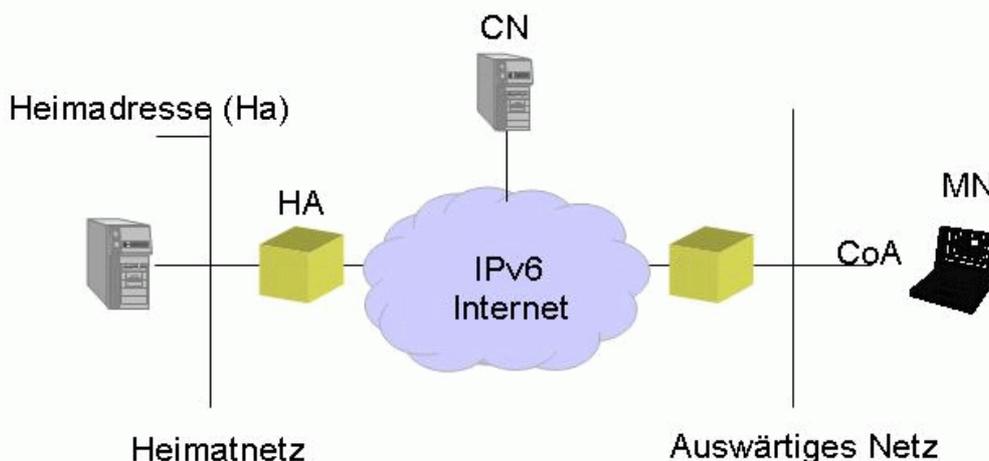


Abbildung 6: IPv6-Mobilitätsverwaltung

Generell wurde bei Mobile IP das Konzept des *Heimatnetzes* eingeführt und faktisch eine Trennung zwischen Endsystemadressierung und Verkehrslenkung eingeführt. Das *Heimatnetz* (Home Network) beschreibt den Teil des Netzes, dem der Heimat-Präfix zugeordnet ist. Die-

ses Präfix wird vom *MN* dazu verwendet, eine Heimat-IPv6-Adresse zu konfigurieren. Diese *Heimadresse* ist eine Adresse, die dem *MN* zugewiesen wird, unter der der *MN* jederzeit erreichbar ist. Pakete, die an den *MN* adressiert sind, werden zum *Heimatnetz* gesendet. Ist der *MN* im *Heimatnetz* anwesend, erfolgt kein Mobilitätsmanagement. Ist der *MN* jedoch nicht im *Heimatnetz* anwesend wird die Mobilitätsverwaltung angestoßen. Ist ein Benutzer in seinem *Heimatnetz* registriert, ist er für jeglichen *CN* mit seiner *Heimadresse* permanent erreichbar. Ist ein Benutzer an einem anderen Netz als dem *Heimatnetz* mit dem globalen Internet verbunden, hat er sein *Heimatnetz* verlassen und man führt für dieses Netz den Begriff *Auswärtiges Netz* ein. Bewegt sich ein mobiles Endsystem außerhalb seines *Heimatnetzes*, so ist ihm eine zweite IP-Adresse, die *Care-Off Adresse (CoA)* zugewiesen, unter der das Endsystem für die Dauer des Aufenthaltes in diesem fremden Netz ebenfalls erreichbar ist [125]. Diese *CoA* wird einem Knoten im *Heimatnetz*, der für die Mobilitätsverwaltung verantwortlich ist, dem so genannten *Heimagent (Home Agent HA)*, bekannt gemacht. Dieses Bekanntmachen impliziert faktisch die Abmeldung am *Heimatnetz* und ist prinzipiell immer an einen definierten Zeitraum gebunden. Ist ein mobiles Endsystem bei seinem *HA* mit einer *CoA* registriert, so fühlt sich der *Heimagent* für die Dauer dieser Abmeldung für die Annahme der für die an das mobile Endsystem adressierten Pakete stellvertretend verantwortlich. Jegliche, an den *MN* adressierten Pakete werden also stellvertretend vom *HA* angenommen und an die registrierte *CoA* - somit an den *MN* - weitergesendet. Dies erfordert implizit, dass die *CoA* weltweite Gültigkeit besitzt. Weiter ist es erforderlich, dass der *HA* im *Heimatnetz* des *MN* positioniert sein muss. In jedem Fall ist eine solche Registrierung in einem anderen Netz grundsätzlich an eine Zeitdauer gebunden, welche je nach Betreiber individuell konfigurierbar ist.

Um auf der Transportschicht das Ändern der IP-Adresse von der ursprünglichen *Heimadresse* zur jeweiligen *CoA* zu ermöglichen, wird in einem IP-Header auf Vermittlungsschicht ein Paket zwischen *CoA* und Adresse des *CN* übertragen, jedoch auf Transportebene zwischen *Ha* und der IP-Adresse des *CN*, wie in Abbildung 7 gezeigt. Dies geschieht bei der Übergabe der Pakete von Schicht 3 nach Schicht 4 durch Austausch der Schicht-3-Adresse bei Übergabe des Paketes von Schicht 3 nach Schicht 4. Der Austausch wird mittels Signalisierung unter Verwendung der optionalen Erweiterungen im IP-Header (siehe 2.1.1.1) gesteuert. Dieses Verfahren erfolgt analog auf dem Rückweg. Somit findet auf Schicht 4 eine Kommunikation stets zwischen der IP-Adresse des *CN* und einer festen Adresse im Heimatnetz (*Ha*) statt. Zur Verkehrslenkung und Wegefindung werden jedoch die Vorzüge des Internet-Protokolls voll genutzt, und es wird ein aufwändiges Tunneling vermieden.

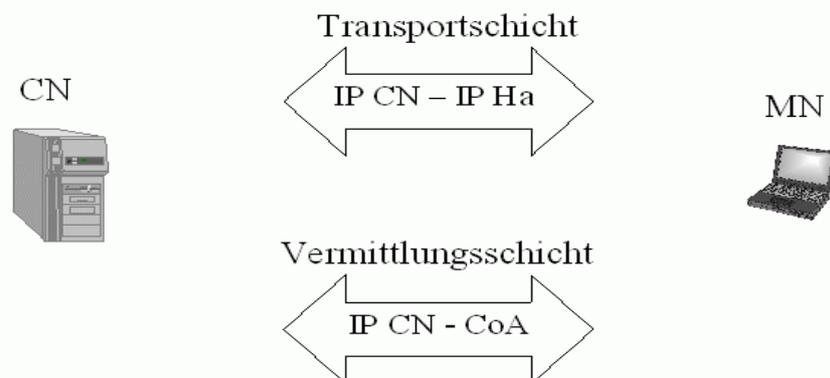


Abbildung 7: Mobile IP Adressierung auf Schicht 3 und Schicht 4

Das mobile Endsystem empfängt die Nachricht und erkennt an der originären Absenderadresse, dass es vom Heimagenten weitergeleitet wurde. Diese Information nutzt das Endsystem aus, um direkt dem *CN* antworten zu können. Dieser wiederum erkennt, dass die Antwort von der *CoA* kam und durch Ausnutzen dieser Information kommt es dann zu einer direkten Kommunikation zwischen den beiden Parteien. Ab diesem Zeitpunkt gibt es

prinzipiell keinen Unterschied mehr zum herkömmlichen IP-Protokoll. Dieses Szenario ist in Abbildung 8 nochmals kurz zusammengefasst dargestellt:

- 1.) Übergang von MN von Zugangsnetz 1 nach Zugangsnetz 2
- 2.) Registrierung des MN am HA (1)
- 3.) Ein CN sendet eine Nachricht an den MN, diese Nachricht wird vom HA aufgegriffen, weil sich der MN am HA mit neuer CoA registriert hat (2)
- 4.) Der HA leitet die Nachricht an die registrierte CoA weiter (3)
- 5.) MN antwortet direkt zu CN und CN direkt an MN (4)

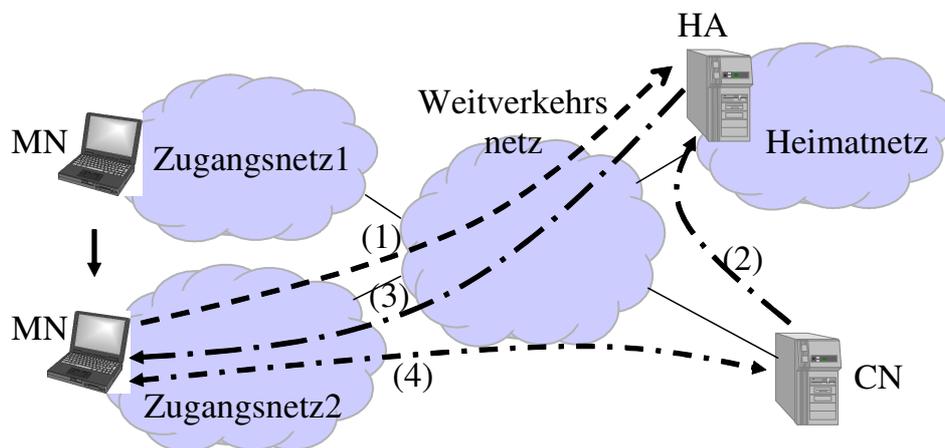


Abbildung 8: Mobile IP Basisoperation

2.1.4.2 IPv6-Nachrichtenformate und Protokollerweiterungen

Um die IP-basierte Mobilitätsverwaltung zu unterstützen, werden so genannte Zieladress-Option Header verwendet. Diese Erweiterungen erfolgen nach Kapitel 2.1.1.1 und erfüllen vier wesentliche Funktionen. Dies sind *Binding Update (BU)*, *Binding Update Acknowledgement (BUAck)*, *Binding Request (BR)* und die *Ha Option*. Für die Mobilitätsverwaltung sind neben diesen Mechanismen noch einige weitere wie *Router Discovery* und *Home Agent Discovery* erforderlich, die sich alle auf das ICMPv6-Protokoll stützen, für den weiteren Verlauf dieser Arbeit jedoch unbedeutend sind, und somit nicht näher betrachtet werden.

Binding Update Option

Die *BU Option* wird von einem mobilen IPv6-Knoten dann verwendet, wenn er sich außerhalb des Heimatnetzes befindet und eine neue *CoA* konfiguriert hat. In diesem Fall wird mit der *BU Option* der jeweilige *CN* über die neue *CoA* informiert.

Diese Option kann in jedes beliebige Datenpaket, welches zum *CN* adressiert ist und den *Destination Option Header* enthält, eingebunden werden. Ein *BU* wird prinzipiell zu zwei verschiedenen Zwecken eingesetzt. Entweder informiert ein *BU* den eigenen *HA* darüber, dass sich die *CoA* eines mobilen Endsystems geändert hat. Dies wird in der Literatur auch als *Home Registration BU* bezeichnet. Weiter kann ein *BU* auch dazu verwendet werden, um einen aktuellen *CN* während einer Sitzung darüber zu informieren, dass sich die *CoA* geändert hat. In diesem Fall verwendet der *CN* die *Ha Option* der Quelladresse eines ankommenden Pakets, um den eigenen Binding Cache-Eintrag zu verändern.

Binding Acknowledgement

Die *Binding Acknowledgement Option* bestätigt den Empfang einer *BU-Anforderung*. Auch diese Information kann innerhalb jedes Pakets, welches vom *CN* zurück zum *MN* adressiert ist, versendet werden. Wenn ein *Binding Acknowledgement* versendet wird, beinhaltet es zu-

sätzlich die Information, wie lange der *HA* diese *CoA* als Adressierungsendpunkt für ein mobiles Endsystem akzeptiert. Dies ist eine Zeitdauer, vor deren Ablauf sich ein mobiles Endgerät neu beim *HA* registrieren muss, soll der Heimagent die Registrierung der *CoA* aufrechterhalten.

Binding Update Request

Ein *Binding Update Request* wird verwendet, wenn das Binding eines *MN* geändert bzw. erneuert werden soll. Wenn ein *MN* eine solche Nachricht empfängt, antwortet er mit einem *BU*. Eine solche Anforderung wird immer dann versendet, wenn sich die gültige Verweildauer eines *BU* zum Ende neigt. Der *Binding Update Request* kann in jedes beliebige Datenpaket, welches zwischen den Kommunikationsgegenstellen ausgetauscht wird, eingebunden werden.

Home Address Option

Die *Home Address Option* verweist auf die *HA* des mobilen Endsystems und wird in jedes Paket eingebunden, das ein mobiles Endsystem versendet, solange sich das mobile Endsystem nicht im Heimatnetz befindet.

2.1.4.3 Mobile IPv6-Datenstruktur

Zur Unterstützung des IP-basierten Mobilitätsmanagements werden folgende Datenstrukturen verwendet: (Auch diese Aufzählung enthält nur die für den weiteren Verlauf der Arbeit relevanten Datenstrukturen)

Binding Cache

Der *Binding Cache* besteht im Wesentlichen aus einer Tabelle, die von jedem *CN/HA* verwaltet wird und die aktuellen Bindings der mobilen Endsysteme enthält. Abbildung 9 zeigt den wesentlichen Aufbau des *Binding Cache*.

HA des Mobiles Endsystems
CoA des mobilen Endsystems
Gültigkeitsdauer des Cache Eintrages
Der Zeitpunkt, zu dem der letzte BU erfolgte

Abbildung 9: Wesentliche Einträge des *Binding Cache*

Binding Update List

Die *Binding Update List* (Abbildung 10) wird vom *MN* verwaltet und hat die zentrale Aufgabe, die jüngst versendeten *BUs* zu speichern, welche zu den *CNs* und *HAs* versendet wurden.

Adresse des Knotens, zu dem BU gesendet wurde
HA des Binding Update
CoA, die beim letzten BU verwendet wurde
Gültigkeitsdauer des BU
Verbleibende Gültigkeitsdauer des Binding
Die absolute Zeit, bei der der letzte BU versendet wurde

Abbildung 10: Wesentliche Einträge der *Binding Update List*

2.1.4.4 Statische Mobile IPv6-Kommunikationsszenarien

Im Folgenden soll die prinzipielle Funktionsweise von Mobile IPv6 anhand von zwei Basis-szenarien vorgestellt werden. Diese Basisszenarien sind zum einen die Kommunikation zwischen einem Mobilten Endsystem (*MN*) und einer Mobilten Kommunikationsgegenstelle (*CN*) und die Kommunikation zwischen einem Mobilten Endsystem und dem eigenen Heimagenten.

Entdecken Care-off-Adresse

Der Prozess, welcher ein Endsystem entdeckt und lokalisiert, ähnelt dem im Internet Control Message Protocol (ICMP) beschriebenen Vorgehen [36]. Hier senden die Zugangsrouters periodisch „Lebenszeichen“ an die ihnen direkt zugeschalteten Netze. Konkret werden die Standard-ICMP-Nachrichten um mobilitätsspezifische Informationen erweitert. Diese Nachrichten werden in der Literatur *Router Solicitation Message* oder auch *Beacon* genannt.

Basierend auf dem Empfang dieser *Beacons* kann ein Endsystem erkennen, dass es sich im Einzugsbereich eines Zugangsrouters befindet. Da diese *Beacons* die Netzpräfix-Information enthalten, hat das Endsystem die Möglichkeit (z.B. unter Verwendung der eigenen Schicht 2-Adresse), sich selbst eine gültige IP-Adresse zuzuweisen. Dies erfolgt, wie in Abbildung 11 gezeigt, durch Aufspaltung der 48-bit langen MAC-Adresse in zwei Teile und dem Hinzufügen der FF:FE Sequenz und ist kompatibel zu dem Aufbau aus Abbildung 2.

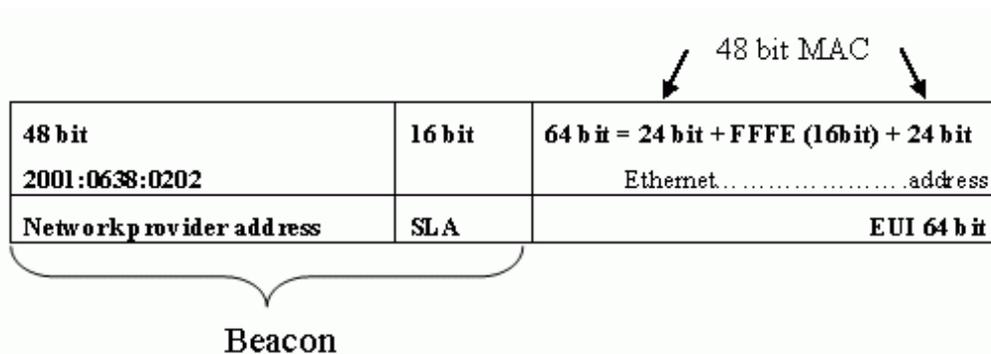


Abbildung 11: Autokonfiguration in Mobile IPv6

2.1.4.5 Sicherheit

Während der Spezifikation von Mobile IPv6 hat man sich entschlossen, Sicherheit nicht auf Anwendungsschicht zu belassen, sondern in das Protokoll selbst Möglichkeiten zur Authentisierung und Verschlüsselung einzubauen. Dies kommt vor allem der IPv6-Mobilitätsunterstützung zu Gute, da die erforderlichen Authentisierungsvorgänge bereits in das Schicht 3-Protokoll eingebunden sind. Jedoch muss erwähnt werden, dass sich jeglicher zusätzliche Mechanismus für Authentisierung negativ auf die Leistungsfähigkeit der Mobilitätsverwaltung auswirkt.

Zu erwähnen ist in diesem Zusammenhang, dass Mobile IP dafür vorgesehen wurde, mit anderen Sicherheitsmechanismen zusammenzuarbeiten. So blockieren in der Regel Firewalls Pakete, die in ein Netz eintreten und scheinbar aus diesem kommen. Dies ist vor allem der Fall, wenn sich ein mobiles Endsystem außerhalb des Heimatnetzes befindet und mit Netzknoten aus dem Heimatnetz kommunizieren möchte. Würde hier im Quelladressefeld die Heimatadresse verwendet werden, würde der Firewall diese Pakete blockieren. So wurde in [91] ein Protokoll vorgestellt, dass in einem solchen Fall die Pakete durch einen Firewall hindurchschleust.

2.1.4.6 Adressierung und Namensvergabe

Aus Mobile IP kann die Annahme abgeleitet werden, dass ein mobiles Endsystem permanent mit dem Heimatnetz verbunden ist. Dies ist eine Grundvoraussetzung für uneingeschränkte Erreichbarkeit eines Endsystems im Internet über eine IP-Adresse und mittels einer Namensverwaltung über den *Fully Qualified Domain Name (FQDN)*. Der *FQDN* kann sowohl statisch, als auch dynamisch mit einer oder mehreren IP-Adressen assoziiert werden. Im Falle dass mehrere IP-Adressen assoziiert sind, haben alle IP-Adressen die gleiche Bedeutung. Dies

kann z.B. für die Optimierung der Verkehrslenkung verwendet werden, um die Verkehrswege geeignet abzukürzen.

2.1.5 Sitzungskonzepte im Internet

2.1.5.1 WWW-basierte Sitzungen

WWW-basierte Sitzungen können als so genannte Inhalte-Dienste bezeichnet werden. Hier werden Inhalte angeboten, die z.B. elektronische Marktplätze oder Webbasierte Nachrichten darstellen können. Das Protokoll, welches den Transport der WWW-Sitzungen unterstützt, ist das Hyper-Text Transfer Protocol (HTTP), welches in [42] spezifiziert ist und zum Austausch von Hypertextdokumenten dient.

Das WWW kann als ein verteilter, hypertextbasierter Informationsdienst im Internet betrachtet werden, dessen Vorteile in einer einheitlichen graphischen Benutzeroberfläche und der Integration auch anderer Informationsdienste, wie z.B. Multimedia-Dokumenten liegen. HTTP setzt auf TCP auf und ist zustandslos.

Bei HTTP-basierten Sitzungen wird eine Sitzungs-ID verwendet um das Benutzerverhalten nachverfolgen zu können, um den Status eines Benutzers bei den verschiedenen HTTP-Ressourcen festzuhalten und in eingeschränkter Form eine Authentisierung durchzuführen [76]. Prinzipiell gibt es drei verschiedene Möglichkeiten Sitzungs-IDs zu verwalten:

- URL-Parameter
- Cookies
- HTTP State Management

URL steht für Uniform Resource Locator und kann als globale, weltweit gültige Adresse für Dokumente und andere WWW-Ressourcen betrachtet werden.

Weiter werden bei HTTP-Sitzungen Cookies eingesetzt. Hierunter versteht man einen Mechanismus, welcher von z.B. einem Web Server benutzt werden kann, um Informationen auf einem Klienten abzuspeichern, um zu einem späteren Zeitpunkt wieder darauf zugreifen zu können. Cookies sind jedoch nicht verlässlich, da der Benutzer die Verwendung von Cookies deaktivieren kann. Eine Sitzungs-ID, die zum Zweck der Authentisierung eingesetzt wird, ist in der Regel durch einen kurzlebigen Cookie repräsentiert, der nur für eine sehr begrenzte Zeit im Speicher des Browsers existiert. Der Cookie repräsentiert somit Zustände welcher ggf. nach Neustart eines Endsystems bestimmte Sitzungen aufrechterhalten kann.

In [76] ist ein Weg beschrieben, wie Zustände mit Hilfe von HTTP-Nachrichten erzeugt werden können. Der Mechanismus, der diese Zustände verwaltet, erlaubt Klienten und Servern das Austauschen von Zustandsinformationen. In [92] wird diese Methode kontrovers diskutiert und auf die Vorteile bzw. Nachteile dieses Mechanismus eingegangen. Der Vorteil liegt darin, dass eine Vielzahl von Anwendungen von der Möglichkeit profitieren, Zustände zwischen verschiedenen HTTP-Transaktionen zwischenspeichern zu können. Der Nachteil ist, dass hierbei die Privatsphäre des Benutzers elementar gestört wird, denn über diese lokal abgelegten Cookies kann nachvollzogen werden, auf welchen Seiten sich der Benutzer aufgehalten hat bzw. welche Ressourcen der Benutzer konsumiert hat. Ein koordiniertes Auswerten Informationen z.B. für Bereiche wie Customer Relation Management (CRM) und Marktforschung hochinteressant und wird oft ohne Wissen des Benutzers durchgeführt.

2.1.5.2 Real-Time Streaming Protokoll der IETF

Das Real-Time Streaming Protocol ist in [116] spezifiziert und kann als Gegenstück zu HTTP betrachtet werden, mit dem Unterschied, dass es sich hier um einen Echtzeitdienst handelt, der auf Transportebene auf UDP zurückgreift. Anders als bei der Übertragung von Webseiten ist es bei der Übermittlung von Audio und Videodaten nicht hinreichend, dass alle Datenpakete beim Empfänger ankommen, hier spielt die Rechtzeitigkeit des Eintreffens eines Pakets eine ebenso wichtige Rolle. Z. B. wird bei einem Videostrom ein nicht rechtzeitig eingetroffenes Paket einfach ausgelassen, was ggf. vom Anwender kaum bemerkbar ist. Weiter ermöglicht das Protokoll Start, Stop, Pause, Vorspulen und Zurückspulen und ähnliche Funktionen.

Innerhalb von RTSP wurde ebenfalls ein Sitzungskonzept eingeführt. Eine RTSP-Sitzung wird von einem Server mit einer *Setup*-Nachricht gestartet und mit einer *Teardown*-Nachricht beendet. Eine Sitzung ist durch eine Sitzungs-ID, welche im RTSP-Kopf definiert ist, eindeutig identifiziert. Diese ID wird vom Media Server festgelegt. Die Eindeutigkeit der Sitzungen erstreckt sich über mehrere Transport Sitzungen oder Verbindungen. Kontrollnachrichten für mehr als eine RTSP-URL können innerhalb einer einzigen URL-Session gesendet werden. Somit können Klienten dieselbe Sitzung verwenden, um mehrere zu einer übergeordneten Sitzung gehörende Datenströme zu kontrollieren, so lange alle Datenströme von einem Server kommen. Jedoch sind bei Mehrbenutzersitzungen einer einheitlichen URL, welche zu einem Klienten gesendet wird, mehrere Sitzungs-IDs notwendig, da über die Sitzungs-ID die vom gleichen Klienten kommenden Anfragen unterschieden werden müssen. RTSP benutzt für die Beschreibung der Sitzungen das Session Description Protocol SDP [48].

2.1.5.3 Session Initiation Protocol

Das *Session Initiation Protocol (SIP)* [49] ist ebenfalls ein Signalisierungsprotokoll auf Anwendungsebene und kann als aktueller Standard der IETF für Multimedia-Konferenzen über IP betrachtet werden. SIP wird verwendet, um Sitzungen zwischen einem oder mehreren Endpunkten zu verwalten. Somit stellt SIP die klassischen Signalisierungsfunktionen für ein paketbasiertes Telefonesystem bereit, bei dem mittels der Sitzungsverwaltung die Attribute einer Ende-zu-Ende Sitzung gesteuert werden können. SIP stellt folgende Kernfunktionen bereit:

- Bestimmung des Zielsystems – unterstützt durch Namensauflösung, Namensabbildung und Rufweitervermittlung.
- Bereitstellung der Eigenschaften des Endsystems. Hier wird mittels eines Sitzungsbeschreibungprotokolls (SDP) ein kleinster gemeinsamer Nenner für die Kommunikation zwischen zwei Endpunkten vereinbart.
- Anzeigen der Rufannahmefähigkeit der Kommunikationsgegenstelle.
- Etablierung einer Sitzung zwischen den beiden Kommunikationsendpunkten. In SIP besteht zusätzlich die Möglichkeit, das Ende-zu-Ende Paradigma aufzubrechen und dediziert so genannte Mittelboxen einzuführen, um z.B. je nach Bedarf die Kodierung eines Medianstromes den aktuellen Gegebenheiten anzupassen.
- Die Verwaltung der Übergabe einer Sitzung von einem Endpunkt zu einem anderen.

Generell kann eine SIP-Konferenz zwischen zwei oder mehreren Benutzern auf Basis von entweder Punkt-zu-Mehrpunkt (Multicast) Technologie, oder mehreren zeitgleichen Punkt-zu-Punkt (Unicast) Datenströmen stattfinden.

2.1.5.4 SIP-Komponenten

SIP ist ein so genanntes peer-to-peer Protokoll. Die Kommunikationsendpunkte einer Sitzung werden Benutzeragenten genannt. Ein Benutzeragent kann folgende Rollen einnehmen:

- Benutzeragenten im Endsystem (*User Agent Client - UAC*); eine Anwendung welche eine SIP-Anforderung initiiert.
- Server eines Benutzeragent (*User Agent Server - UAS*); eine serverseitige Anwendung, welche den Benutzer kontaktiert nachdem eine SIP-Anforderung empfangen wurde und in Vertretung des Benutzers diese beantwortet.

In der Regel kann ein jeder SIP-Endpunkt sowohl die Rolle des *UAC*, als auch die Rolle des *UAS* einnehmen. Die physikalischen Komponenten eines SIP-Netzes können in zwei Gruppen eingeteilt werden. In Abbildung 12 ist die generische SIP-Architektur einschließlich der Kernkomponenten schematisch dargestellt. In den folgenden Abschnitten wird auf die einzelnen SIP-Komponenten näher eingegangen.

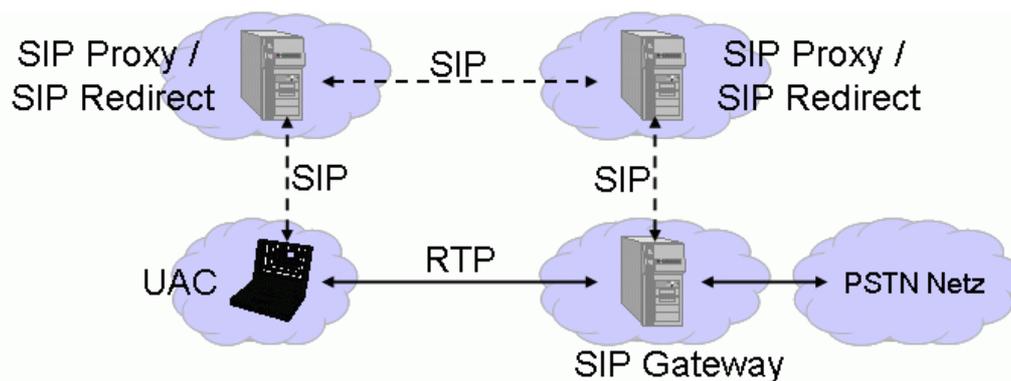


Abbildung 12: SIP-Architektur und Komponenten

2.1.5.5 SIP-Endsysteme

SIP-Endsysteme sind im Regelfall Telefone, die entweder als UAS oder als UAC agieren, oder aber so genannte Softphones, die z.B. die Funktionalität eines Telefons übernehmen. SIP Endsysteme sind in der Lage, SIP-Sitzungen zu initiieren, oder auf initiierte SIP-Anforderungen zu antworten.

SIP-Sitzungen werden in Richtung des öffentlichen Sprachtelefonnetz am SIP Gateway terminiert. SIP Gateways verwalten den Übergang zwischen SIP-fähigen Endsystemen und nicht-SIP-fähigen Endgeräten. Diese Verwaltung bezieht sich sowohl auf die Sitzungsverwaltung (Verbindungsaufbau, Verbindungsabbau usw.), als auch die Sitzung selbst. Hier kann z.B. eine Anpassung einer Audio oder Videokodierung erforderlich sein.

2.1.5.6 SIP-Server

Generell gibt es drei verschiedene Arten von SIP-Server. Dies sind ein Proxyserver, ein Redirect-Server und ein Registrierungsserver.

Der Proxyserver ist ein zwischen den Kommunikationssystemen liegender Server, welcher Anforderungen von einem Endsystem empfängt und diese dann im Namen des Endsystems weiterleitet (in der Regel zum nächsten SIP-Server). SIP-Proxyserver stellen zusätzlich die Funktionalität Authentisierung, Autorisierung, Netzzugangskontrolle Verkehrslenkung und Sicherheit zur Verfügung.

Der Redirect-Server informiert ein Endsystem über den nächsten Knoten, der die aktuelle SIP-Nachricht erhalten soll.

Der Registrierungsserver ist ein Server, der den aktuellen Aufenthalt eines UAC verwaltet.

2.1.5.7 Grundlegende Arbeitsweise von SIP

Endsysteme sind in SIP über eine eindeutige SIP-Adresse identifiziert. Diese Adresse ähnelt dem Format einer E-Mail Adresse (z.B. *sip:Benutzer-ID@rus.unistuttgart.de*). Die *Benutzer-ID* kann hier sowohl eine E.164 Adresse [65] sein, als auch nur ein Benutzernamen.

Ein Endsystem muss sich mit der SIP-Adresse beim Registrierungsserver melden, welcher diese Information auf Anforderung an andere SIP-Komponenten weiterleitet. Wenn nun ein Benutzer eine SIP-Sitzung initiiert, wird eine *SIP-Request* Nachricht zum SIP-Server gesendet. Diese Nachricht beinhaltet SIP-Ziel und Quelladresse und kann prinzipiell zum *SIP-Proxy* oder *SIP Redirect-Server* gesendet werden. Beide Möglichkeiten werden in den folgenden Abschnitten dieses Kapitels kurz beschrieben. Prinzipiell ist es möglich, dass ein Benutzer zwischen verschiedenen Endsystemen wechselt. Ein Benutzer kann somit seinen Netztopologischen Standort dynamisch beim SIP-Registrierungsserver erneuern.

Bei einer Einbindung eines Proxyserver beginnt der *UA* mittels einer *Invite*-Nachricht, welche zum SIP-Proxy gesendet wird, den Sitzungsaufbau. Der Proxy-Server hat nun die Aufgabe den Weg zum Empfänger zu suchen und die Nachricht entsprechend weiterzuleiten. Der Empfänger bestätigt diese Anforderung und sendet sie zurück an den Proxy-Server, welcher die Antwort zurück an den Initiator sendet. Im Anschluss daran schickt der Proxyserver an beide Parteien eine abschließende Bestätigung, welche die Sitzungsaufbauphase beendet und eine direkte RTP-Kommunikation zwischen den Parteien ermöglicht.

Bei der Verwendung eines Redirect-Servers sendet der Anrufer eine *Invite*-Nachricht zum Redirect-Server, welcher dem Anrufer den aktuellen Aufenthaltsort des Ziels zurückgibt. Mit dieser Information ist es dem Anrufer nun möglich direkt mit dem Zielsystem Kontakt aufzunehmen und die Kommunikation aufzubauen.

2.2 Das Mobilkommunikationsnetz

Das Kernkonzept der Mobilkommunikation ist die Netzzelle. Dies bedeutet, dass regionale Netze in verschiedene Netzzellen eingeteilt sind. Jede Netzzelle deckt dann ein begrenztes Gebiet ab. In jeder dieser Zellen ist eine stationäre Empfangsstation (*Base Transceiver Station BTS*) platziert, welche die Kommunikation zwischen einem Endsystem (*Mobile Terminal, MT*), das sich im Einzugsbereich einer Netzzelle befindet, und der entsprechenden Kommunikationsgegenstelle geeignet unterstützt. Die Größe einer solchen Netzzelle hängt von mehreren Faktoren ab. Dies sind unter anderem die Teilnehmerdichte, die Höhe der Antenne, die Sendeleistung der Antenne, die Sendeleistung der Endgeräte und die Topologie des Einzugsbereiches der Netzzelle. Die grundlegende Netzarchitektur ist in Abbildung 13 dargestellt.

Mehrere *BTS* sind an einen *BSC (Base Station Controller)* angeschlossen. Dieser ist für die logische Steuerung jeder einzelnen stationären Empfangsstation verantwortlich. Hauptaufgabe des *BSC* ist die Steuerung und Koordination von Zellübergängen. Der *BSC* ist mit einer Mobilvermittlungsstelle (*MSC - Mobile Switching Center*) verbunden. Die *MSC* ist für die Signalisierung der Verbindungen zu und von einem mobilen Endgerät verantwortlich. Somit sind in der *MSC* viele der Funktionen implementiert, die auch in einem entsprechenden Netzknoten der Festnetztelefonie ausgefüllt werden.

Eine weitere Kernkomponente der allgemeinen Mobilkommunikationsarchitektur ist das Heimatregister (*Home Location Register - HLR*). Das *HLR* ist ein dynamisches Verzeichnis, welches die benutzerspezifischen Informationen der bei einem Netzbetreiber unter Vertrag stehenden Endkunden enthält. Neben der Beherbergung dieser benutzerspezifischen Daten ist

das *HLR* auch in die Mobilitätsverwaltung eingebunden. Dies erfolgt nach folgendem Schema:

Wenn ein Benutzer mit seinem *MT* logisch von einer *MSC* zu einer anderen *MSC* wechselt, so muss die *MSC* diesen neuen Aufenthaltsort des Endsystems dem *HLR* melden. Sollte dann ein ankommendes Gespräch aus dem Kernnetz eintreffen, fragt die *MSC* beim *HLR* nach dem letzten gemeldeten Aufenthaltsort des Teilnehmers, damit das Gespräch weitergeleitet werden kann.

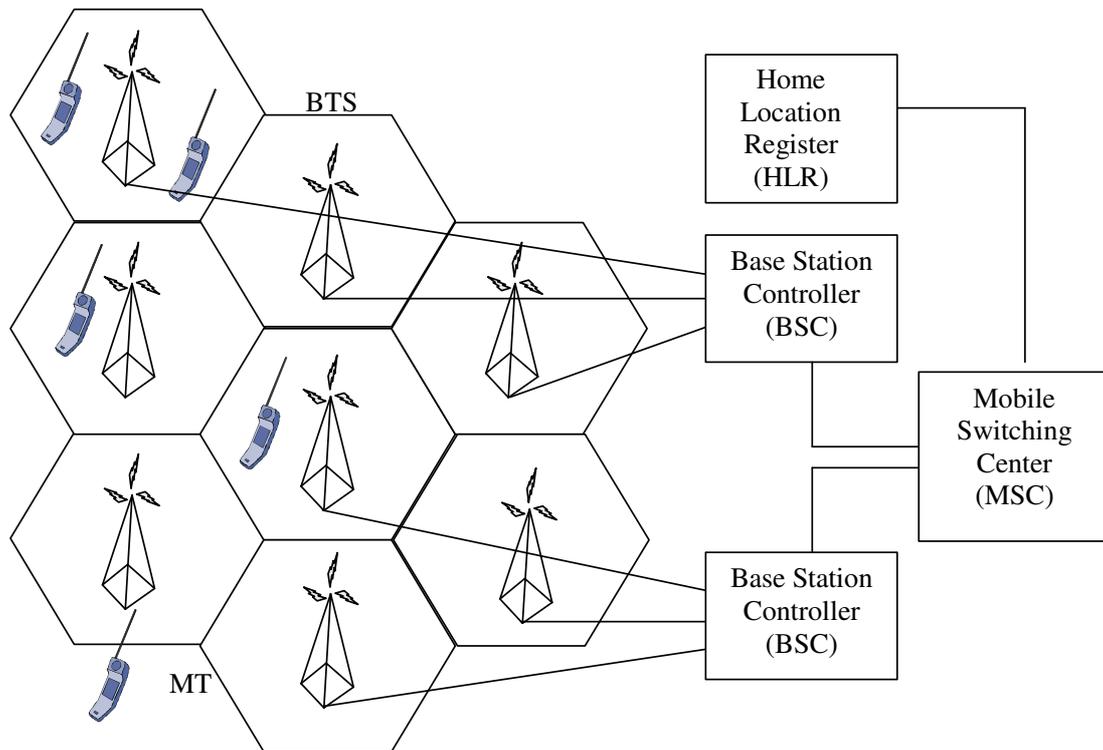


Abbildung 13: Grundlegende Netzarchitektur der Mobilkommunikationsnetze

Eine grundlegende Eigenschaft der Luftschnittstelle sind die sehr begrenzt zur Verfügung stehenden Ressourcen. Im Gegensatz zu anderen Übertragungsmedien, wie z.B.: eine Kupferleitung oder eine Glasfaser, ist es beim Übertragungsmedium Luft nicht möglich, ein weiteres Spektrum vorbehaltlos bei zusätzlichem Bedarf hinzuzufügen. Da das zur Verfügung stehende Spektrum sehr begrenzt ist, ist es bei der drahtlosen Mobilkommunikation umso wichtiger, mit den verfügbaren Ressourcen sehr effektiv umzugehen. Dies führte zu der Standardisierung verschiedener Verfahren, welche alle das Ziel haben, das verfügbare Spektrum geeignet für die Kommunikation zu verwalten.

2.2.1 Die Evolution der Mobiltelefonie

Das Mobiltelefonnetz folgt einer aus dem Telefonnetz abgeleiteten Netzarchitektur und ist somit ein weiterer Vertreter der klassischen verbindungsorientierten Betriebsart. Die wesentlichen Meilensteine der Evolution der Mobiltelefonie sind in Abbildung 14 dargestellt.

Das Mobiltelefonnetz wurde ursprünglich in analoger Technologie konzipiert und betrieben. Eine nennenswerte Penetration wurde in Deutschland jedoch erst mit der zweiten Generation der Netzarchitektur (2G) erreicht und wird nun sukzessive durch die dritte Generation (3G – UMTS) abgelöst. Die 2G-Netzarchitektur ist auch unter dem Akronym *GSM* (*Global System*

for Mobile Communication) bekannt. In den folgenden Abschnitten soll auf ausgewählte Architekturen in dieser Evolution näher eingegangen werden.

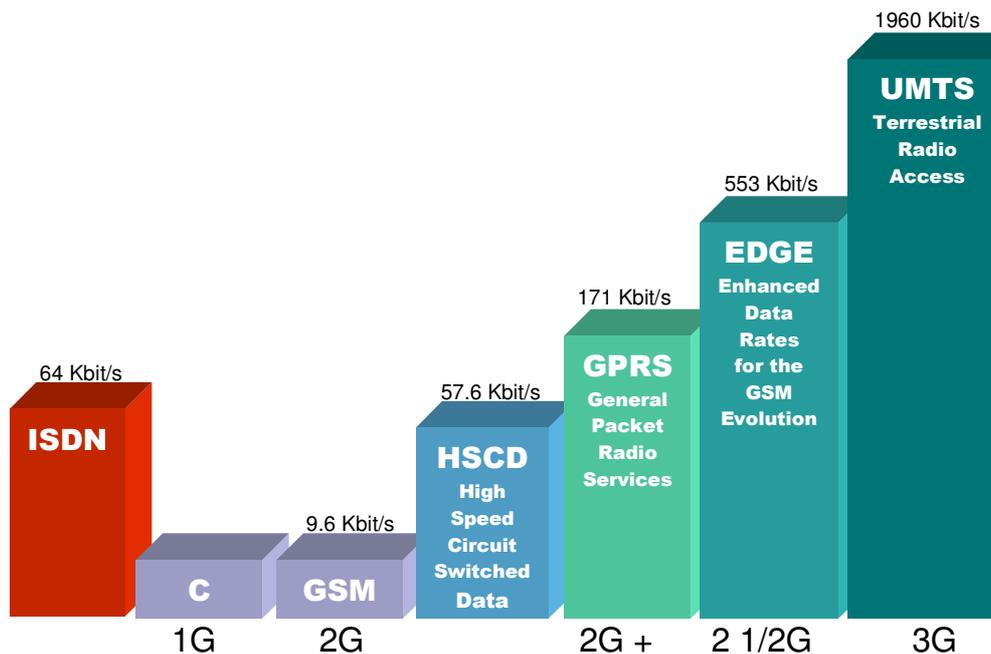


Abbildung 14: Evolution der Mobilfunktechnologie

2.2.1.1 GSM

Das *GSM*-Netz wird im 900 MHz und 1800 MHz (1900 MHz in den USA) Frequenzband betrieben und ist ein digitales, zellulares System, das in erster Linie sprachbasierte Dienste bei netzübergreifender Erreichbarkeit und automatischer Übergabe von Verbindungen bei Zellwechsel ermöglicht.

GSM bietet insgesamt vier Dienste. Neben der Sprachübertragung gibt es die Dienste Kurznachrichten, Fax und Datenkommunikation und eine Reihe von ISDN-ähnlichen Dienstmerkmalen. *GSM* stellt einen verbindungsorientierten Dienst bereit, für den unter Verwendung des *GMSK* Zeitmultiplex-Modulationsverfahrens (*Gaussian Minimum Shift Keying*) an der Luftschnittstelle, zunächst für Sprache 13 kbit/s und für Daten 9,6 kbit/s spezifiziert sind. Diese Geschwindigkeit reicht lediglich für Anwendungen wie E-Mail und elementare Sprachkommunikation aus, ist aber für synchrone Anwendungen mit größerem Datenaufkommen mangels limitierten Kommunikationskanälen nicht ausreichend. Obwohl mit Hilfe verschiedener Verfahren höhere Datenraten erreicht werden können, hat *GSM* große Nachteile, da es Verbindungen mittels Kanalreservierungen erstellt, die sich nur bedingt für die asynchrone Datenübertragung eignen. Aktuell planen die Netzbetreiber in Deutschland den Betrieb des *GSM*-Netzes in Deutschland im Jahre 2009 einzustellen.

Abbildung 15 zeigt einen Überblick über wesentliche Komponenten der *GSM*-Netzarchitektur. Diese sind neben dem Endsystem, das aus dem Endgerät und einem *SIM* (*Subscriber Identity Module*) besteht, die *BTS*, der *BSC*, das Mobile Switching Center (*MSC*), das Home Location Register (*HLR*) das *Gateway MSC* (*GMSC*) und das *Equipment Identity Register* (*EIR*).

Das *SIM* ist als Chipkarte implementiert, welche benutzerspezifische Informationen, die Identität des Vertragsnehmers, Informationen, welche für die Benutzerauthentisierung erforderlich sind, und ggf. spezifische Dienstinformationen enthält. Durch das Einfügen einer *SIM*-

Karte in ein Endgerät werden die mit dieser *SIM*-Karte in Zusammenhang gebrachten Eigenschaften, die ein Abbild des Vertrages zwischen dem Inhaber der *SIM*-Karte und dem Netzbetreiber darstellen, auf das Endgerät übertragen.

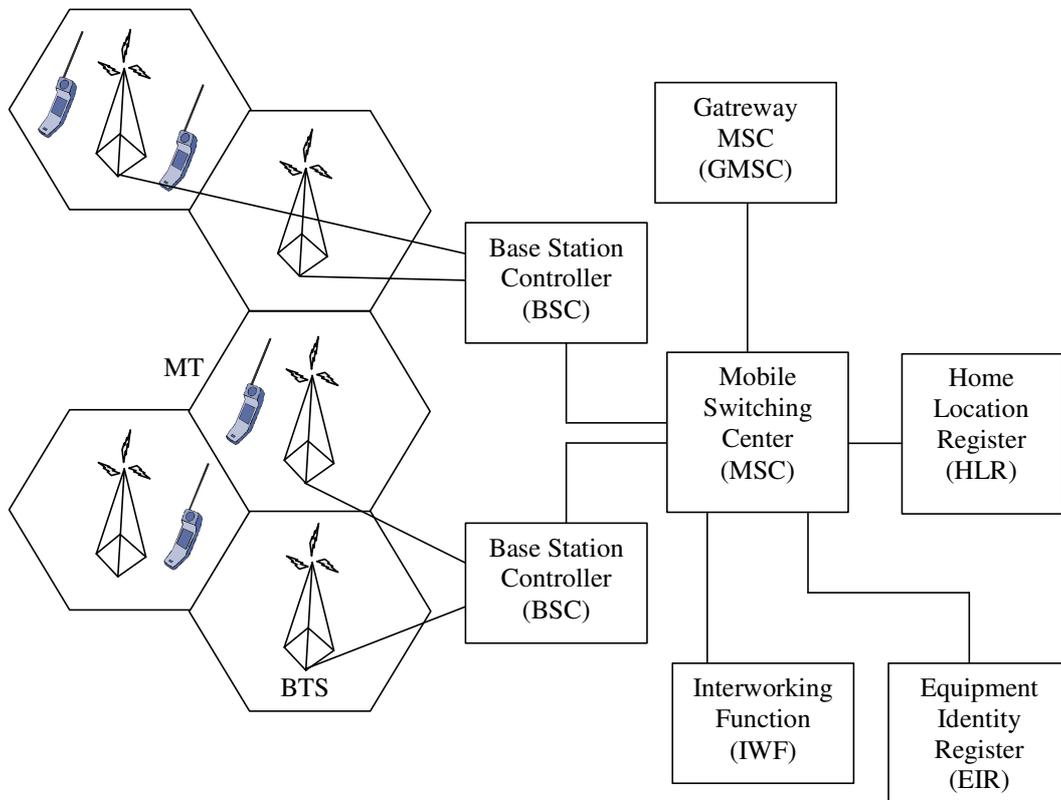


Abbildung 15: Komponenten der GSM-Architektur

Der *BTS* stellt einen physikalischen Übergang zwischen dem drahtlosen und dem drahtgebundenen Abschnitt der Kommunikationsarchitektur her. Mehrere *BTS* werden mit einem *BSC* verbunden, der für Ressourcenverwaltung, lokale Mobilitätsverwaltung und einige Betriebsaufgaben verantwortlich ist. Die Schnittstelle zwischen *BSC* und *MSC* wird in der Literatur auch als A-Schnittstelle bezeichnet und basiert auf SS7 [120], einem Signalisierungsprotokoll aus der Sprachkommunikation im Festnetzbereich.

Eine wesentliche Funktion innerhalb des *MSC* ist das für das Mobilitätsmanagement verantwortliche *VLR* (*Visiting Location Register*). Dies ist eine Datenbank, die die benutzerspezifischen Informationen der sich im Einzugsbereich des *MSC* befindlichen Benutzer enthält.

Das Gegenstück zum *VLR* ist das *HLR*. Dies ist ebenfalls eine Datenbank, jedoch werden hier die aktuellen Aufenthaltsorte der sich „auf Reise“ befindlichen Benutzer verwaltet und mit den jeweils anderen *VLR* ausgetauscht. Das *HLR* ist noch mit einer Instanz, die für die Authentisierung verantwortlich ist, verbunden.

Das *EIR* registriert eine eindeutige Kennung der Endsysteme, die *IMEI* (*International Mobile Equipment Identity*) der Mobilstationen. Somit kann dann dediziert ein Endsystem vom Netz genommen werden.

2.2.1.2 HSCD

HSCD (*High Speed Circuit Switched Data*) ist ein leitungsvermittelndes Verfahren, welches auf der *GSM*-Architektur aufbaut. Durch die gleichzeitige Nutzung von bis zu sechs Zeit-

schlitzen, ist es einem einzelnen Benutzer möglich, Transferraten von bis zu 57.6 kbit/s in Anspruch zu nehmen. Ansonsten besitzt dieses Protokoll keinerlei weitere Besonderheiten, welche im Rahmen dieser Arbeit von Bedeutung sind. Bei *HSCD* können wie bei *GSM* nicht benutzte Ressourcen nicht weitergegeben werden. *HSCD* hat also einen rein statischen Reservierungsmechanismus implementiert, welcher keine weitere Unterstützung für IP-basierten Datenverkehr bietet. Theoretisch können mit dieser Technologie einem Benutzer mehr als 57,6 kbit/s zur Verfügung gestellt werden, jedoch ist die Schnittstelle zwischen *MSC* und *BSC* (A-Schnittstelle), auf 64 kbit/s pro Kanal begrenzt. Für weiterführende Informationen wird auf [120] verwiesen.

2.2.1.3 GPRS

GPRS (*General Packet Radio Service*) ist eine weitere auf *GSM* aufbauende Architektur, welche wesentliche Neuerungen bietet und in der Fachliteratur häufig als 2G+ bzw. 2,5G bezeichnet wird. *GPRS* ermöglicht dem mobilen Benutzer sowohl schnellen Netzzugang, als auch eine effektivere Nutzung der Ressourcen, da hier erstmals im Ansatz Paketvermittlungsverfahren eingesetzt werden. Konkret werden mit Hilfe eines Multiplexverfahrens Pakete einer Sitzung auf verschiedenen Kanälen übertragen. *GPRS* erlaubt, wie in Abbildung 14 dargestellt, eine maximale theoretische Übertragungsrate von bis zu 171 kbit/s und arbeitet ähnlich wie *HSCD* mit Kanalbündelung. Der große Vorteil von *GPRS* ist allerdings, dass die Kanäle somit Ressourcen nur dann belegt werden, wenn Daten auch wirklich gesendet oder empfangen werden. Somit verbraucht ein zeitweise inaktiver Teilnehmer keine Ressourcen und konkurriert während dieser inaktiven Zeit nicht mit aktiven Benutzern einer Netzzelle um Bandbreite. Dieser Vorteil kann als wesentliche Voraussetzung betrachtet werden, um das so genannte „Always-On“ Paradigma des Internets zu unterstützen, auf das in Kapitel 5.3.1 noch näher eingegangen wird.

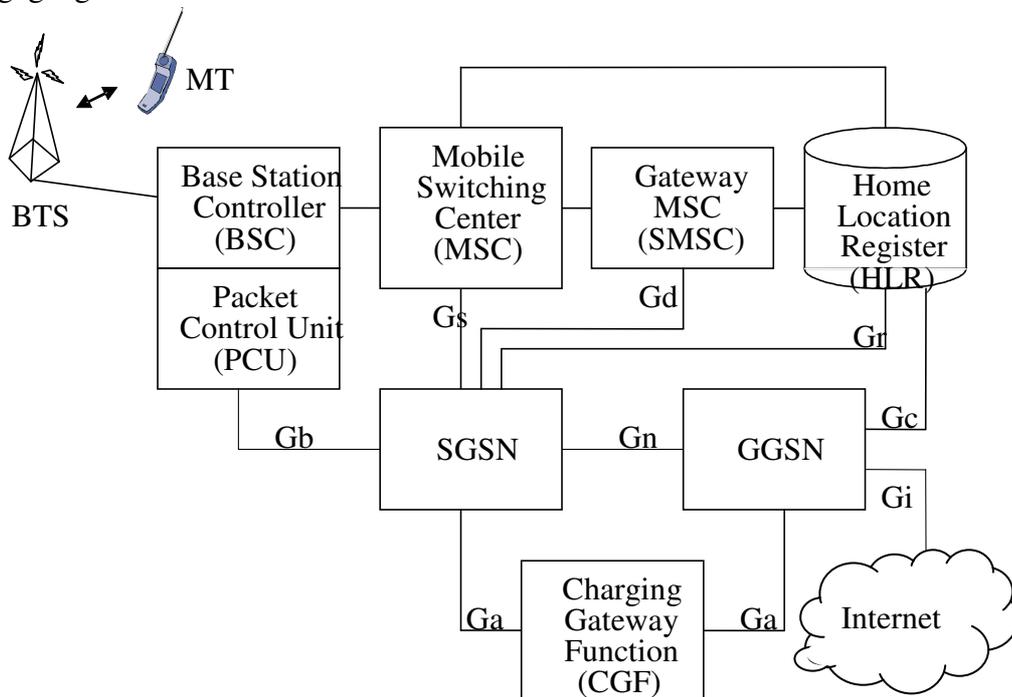


Abbildung 16: Komponenten der GPRS-Netzarchitektur

Hier ist der Benutzer permanent mit dem Netz verbunden ("online"), braucht keine Zeit mit dem Einwählen und Registrieren am Netz aufzuwenden, ist also ständig erreichbar. Dieser so genannte natürliche Multiplex-Gewinn eines paketvermittelnden Netzes führt zu einem sehr viel effektiveren Gebrauch der verfügbaren Ressourcen im Vergleich zu z.B. bei *HSCD* [120].

Abbildung 16 zeigt eine leicht vereinfachte GPRS-Netzarchitektur mit den wesentlichen Schnittstellen. Es erfolgt hier eine strikte Trennung zwischen Datenpfad und Signalisierungspfad. Die Benutzerdaten wandern entlang der Schnittstellen „Gb“, „Gn“ und „Gi“ ins Internet. Alle weiteren Schnittstellen aus Abbildung 16 sind zur Signalisierung vorgesehen.

2.2.1.4 EDGE

Enhanced Data Rates for Global Evolution (EDGE) ist eine auf *GPRS* beruhende Architektur, welche dem Benutzer eine max. Datenrate von 384 kbit/s ermöglicht. *EDGE* bringt im Wesentlichen zwei Neuerungen mit sich. Dies sind das 8-PSK-Modulationsverfahren (*Phase Shift Keying*), sowie einen adaptiven Wechsel zwischen verschiedenen Modulationsarten. Das 8-PSK Verfahren erreicht gegenüber dem bisherigen Verfahren die dreifache Datenrate pro Übertragungsschritt. Der Vorteil von *EDGE* ist, dass kein neues Frequenzband für die Kommunikation erforderlich ist, sondern das vorhandene GSM-Band effektiver ausgenutzt wird. Das *EDGE*-Prinzip lässt sich für die zukünftige *GSM/GPRS*-Netzevolution ausgehend von *HCS*D und *GPRS* einsetzen. Im Falle von *HCS*D spricht man dann von einem *ECSD*-Netz (*Enhanced Circuit Switched Data*) und im Falle von *GPRS* spricht man von einem *EGPRS*-Netz (*Enhanced General Packet Radio Service*). Da nahezu alle GSM-Netzbetreiber *GPRS* in ihren Netzen integriert haben, werden vor allem *EGPRS*-Dienste in Zukunft von größerem Interesse sein. Für weitere Informationen bezüglich *EDGE* sei auf [120] verwiesen.

2.2.2 UMTS

Als Nachfolge von *GSM* wurde bei ETSI [40] das *UMTS (Universal Mobile Telecommunication System)* standardisiert [129]. Entsprechende weltweite Aktivitäten werden unter dem Überbegriff *IMT-2000* zusammengefasst [120]. Als das Mobilkommunikationssystem der dritten Generation (3G) soll *UMTS* vor allem höhere Bandbreiten für multimediale Telekommunikations und Datendienste bereitstellen. Hier kommt das *WCDMA (Wideband-Code Division Multiple Access)* Multiplexverfahren zum Einsatz, mit dem sich in einem relativ schmalen Frequenzband mit Hilfe eines Codemultiplexverfahren eine signifikant erhöhte Übertragungsrate erzielen lässt. Für das Zugangsnetz wird auf der Luftschnittstelle der Bereich um 2 GHz verwendet. *UMTS* stellt nach aktuellem Stand bei großer geographischer Abdeckung bis zu 384 kbit/s pro Endgerät zur Verfügung. Die maximale Übertragungsbandbreite einer Sendezelle ist auf ca. 2 Mbit/s begrenzt. Diese Bandbreite ist theoretisch für einen einzelnen Benutzer verfügbar, sollte sich dieser alleine im Bereich einer Sendezelle befinden. Ein sich an *UMTS* anlehrender Standard wird unter dem Akronym *CDMA 2000* entwickelt. Hierbei handelt es sich um die nordamerikanische Form des *UMTS*-Standards, der auf einem anderen Frequenzband arbeitet.

Für die exklusive Nutzung eines Frequenzbandes sind Lizenzen erforderlich, deren Vergabe in der Regel unter staatlicher Hoheit liegt. Abbildung 17 zeigt die relevante Frequenzaufteilung von *UMTS* und *IMT-2000* in Europa, den USA und Japan.

In seiner ersten Version ist die *UMTS*-Architektur noch sehr stark mit klassischen, historisch aus der Sprachtelefonie bekannten Konzepten versehen. Es gibt aber in Europa schon Anstrengungen, diese Technologie in ihrer nächsten Ausbaustufe aus konzeptioneller Sicht weiter mit dem Internet zu migrieren, wie es z.B. in dem von der Europäischen Union geförderten Projekt *Moby Dick* [39], [50] angestrebt wurde. Auf diese Architektur wird in Kapitel 4 noch näher eingegangen.

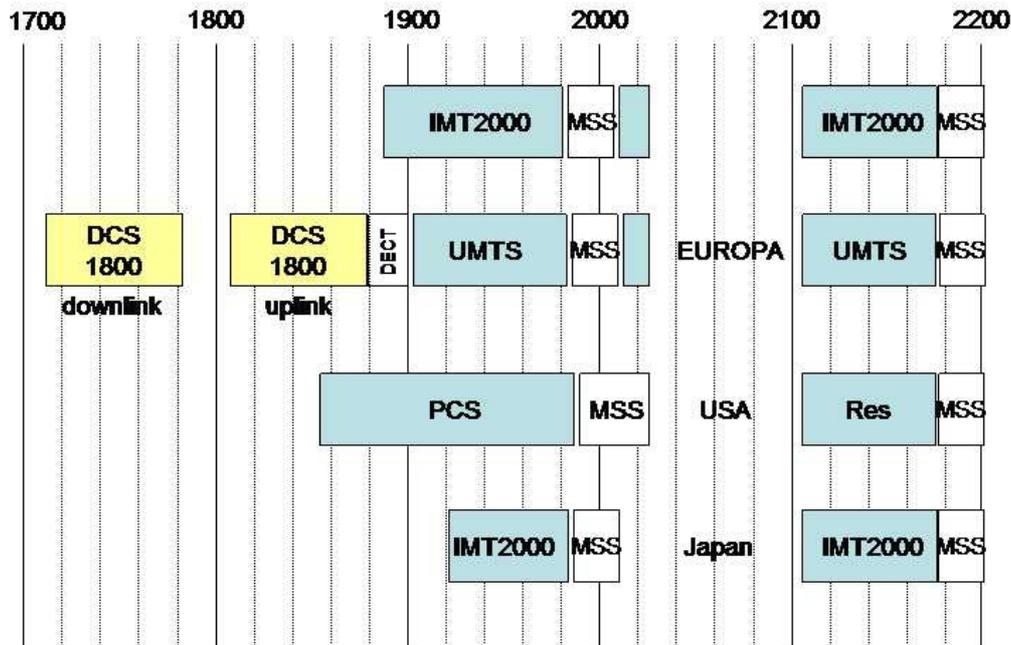


Abbildung 17: Aufteilung der Frequenzbänder in IMT-2000 (Quelle: Bundesnetzagentur]

2.2.2.1 Basiskonzepte von UMTS

Die UMTS-Technologie wird es ermöglichen, Echtzeit-Audio und Videodaten von und zum mobilen Teilnehmer zu übertragen. Es wird somit das erste System sein, das mobile Sprach und Datendienste auf einer einheitlichen Plattform vereinheitlicht.

2.2.2.2 UTRAN

Der Funknetzteil von *UMTS* (*UTRAN – UMTS Terrestrial Radio Access Network*) [1] kann als Weiterentwicklung des Funknetzteils von *GSM* betrachtet werden, bei dem zusätzliche Funktionalitäten unterstützt werden, welche für *GSM* nicht notwendig sind. So verwendet *GSM* ein *TDMA*-Verfahren, während *UMTS* ein *CDMA*-Verfahren benutzt. Abbildung 18 zeigt die Komponenten des *UTRAN*, die primär aus zwei Baugruppen bestehen. Dies sind die Basisstation (Base Station), welche in der Literatur *Node B* genannt wird, und der *Radio Network Controller (RNC)*. Die *Node B* ist für die Verwaltung der Luftschnittstelle verantwortlich und kann mehrere Zellen separat unterstützen. Der *RNC* ist für die Mobilitätsverwaltung einer bestimmten Anzahl von *Node B*'s verantwortlich. Die *RNC*'s sind untereinander über die Iur-Schnittstelle verbunden und können somit direkt miteinander kommunizieren. Das *UTRAN* ist dadurch in der Lage, Netzübergänge eigenständig durchzuführen. Dies ist gegenüber dem *GSM*-Funknetz eine wesentliche Neuerung, da hier bei einem Zellwechsel zwischen verschiedenen *RNC*-Bereichen die Mobilitätsverwaltung über eine Vermittlungsstelle im Kernnetz abgewickelt werden kann.

Bei Zellenübergängen senden die mobilen Endgeräte mit sehr hoher Leistung, da die maximale Entfernung zur *Node B* überbrückt werden muss. Diese erhöhte Sendeleistung verursacht Reststörungen in der Zelle, in welche das mobile Endgerät wandert. Dies hat zur Folge, dass alle Teilnehmer in der neuen Sendezelle ihrerseits die Sendeleistung erhöhen müssen, um die aktuellen Kommunikationssitzungen aufrechterhalten zu können. Diese Leistungssteigerung aller Sendeteilnehmer in einer Zelle führt dann zu einer Reduzierung der Kapazität in dieser Zelle. Dieser Sachverhalt macht im *UTRAN* ein Regulierungsverfahren notwendig und wird als Makrodiversität bezeichnet. Hier wird beim Zellwechsel ein zweiter Kommunikationskanal aufgebaut und das Endsystem ist somit für eine gewisse Zeit zeitgleich über zwei *Uu*-Schnittstellen mit einer oder verschiedenen *Node B*'s verbunden. Hierüber wird die

identische Information doppelt übertragen. Die *Node B*'s leiten die empfangenen Informationen über deren *Iub*-Schnittstelle an den *RNC* weiter, der die beiden Informationen so kombiniert, dass sich ein fehlerfreies Signal ergibt, welches dann über die *Iu*-Schnittstelle an das Kernnetz weitergeleitet wird. In umgekehrter Richtung teilt der *RNC* das vom Kernnetz kommende Signal auf und leitet es an die jeweiligen *Node B*'s weiter. In diesem Fall muss das Endsystem beide empfangenen Signale kombinieren. Dieser Vorgang wird auch Combining und Splitting genannt.

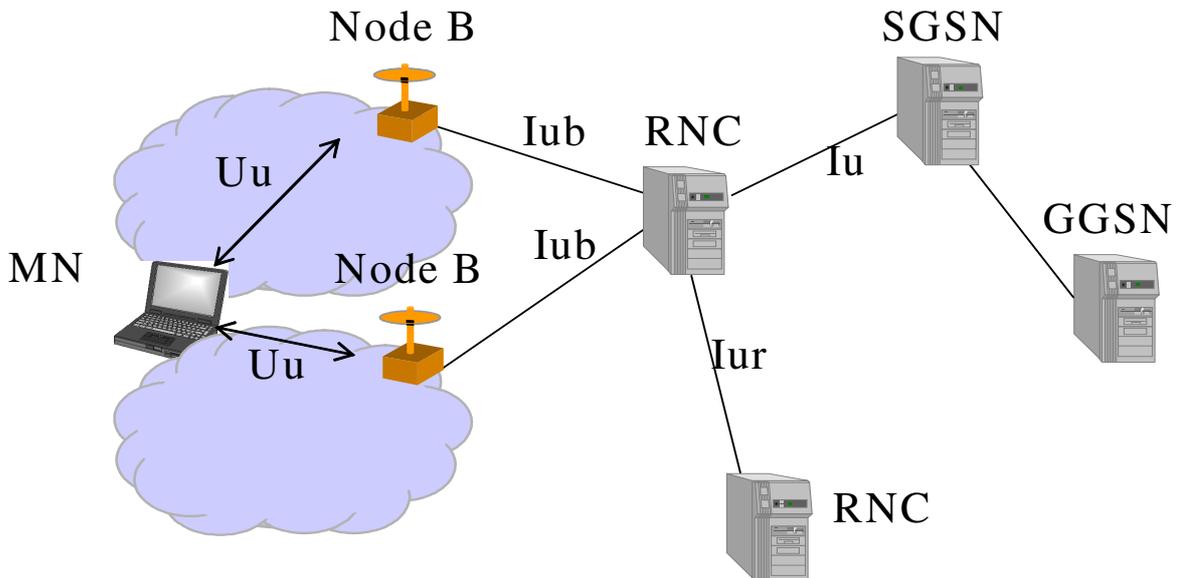


Abbildung 18: UTRAN-Architektur

Die Makrodiversität trägt somit zusätzlich zu einer effizienten Mobilitätsverwaltung bei. Bewegt sich ein Endsystem auf eine neue Zelle zu, so muss nicht erst ein neuer Kanal zugewiesen bzw. aufgebaut werden, da dieser Kanalaufbau bereits im Überlappungsgebiet zweier in ein Mobilitätsszenario involvierter Zellen vollzogen wurde und das Endsystem sowohl über die alte Zelle als auch die neue Zelle mit der Kommunikationsgegenstelle verbunden ist. Der Zellenübergang wird in der Regel vom Benutzer nicht wahrgenommen und wird in der Literatur als Soft-Handover bezeichnet [120].

Dies bedeutet, dass kein Frequenzwechsel beim Übergang von der alten Zelle in die neue Zelle erforderlich ist und das Endsystem innerhalb einer so genannten Makrozelle bleibt. Ist beim Kanalwechsel ein Frequenzwechsel verbunden, so spricht man von einem harten Netzzellenwechsel (Hard-Handover). In *UTRAN* kommt es zu einem Hard-Handover, wenn entweder die Zellularebene gewechselt wird (z.B. von Makrozelle zu Mikrozone) oder wenn ein Wechsel der Technologie (z.B. in das *GSM/GPRS*-Netz) durchgeführt wird. Bei einem Soft-Handover kann man zwischen drei verschiedenen Arten unterscheiden:

- Netzzellenwechsel innerhalb des Bereiches einer *Node B*: Beim Soft-Handover innerhalb des Bereiches einer *Node B* findet das „Combining/Splitting“ in der *Node B* statt, sodass nur eine einzige *Iub*-Schnittstelle erforderlich ist, um die Information zum *RNC* weiterzuleiten.
- Netzzellenwechsel innerhalb des Bereiches eines *RNC*: Hier befindet sich das Endsystem im Überlappungsgebiet zweier Zellen, welche von verschiedenen *Node B* gesteuert werden, jedoch an denselben *RNC* angeschlossen sind. Das „Combining/Splitting“ erfolgt im *RNC*.

- Netzzellenwechsel zwischen *RNC*s: Hier befindet sich das Endsystem im Überlappungsgebiet zweier Zellen, die zu unterschiedlichen *RNC*-Bereichen gehören. Ein *RNC* erfüllt nun die Rolle eines so genannten Ausführenden *RNC*s (*Serving RNC SRNC*), der die Anbindung an das Kernnetz durchführt und auch das „Combining/Splitting“ übernimmt. Der zweite *RNC* erfüllt die Rolle eines Zubringer *RNC* (*Drift RNC - DRNC*), der die Informationen zum *SRNC* durchschleust und die angeschlossenen *Node B*'s verwaltet.
- Abbildung 19 zeigt die Protokollarchitektur von *UMTS*. Hier wird ersichtlich, dass die IP-Schicht nicht als Konvergenzschicht bezeichnet werden kann, da bei *UTRAN* wesentliche Funktionalitäten von der darunter liegenden *ATM*-Schicht, beziehungsweise dem *GTP*-Protokoll (*GPRS Tunneling Protocol*) und dem *PDC*-Protokoll (*Packet Data Convergence Protocol*) erbracht werden.

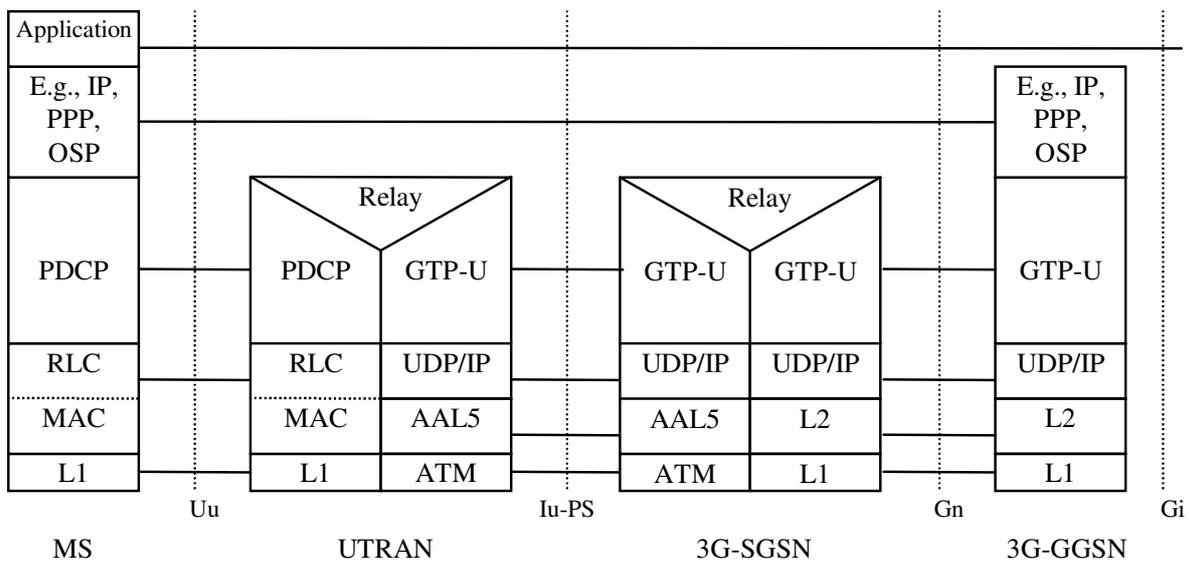


Abbildung 19: IP Tunneling in UMTS

Das *GTP*-Protokoll wird für den Paketdatenverkehr zwischen *SGSN* und *RNC/UTRAN* verwendet. Als Transportmechanismus zwischen *SGSN* und *RNC* wird *ATM* verwendet. Um perspektivisch auch *IPv6* zu unterstützen, wurde die *PDCP*-Schicht eingeführt. Folgende weitere Funktionen führt das *PDCP* aus:

- Header-Kompression und Dekompression von IP-Datenströmen (z. B. Transportkontroll-Protokoll *TCP/IP*).
- Übertragung von Nutzerdaten, d.h. *PDCP* überträgt eine *PDCP*-Dienstdateneinheit (*Service Data Unit, SDU*) der Nichtzugriffsschicht, indem es sie an die *RLC* weiterreicht (Senderseite) und sie von der *RLC* entgegen nimmt (Empfängerseite).

2.2.2.3 UMTS Dienstgütequalität

Das *UMTS* stellt den Endgeräten Dienstgüte (*QoS*) durch den so genannten *PDP-Kontext* Mechanismus bereit. Um Daten versenden oder empfangen zu können, muss ein Endgerät, das sich einmal am *UMTS* Netz angemeldet hat, einen *PDP-Kontext* aktivieren. Diese *PDP-Kontext* Prozedur besteht aus einer Reihe von Signalisierungsnachrichten, die einem *UMTS*-Netzknoten erlauben, einen virtuellen Kanal zum Endgerät aufzubauen. Dieser Kanal wird zwischen *GSGN* und Endsystem etabliert und beinhaltet als Parameter die Dienstgüteanforderungen der Anwendung für die entsprechende Verbindung (Abbildung 20).

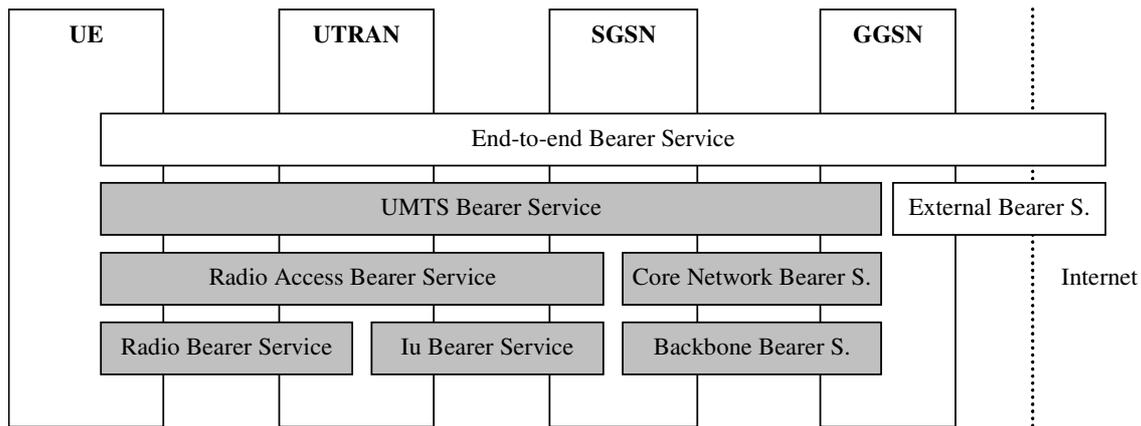


Abbildung 20: UMTS QoS-Architektur

2.2.2.4 3GPP

Das 3rd Generation Partnership Projekt oder 3GPP ist eine weltweite Kooperation von Standardisierungsgremien für die Standardisierung im Mobilfunk. Abbildung 21 zeigt die 3GPP-Referenzarchitektur [101].

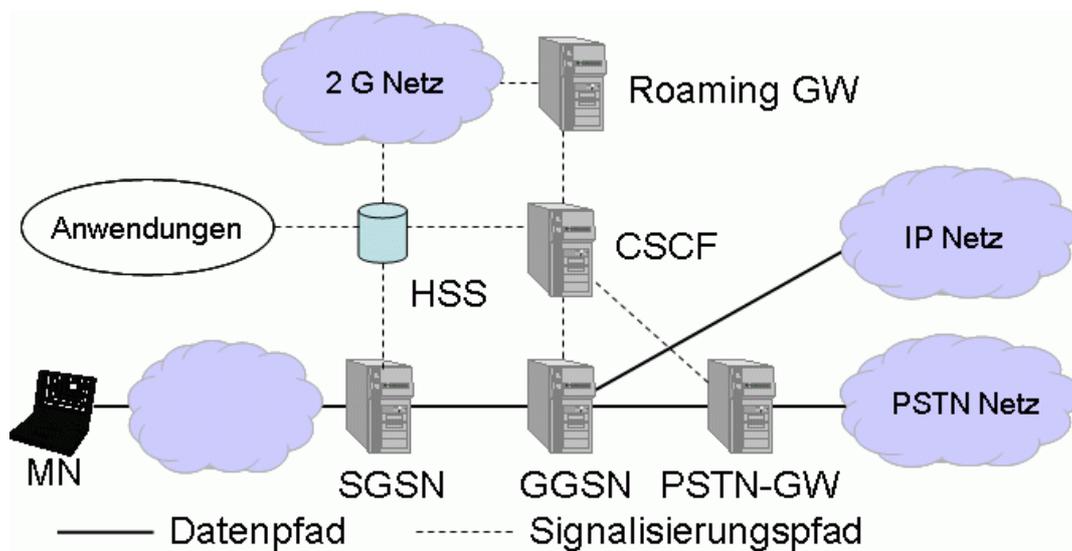


Abbildung 21: 3GPP-Referenzarchitektur

3GPP spezifiziert zwischen *UTRAN* und dem Kernnetz drei verschiedene Schnittstellen. Dies sind Schnittstellen zum leitungsvermittelnden Netz, zum paketvermittelnden Netz und zum Verteilnetz. Alle drei benutzen zur Übertragung der Nutzdaten ATM und hier speziell die AAL-Schicht (AAL5 und AAL2). Die Signalisierung zum leitungsvermittelnden und paketvermittelnden Netz erfolgt über die ATM-Signalisierung, während zum Verteilnetz keine eigentliche Signalisierung benötigt wird.

IPv6 wurde als Routingprotokoll des Kernnetzes in 3GPP vorgesehen. Der größte Vorteil hierbei wird im erweiterten Adressraum gesehen. Aktuell wird nicht mehr von einem sehr schnellen Übergang von IPv4 zu IPv6 ausgegangen, sondern eher von einem langsamen und in mehreren Einzelschritten stattfindenden Prozess, der dazu führt, dass diese Übergangsphase durch ein problemloses Zusammenarbeiten von IPv4 und IPv6 erleichtert bzw. unterstützt werden muss.

2.2.2.5 3GPP2

Unter dem Akronym *3GPP2* (*3rd Generation Partnership Project 2*) wurde eine Organisation mit dem Ziel gegründet, eine ALL-IP-Netzarchitektur für CDMA 2000-Netze zu spezifizieren, da die aktuelle *3GPP*-Architektur z.B. IPv6, jedoch kein Mobile IPv6 unterstützt. Die Architekturen von *3GPP* und *3GPP2* ähneln sich sehr. Ebenso wie die *3GPP*-Architektur kann die *3GPP2*-Architektur in verschiedene Teile eingeteilt werden. Das Zugangnetz und der Zugangsknoten stellen dem mobilen Endsystem IP-Dienste zur Verfügung.

Abbildung 22 gibt einen Überblick über die *3GPP2*-Referenzarchitektur. Während in *3GPP* alle Datenbanken in einem so genannten *HSS* (*Home Subscriber Server*) zusammengeführt sind, ist in der *3GPP2*-Architektur äquivalent dazu eine Instanz von Datenbanken mit vergleichbaren Funktionen vorgesehen. In *3GPP* wurde das Dienstgütemanagement in die *RAN*-Instanzen verlagert:

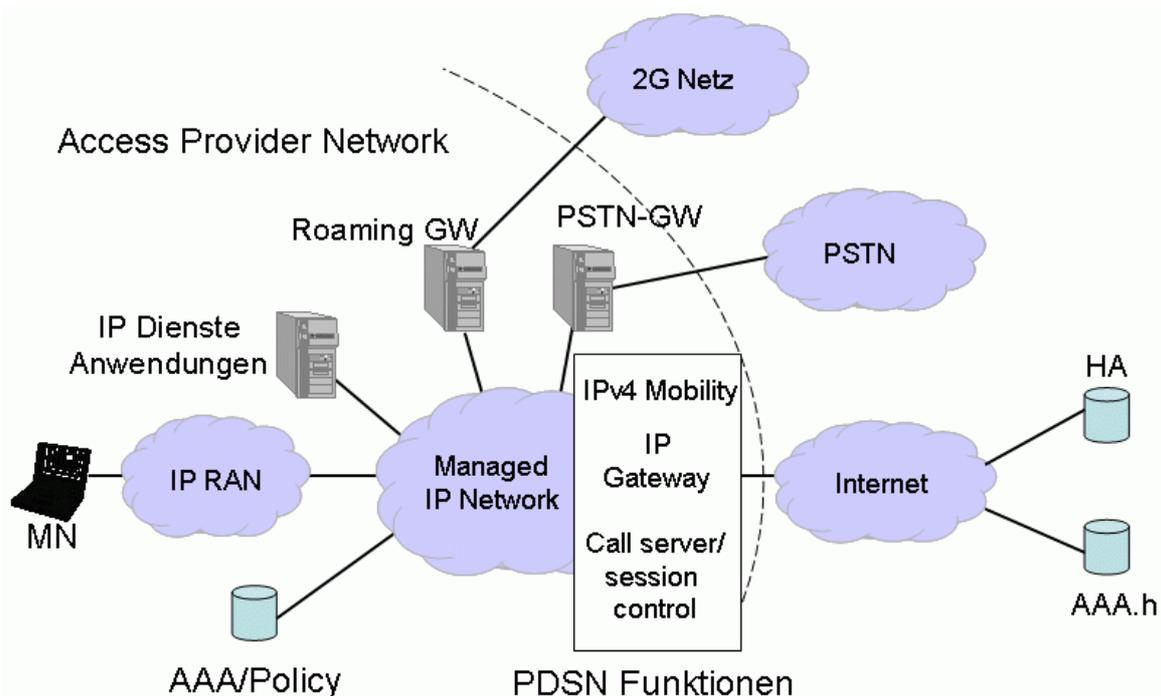


Abbildung 22: 3GPP2 All IP-Referenzarchitektur

Hierfür wurde eine eigene Instanz vorgesehen. Diese ist verantwortlich für Policy-Entscheidungen, die benutzerspezifisch sind und in einem Benutzerprofil abgelegt sind. Zusätzlich ist ein Dienstgütemanager vorgesehen, der die Ressourcen auf dem Kernnetz verwaltet. Die *3GPP2*-All-IP-Architektur sieht eine Mobilitätsverwaltung auf Netzschicht vor. Konkret wurde zur Mobilitätsverwaltung ausschließlich Mobile IP als Protokoll vorgesehen, jedoch noch in der Version Mobile IPv4. Weiter wird hier jedoch nur die Mobilitätsverwaltung zwischen Zugangsknoten betrachtet. Mobilitätsszenarien, bei denen lediglich ein Zugangsknoten involviert ist, werden mit Hilfe eines proprietären Protokolls unterstützt.

Die Gateways der *3GPP2*-Architektur sind prinzipiell von den Gateways der *3GPP*-Architektur abgeleitet. Die Integration der *3GPP2*-Architektur in das Internet kann aktuell als noch nicht abgeschlossen betrachtet werden, so dass hier noch einige Verfeinerungen und Anpassungen zu erwarten sind. Jedoch wurden erste Lösungen für Sicherheit, QoS, Mobilitätsverwaltung, AAA, Transport und Policy angegangen. Sowohl *3GPP*, als auch *GPP2* haben die IETF AAA-Architektur für den jeweiligen ALL-IP Ansatz vorgesehen (siehe hierzu auch Kapitel 4.4). In diesem Zusammenhang hat *3GPP2* die Verwendung von AAA für Mobile

IPv4 das RADIUS-Protokoll [111] festgelegt. Für weiterführende Informationen wird auf [101] verwiesen.

2.3 Zusammenfassung

In diesem Kapitel wurden die für den weiteren Verlauf dieser Arbeit relevanten netztechnischen Grundlagen vorgestellt. Dieses waren im Wesentlichen die fundamentalen Eigenschaften des Internets und des Telefonnetzes, einige Evolutionsstufen des Mobilkommunikationsnetzes und die jeweiligen mobilen Ausprägungen dieser Konzepte. Da allgemein davon ausgegangen wird, dass diese beiden Netzarchitekturen migrieren, soll nun abschließend zusammenfassend ein Überblick über die fundamentalen Unterschiede gegeben werden.

Die aktuell unter dem Begriff „ALL-IP“ arbeitende Aktivität der Mobiltelefonie im Sinne der 3GPP- und vor allem der 3GPP2-Gruppe hat sich zwar einen Schritt in Richtung Internet bewegt, hat jedoch nicht alle leitungsvermittelnden Konzepte durch paketvermittelnde Konzepte ersetzt. Somit hat der Begriff ALL-IP in diesem Zusammenhang eine irreführende Bedeutung und hat z.B. im Projekt Moby Dick [81] dazu geführt, Begriffe wie *Pure-IP* und *Real All-IP* einzuführen, mit dem eine konsequentere Adaption der traditionellen Internet-Philosophie verbunden ist. Auf diese Thematik wird in Kapitel 4 noch näher eingegangen.

Fundamentale Unterschiede, die im weiteren Verlauf dieser Arbeit von Bedeutung sind, ist das „Always-On“ Paradigma des Internets, welches bei der Mobilkommunikation so nicht implementiert ist, da dies, vor allem durch die implementierten verbindungsorientierten Prinzipien, Ressourcen verschwenden würde. Weiter sind die Charakteristika der verschiedenen Internet-Transportprotokolle in der Mobilkommunikation nicht strukturell und ressourcenschonend implementiert, da dort durch Reservierungsmechanismen explizit Kapazitäten reserviert – d.h. konkurrierenden Benutzern vorenthalten – werden. Als weiterer fundamentaler Unterschied zwischen den beiden konvergierenden Prinzipien kann festgehalten werden, dass das Internet über keine strikte Trennung zwischen Signalisierung und Daten verfügt. Im weiteren Verlauf dieser Arbeit wird sich dies noch als wesentlicher Nachteil herausstellen.

Es kann an dieser Stelle vorab festgehalten werden, dass in der Internet-Gemeinde nicht alle Mechanismen bereitgestellt werden, welche notwendig sind, um ein zukünftigen migriertes Kommunikationsnetz im Sinne der traditionellen Netzbetreiber zu betreiben. Um diese betrieblichen und teilweise nichttechnischen Aspekte besser einordnen zu können, werden im anschließenden Kapitel die Grundlagen der nichttechnischen Aspekte eines kommerziell arbeitenden Netzbetreibers vorgestellt.

3 Ökonomische Aspekte eines Netzbetreibers

3.1 Einleitung

In diesem Kapitel werden die nichttechnischen Aspekte eines kommerziellen Netzbetreibers vorgestellt, welche dazu dienen, für im weiteren Verlauf dieser Arbeit relevante Anforderungen an die Netzarchitektur eines kommerziellen Netzes ableiten zu können.

Im weiteren Verlauf dieses Kapitels wird herausgearbeitet, dass es für einen kommerziellen Netzbetreiber unerlässlich ist, jegliche Regelwerke, Mechanismen und Dienste in einer benutzerorientierten Sichtweise zu implementieren, anzuwenden und bereitzustellen. Diese benutzerorientierte Sichtweise ist zwingend notwendig, um in einem kommerziellen Netz Anforderungen (z.B. aus dem Marketing) entsprechend angemessen berücksichtigen zu können. Das Kapitel gibt entsprechend einen Überblick über wesentliche Mechanismen und Konzepte und zeigt beispielhaft auf, wie diese in aktuellen Netzen der Mobilkommunikation realisiert sind. Weiter wird gezeigt, dass diese Mechanismen im Internet – und somit auch im Mobil Internet – nicht ausreichend berücksichtigt worden sind. Ein weiterer wesentlicher Punkt ist die Entkopplung von Kosten und Preis in einem kommerziellen Umfeld [127]. Im ersten Teil dieses Kapitels spielt der Preis, d.h. das Entgelt, welches ein Benutzer für eine bestimmte Leistung bezahlen muss, eine zentrale Rolle. Im zweiten Teil dieses Kapitels werden die Kosten eines Netzbetriebs genauer betrachtet. In diesem Zusammenhang sind Kosten ein Maß für den Aufwand, den ein Netzbetreiber erbringen muss, um eine Leistung bereitstellen zu können, für die ein Benutzer einen Preis bezahlt. Insgesamt legt dieses Kapitel den Grundstein für die Kernmotivation dieser Arbeit – das Entwickeln Internetspezifischer Mechanismen, welche erforderlich sind, um eine Mobile Internet-Infrastruktur im Sinne eines Betriebskonzeptes, wie aus 2G bekannt, realisieren zu können.

Prinzipiell ist im Kontext dieser Arbeit zwischen zwei verschiedenen Disziplinen der Betriebswirtschaftslehre zu unterscheiden. Dies sind zum einen die betriebswirtschaftliche Kostenstruktur eines Netzbetreibers und zum anderen das zugehörige Marketing. Die betriebswirtschaftliche Kostenstruktur ist aus Sicht eines Unternehmens eine nach innen gerichtete Sichtweise, in der interne Abläufe und deren Kosten betrachtet werden. In der Regel gilt es diese zu minimieren, um den Ertrag und letztendlich den Gewinn zu maximieren. Weiter gibt es eine aus Sicht eines Unternehmens nach außen gerichtete Sichtweise, welche im Rahmen des Marketings eines Unternehmens abgedeckt wird. In der klassischen Marketing-Lehre wird hier zwischen 4 verschiedenen Unterdisziplinen unterschieden. Dies sind Kommunikationspolitik, Produktpolitik, Distributionspolitik und Preispolitik [89].

Kommunikationspolitik bezeichnet das strategische Konzept eines Unternehmens sich in der Öffentlichkeit darzustellen. Diese Unterdisziplin der klassischen Marketing-Lehre wird in dieser Arbeit nicht weiter verfolgt.

Produktpolitik umfasst alle Aktivitäten zur Gestaltung der Sach- und Dienstleistungen, sowie deren Kombination. Im Zusammenhang dieser Arbeit ist das Produkt der internetbasierte Netzzugang. Jegliche Serviceleistungen, die mit diesem Zugang als Bündel komponiert werden könnten (z.B. vor Ort Installationshilfe usw.), sind von den weiteren Betrachtungen in dieser Arbeit ausgeschlossen.

Distributionspolitik beschreibt die Strategie, mit der das Unternehmen das Produkt verteilt. Hier findet man in der Praxis Direktvertreiber, Vertrieb über Großhandel, usw. Auch diese

Marketing-Linie ist im Falle eines Telekommunikationsnetzes im Sinne dieser Arbeit nicht weiter relevant.

Preispolitik beschreibt die Konfiguration der Preise und die verschiedenen Strategien, wie ein Produkt über den Preis (mit Unterstützung anderer Marketing-Elemente) am Markt positioniert wird. Im weiteren Verlauf dieser Arbeit wird sich zeigen, dass die preispolitischen Maßnahmen eines Netzbetreibers einen sehr starken Einfluss auf die Netzinfrastruktur haben, so dass im weiteren Verlauf dieser Arbeit die Aspekte der Preispolitik und mit Einschränkungen einige Aspekte der Produktpolitik näher betrachtet werden.

3.2 Grundlagen der Preispolitik

Bevor nun auf die Preispolitik und die dazu verbundenen theoretischen Grundlagen näher eingegangen wird, sollten zunächst die Begriffe Preis, Preispolitik und Preistheorie eindeutig definiert werden. Als Preis versteht man das Entgelt für die Leistungen, die ein Unternehmen auf einem Markt anbietet. Unter Preispolitik werden alle marktbezogenen Maßnahmen und Entscheidungen des Unternehmens verstanden, mit denen Preise beeinflusst und durchgesetzt werden können. Hier muss sich ein Unternehmen im Rahmen seiner globalen Strategie für eine spezielle Preisstrategie entscheiden, welche dann unter Ausschöpfung des Entscheidungsspielraums durchgesetzt wird. Die Preispolitik kann sich generell auf das Gesamtangebot eines Unternehmens, auf Teilbereiche oder auf Einzelprodukte oder Einzelleistungen beziehen. Unter Preistheorie versteht man die Methode, mit denen sich die Auswirkungen einer Preispolitik auf bestimmten Märkten vorhersagen lassen.

Generell müssen Preise weniger gerecht gestaltet werden, sondern zweckgerichtet. Die Beurteilung, ob ein Preis zweckgerichtet ist, gestaltet sich als nicht trivial, da hier sowohl technische, als auch ökonomische Betrachtungsweisen von wesentlicher Natur sind. Die Entscheidung für einen bestimmten Preis ist von den globalen Unternehmenszielen eines Unternehmens abhängig. Sicherlich ist in erster Linie die Erzielung von Gewinn ein wichtiges preispolitisches Ziel eines jeden Unternehmens. Die Gewinnmaximierung ist jedoch nicht das alleinige Ziel, vielmehr sind Preisentscheidungen letztlich von weiteren, zum Teil kurzfristigeren Unternehmenszielen abhängig. In den seltensten Fällen strebt ein Unternehmen mit seiner Preispolitik nur ein Ziel an, sondern es wird eine Kombination von Zielen verfolgt, wie z.B. Gewinnmaximierung und gleichzeitige Marktanteilssteigerung. Hier spielt der zeitliche Bezug eine wichtige Rolle, was erfordert, dass die einzelnen Maßnahmen zeitlich optimiert aufeinander implementiert werden müssen, was wiederum eine Abstimmung der Ziele erfordert. Wenn sich beispielsweise ein Unternehmen ein Image als exklusiver Anbieter hochwertiger Ware aufbauen will, können Aktionen mit Billigstware zu Tiefstpreisen die Erreichung des langfristigen Zieles nachhaltig stören. Die beiden Komponenten „Billigstware“ und „Tiefstpreise“ würden in diesem Fall nicht in die übergeordnete Unternehmensstrategie passen, da Begriffe wie „Billigstware“ und „Tiefstpreise“ zu einem Image führen, welches sich nicht fördernd für die Erreichung des Ziels „exklusiver Anbieter hochwertiger Ware“ auswirkt. Generell lassen sich Preisziele in zwei Gruppen einteilen. Dies sind unternehmensbezogene Ziele und marktbezogene Ziele, von denen nun einige exemplarisch aufgezählt werden [89].

Unternehmensbezogene Ziele:

- Erzielung einer bestimmten Gewinnrate, bezogen auf den Gesamtumsatz pro Produkt, pro Produktgruppe bzw. pro Markt
- Erzielung einer bestimmten Gewinnrate, bezogen auf die Kosten pro Leistungseinheit
- Erzielung einer bestimmten Kapitalrentabilität (Return on Investment, ROI)
- Erreichen bzw. Beibehalten eines bestimmten Beschäftigungsgrades im Unternehmen

- Realisierung einer bestimmten Kostensituation

Marktbezogene Ziele:

- Erschließung neuer Märkte
- Erhaltung bzw. Ausbau bzw. Verbesserung der eigenen Marktposition
- Gewinnen von Marktanteilen
- Eindringen in neue Marktsegmente
- Verhinderung des Markteintritts, bzw. Ausschaltung von Mitbewerbern

Je nach Zielsetzung kann ein Unternehmen unterschiedliche Preisstrategien verfolgen. Dabei wird vor allem danach unterschieden, ob das Unternehmen einen Markt erschließen will, ob eingeführte Produkte abgesetzt, oder ob mit differenzierten Preisen eine Marktanteilssteigerung erreicht werden soll.

Man kann festhalten, dass preispolitische Entscheidungen in erster Linie auf dem Markt getroffen werden. Eine Orientierung an den Kosten, die ein Unternehmen für die Bereitstellung eines Dienstes aufwendet, ist in diesem Zusammenhang nur insofern wichtig, als dass sie einem Unternehmen einen Anhaltspunkt geben, ob das aktuelle Preisniveau, zu dem ein Unternehmen seine Produkte am Markt absetzen kann, positiv oder negativ zum Unternehmenserfolg beiträgt. Langfristig sollte natürlich jede Werteeerstellung kostendeckend sein, jedoch ist es nicht ungewöhnlich, dass ein Unternehmen zum Erreichen anderer Ziele wie z.B. Steigerung des Marktanteils, von dem Ziel Kostendeckung zeitweise abrückt.

Abschließend lässt sich festhalten, dass der Preis in direktem Zusammenhang mit den Zielgrößen Umsatz, Gewinn und Marktanteile steht. Preisänderungen sind meist einfach durchzuführen, jedoch ist die Vorhersagbarkeit ihrer Wirkung nicht immer einfach, da sie von weiteren Faktoren abhängen. Nun soll auf die preispolitischen und preistheoretischen Grundlagen, deren Verständnis für den weiteren Verlauf dieser Arbeit notwendig ist, näher eingegangen werden.

3.2.1 Relevante Marktformen

In einem Wirtschaftsraum müssen bei der Betrachtung der Preisbildung zunächst die Marktverhältnisse auf dem relevanten Markt erörtert werden. Diese sind ausschlaggebend für die Höhe eines Preises, der für ein Produkt oder eine Dienstleistung am Markt erzielt werden kann. Die Betrachtung eines Marktes kann sowohl aus Sicht der Anbieter, als auch aus Sicht der Nachfrager erfolgen.

Prinzipiell unterscheidet man zwischen idealen Märkten, also Märkten, in denen es ausreichend viele Anbieter bzw. Nachfrager gibt, oligopolistischen Märkten, also Märkten mit einer limitierten Anzahl von Anbietern bzw. Nachfragern und monopolistischen Märkten, also Märkten mit einem Anbieter bzw. Nachfrager. Tabelle 1 zeigt noch einmal die charakteristischen Marktformen auf. Ein monopolistischer Markt hat entweder das Monopol, also die allein herrschaftliche Marktstellung auf der Käufer oder Verkäuferseite. Ein oligopolistischer Markt ist durch eine begrenzte Zahl von Anbietern bzw. Nachfragern gekennzeichnet, die dann untereinander durchaus in einem Wettbewerb stehen, der mehr oder weniger stark ausgeprägt sein kann. Im Idealen Markt treffen infinite Nachfragen auf infinite Angebote.

Beim Mobiltelekommunikationsmarkt handelt es sich um einen Markt, in dem eine überschaubare Zahl von Anbieter ihre Produkte einer sehr viel größeren Zahl von Nachfragern anbieten. So gibt z.B. es beim 2G-Netz in Deutschland genau vier Netzbetreiber (T-Mobile, Vodafone, e-plus und O2) und ca. 68 Millionen Kunden. Diese Marktform kommt der Be-

schreibung des Angebots-Oligopols aus Tabelle 1 am nächsten. Da diese Marktform charakteristisch für die gesamte Branche ist, werden alle weiteren Marktformen aus Tabelle 1 im Rahmen dieser Arbeit nicht weiter verfolgt.

Tabelle 1: Marktformen [89]

Angebot Nachfrage	Viele (Ideal unendlich)	Oligopolistisch (wenige)	Monopolistisch (einer)
Viele (Ideal unendlich)	Idealer Markt	Angebots- Oligopol	Angebots- Monopol
Oligopolistisch (wenige)	Nachfrage- Oligopol	Bilaterales Oligopol	Beschränktes Angebots-Monopol
Monopolistisch (einer)	Nachfrage Monopol	Beschränktes Nachfrage-Monopol	Bilaterales Monopol

Zusätzlich wird zwischen offenen und geschlossenen Märkten unterschieden. Bei einem offenen Markt handelt es sich um einen Markt, zu dem jeder Interessent freien Zugang hat. Ein geschlossener Markt ist ein Markt, bei dem der Zutritt mit Genehmigungen und Auflagen verbunden ist, deren Vergabe von verschiedenen Faktoren abhängig sein kann. Die meisten Märkte sind offene Märkte. Für weiterführende Informationen wird auf [130] verwiesen.

Im Falle des Telekommunikationsmarktes handelt es sich um einen geschlossenen Markt, da Funklizenzen, also das Recht das Medium Luft zur Übertragung von Sprache und Information auf einer bestimmten Frequenz kommerziell nutzen zu dürfen, staatlich reglementiert sind. In der Bundesrepublik Deutschland gibt es eine Regulierungsbehörde für Post und Telekommunikation [109], die im Jahr 2001 das Recht versteigerte, das für das 3G-Netz vorgesehene Frequenzband für eine begrenzte Zeitdauer für die Übertragung von Sprache und Daten ökonomisch nutzen zu dürfen (vergleiche Abbildung 17). Anbieter, die in dieser Versteigerung nicht zum Zuge gekommen sind, werden vom Eintritt in diesen Markt als Netzbetreiber ausgeschlossen. Jedoch muss man in diesem Zusammenhang erwähnen, dass sich beim „2G-Markt“ eine Art Zweitmarkt gebildet hat. Hier treten Großhändler auf, die Sprachkapazität in großen Mengen von einem oder mehreren dieser vier lizenzierten Anbieter aufkaufen, und diese dann weiterverkaufen.

Betrachtet man jedoch das Mobile Internet ausgehend von IEEE 802.11 [56] und die jeweiligen Erweiterungen wie IEEE 802.16, IEEE 802.20 usw., so kann man festhalten, dass für das Betreiben eines so genannten WLAN-Hotspots keine explizite Lizenz notwendig ist, da der Staat einige Funkbereiche freigestellt hat. Dieser Markt ist sozusagen für jeden potenziellen Anbieter frei zugänglich und es handelt sich um einen offenen Markt. Somit kann von Migration von Mobilem Internet und Mobiltelekommunikation in Richtung B3G-Netz (Beyond 3G) aus regulatorischer Sicht keine Rede sein, was zu einer sehr unübersichtlichen Marktform führt, in der offene und geschlossene Märkte konkurrieren, ein Betreiber jedoch in beiden aktiv sein kann.

3.2.2 Überblick über Preisstrategien

Wie bereits erwähnt, ist Preispolitik nie isoliert zu betrachten, da mit ihrer Hilfe wesentliche strategische Unternehmensziele verfolgt werden. Dies sind beispielsweise Gewinnmaximierung, Umsatzmaximierung, maximale Kapazitätsauslastung, Verdrängung von Mitbewerbern und maximale Marktabschöpfung.

Die Preisgestaltung wird allgemein als eine der wichtigsten unternehmerischen Entscheidungen betrachtet. Preisentscheidungen beeinflussen in der Regel sehr kurzfristig den Unter-

nehmensgewinn, da sich Preisänderungen unmittelbar auf das Umsatzniveau auswirken. Aus diesem Grund sind Preisänderungen oft eine der ersten Sanierungsmaßnahmen in Unternehmen. Darüber hinaus haben Preisentscheidungen auch erhebliche langfristige Auswirkungen. Für die meisten Unternehmen besteht das grundsätzliche Ziel in der Gewinnerzielung und Gewinnmaximierung. Dies bedeutet, dass das Produkt aus Sicht des Kunden einen entsprechenden Nutzen haben muss, der den dafür verlangten Preis rechtfertigt.

Preisentscheidungen erfordern stets die Balance zwischen zwei gegensätzlichen Kräften: Der Preis muss hoch genug sein, um Erträge für das Unternehmen zu gewährleisten, und niedrig genug, um den Kunden einen ausreichenden Kaufanreiz zu bieten. In diesem Sinne haben Preisentscheidungen eine interne und eine externe Komponente.

Der interne Blickwinkel orientiert sich an den strategischen Bedürfnissen des Unternehmens. Hier wird der Preis so festgesetzt, dass mindestens die Fixkosten und die Variablen Kosten des Produktes bzw. der Dienstleistung gedeckt werden können. In der Kombination aus dem Preis je Einheit und der zu diesem Preis absetzbaren Menge soll ein Gewinn erzielt werden.

Externe Gesichtspunkte der Preisgestaltung orientieren sich mehr an den strategischen Zielsetzungen des Unternehmens. Diese können, wie bereits erwähnt, in der Ausdehnung des Marktanteils, im Setzen neuer Standards durch das eigene Produkt, oder in der Aufrechterhaltung eines bestimmten Unternehmensimages liegen. Diese Aussagen zur Bedeutung der Preispolitik für das Unternehmen weisen darauf hin, dass Preise von verschiedenen Faktoren beeinflusst werden, bei denen die zur Leistungserstellung ermittelnden Fixkosten und Variablen Kosten in gewissen Grenzen unberücksichtigt bleiben.

Der Preis für ein Produkt oder für eine Dienstleistung muss erstmalig gebildet werden, wenn ein Unternehmen ein Produkt neu entwickelt bzw. neu in sein Sortiment aufnimmt, oder wenn laufende Produkte in einen neuen Markt, z. B. in einen Auslandsmarkt, oder über einen neuen Absatzweg einführt werden. Zur Positionierung eines Produktes in der richtigen Kombination von Preis und Qualität wurde ein Neun-Strategien-Modell von Kotler entwickelt [75] (Abbildung 23).

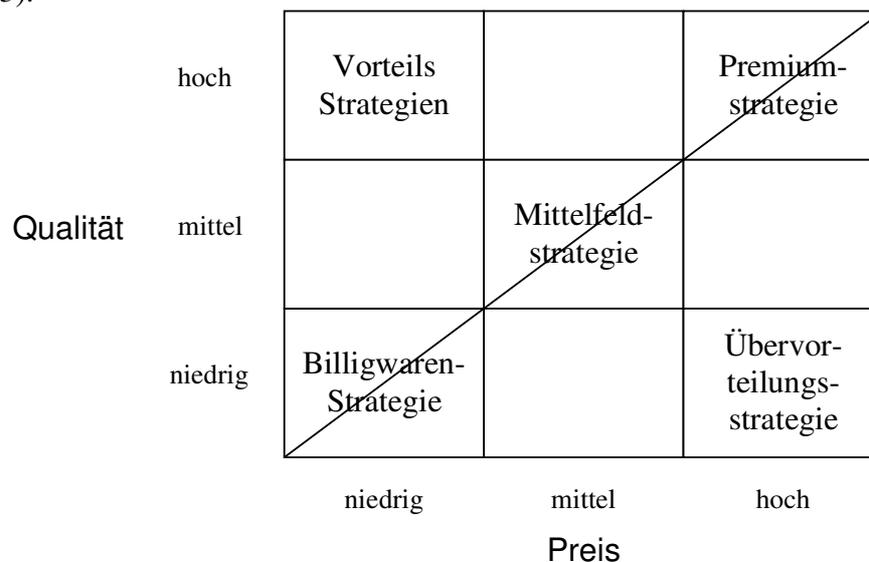


Abbildung 23: Neun-Strategien-Modell

Die Preisstrategien entlang der eingezeichneten Diagonalen weisen ein ausgewogenes Verhältnis zwischen Qualität und Preis aus. Premiumstrategien sind Strategien, bei denen Produkte von hoher Qualität zu einem hohen Preis angeboten werden. Entsprechend wird bei

Mittelfeldstrategien durchschnittliche Qualität zu durchschnittlichen Preisen angeboten und bei Billigwaren-Strategien niedere Qualität zu niedrigen Preisen. In diesen Fällen besteht ein ausgewogenes Preis-Leistungs-Verhältnis. Diese Strategien können gleichzeitig auf einem Markt nebeneinander bestehen und die Wettbewerber können erfolgreich agieren, sofern es folgende drei Käuferschichten gibt:

- Käufer mit hohem Qualitätsanspruch
- Käufer mit hohem Preisbewusstsein
- Käufer, die zwischen Preis und Qualität abwägen

Die Strategie-Positionen in der linken oberen Hälfte (oberhalb der eingezeichneten Diagonalen) stellen Vorteilsstrategien dar, die für Wettbewerber mit den Premium, Mittelfeld und Billigwarenstrategien gefährlich werden können. Die Vorteilsstrategien eignen sich insbesondere, um Wettbewerber mit ausgeglichenem Preis-Qualitäts-Verhältnis anzugreifen. Der Anbieter offeriert dem Kunden hier entweder die gleiche Qualität zu einem günstigeren Preis oder höhere Qualität zum gleichen Preis. Das Auftreten von Marktteilnehmern mit aggressiven Vorteilsstrategien kann unter Umständen dazu führen, dass sich das aus Kundensicht angemessene Verhältnis von Preis und Qualität verschiebt. Schutz vor Wettbewerbern mit solchen Strategien bietet eine aktive Markenpolitik. Die Marke verleiht dem Produkt zusätzlich zu imitierbaren physischen Leistungs- und Qualitätsmerkmalen zusätzliche Eigenschaften wie Status oder Image. Diese sind wesentlich schwerer zu kopieren und können so durch eine starke Differenzierung dauerhaft höhere Preise gewährleisten. Preisbewusste Kunden, die den Qualitätsaussagen dieser Anbieter vertrauen, werden deren Produkte kaufen, weil sie dann weniger Geld ausgeben müssen. Die Übervorteilungsstrategien in der unteren rechten Hälfte (unterhalb der eingezeichneten Diagonalen) weisen ein unausgeglichenes Preis-Leistungs-Verhältnis auf. Der Preis für das Produkt wird als zu hoch empfunden bzw. ist viel zu hoch für den Nutzen, den das Produkt dem Kunden bietet. Kunden, die solche Produkte kaufen, werden sich ggf. übervorteilt fühlen und dieses Produkt nicht wieder kaufen bzw. das Produkt erfüllt trotz mangelnder Qualität andere Werte wie z.B. Image welches den Preis gerechtfertigt. Ein weiteres Beispiel, bei dem in der Praxis solche Angebote akzeptiert werden sind z.B. höhere Lebensmittelpreise an Tankstellen, oder ein höheres Preisniveau in der Einführungsphase eines Produktes.

3.2.2.1 Niedrigpreisstrategien

Zur Niedrigpreispolitik zählen hauptsächlich die Strategien:

- Promotionspreisstrategie
- Penetrationspreisstrategie

Typisch für die Promotionspreisstrategie sind Niedrigpreise. Dies bedeutet, dass der Preis für ein Produkt entweder unter dem Preis vergleichbarer Produkte liegt, oder der Preis erscheint in der Käufer-Wahrnehmung als entsprechend niedrig. Bei der Promotionspreisstrategie ist der Preis das wichtigste Werbeargument. Deshalb wird bei allen Werbemaßnahmen der günstige Preis der Produkte besonders herausgestellt. Diese Preisstrategie kann sowohl für ein gesamtes Produktprogramm bzw. Sortiment gelten, als auch für einzelne Produkte innerhalb breiter Produktlinien. Mit der Promotionspreisstrategie wird versucht, auf niedrigen Preisen beruhend ein bestimmtes Preisimage zu schaffen, das längerfristig beibehalten werden soll. Somit soll ein mittel bis langfristiger Wettbewerbsvorteil gegenüber Konkurrenten erzielt werden. Dieser Wettbewerbsvorteil wird in der Regel durch Skaleneffekte und die Erfahrungskurve begründet. In diesem Zusammenhang sei auch auf [117] verwiesen.

Die Penetrationspreisstrategie zielt darauf ab, mit relativ niedrigen Preisen für neue Produkte schnell Massenmärkte zu erschließen und große Absatzmengen bei niedrigen Stückkosten zu

erzielen. Zusätzlich sollen mögliche Imitatoren und künftige Mitbewerber vor einem Markteintritt abgeschreckt werden. Die Einführungspreise werden dann mit der Zeit schrittweise erhöht. Der Einsatz der Penetrationspreisstrategie ist sinnvoll, wenn:

- Niedrige Preise momentane und potentielle Mitbewerber vom Markteintritt abhalten.
- Durch schnelles Absatzwachstum trotz niedriger Stückdeckungspreise hohe Gesamtdeckungsbeiträge erzielt werden können.
- Die Nachfrager sehr preissensibel reagieren und niedrige Preise höhere Marktanteile erbringen.

Mit dem Einsatz der Penetrationspreispolitik sind auch Risiken verbunden. So dauert es bei Niedrigpreisen in der Regel länger, bis sich die Investitionen in ein neues Produkt amortisiert haben. Auch werden von Abnehmern niedrige Preise oft mit einer geringen Produktqualität assoziiert. Zudem ist der preispolitische Spielraum begrenzt, wenn während der Markterschließung Mitbewerber auftreten oder wenn die Nachfrage falsch eingeschätzt wird. Oft ist es nicht möglich, geplante Preiserhöhungen zu einem späteren Zeitpunkt am Markt durchzusetzen. Als Variation dazu peilen Unternehmen häufig einen hohen Marktanteil mit einem niedrigen Produktpreis an und setzen, wie bereits erwähnt, mit zunehmender Fertigungserfahrung auf sinkende Stückkosten um dadurch Preise weiter reduzieren zu können.

3.2.2.2 Hochpreisstrategien

Zur Hochpreisstrategie, in der Literatur auch Hochpreispolitik genannt, zählen vor allem zwei Strategien:

- die Premiumstrategie
- die Abschöpfungspreisstrategie oder Skimming-Preisstrategie

Kennzeichen der Premiumstrategie ist das Festlegen hoher Preise. Verbunden damit ist in der Regel eine hohe Produktqualität. Die Produkte haben meistens einen bestimmten Zusatznutzen und ein bestimmtes Produktimage. Diese werden durch entsprechende Werbemaßnahmen unterstützt und durch entsprechende Distributionssysteme vertrieben. Typische Beispiele für die Premiumstrategie sind hochwertige Markenprodukte oder Luxusprodukte wie sie bei Uhren, Parfüms, speziellem Interieur oder bestimmten Autos zu finden sind, welche im Zusammenspiel von Produkt- und Kommunikationspolitik geeignet dargestellt werden.

Im Gegensatz zur Premiumstrategie ist die Abschöpfungsstrategie nicht auf die Dauer angelegt, sondern es wird nur in der Einführungsphase eines neuen Produktes ein hoher Einführungspreis verlangt. Die Abschöpfungspreisstrategie ist mit hohem Werbeaufwand verbunden. Meist handelt es sich um Produktneuheiten, für die kurzfristig keine großen Produktionskapazitäten vorhanden sind. Die Absatzmengen sind daher relativ niedrig und die Stückkosten relativ hoch. Mit zunehmender Markterschließung oder mit aufkommendem Wettbewerbsdruck wird der Produktpreis dann gesenkt, damit neue Käuferschichten gewonnen werden können. Der Einsatz der Abschöpfungsstrategie ist beispielsweise unter folgenden Bedingungen sinnvoll:

- Der hohe Preis wird mit hoher Produktqualität und hohem Prestigewert verbunden.
- Es gibt genügend viele Abnehmer, die bereit sind, einen hohen Preis für das Produkt zu zahlen. Bei nachfolgenden Preissenkungen ist es möglich, breitere Konsumentenschichten anzusprechen, die preiselastisch reagieren.
- Die kleineren Absatzmengen bei Neueinführung des Produktes bringen trotz höherer Stückkosten eine höhere Gewinnspanne. Dadurch werden der erhöhte Werbeaufwand und eventuell die späteren Kosten zur Erschließung des Massenmarktes gedeckt.

- Durch den hohen Einführungspreis bei echten Produktneuheiten ist eine schnelle Amortisation des Forschungs- und Entwicklungsaufwandes möglich.

Jedoch macht ein hoher Einführungspreis den Markt für potentielle Konkurrenten attraktiver, was sich langfristig als negativ auswirken kann, sollten diese vermehrt in den Markt eintreten. Deshalb versuchen Unternehmen beim Einsatz der Abschöpfungsstrategie Marktbarrieren für Imitatoren aufzubauen. Diese können beispielsweise durch Patente, speziellem Know-how oder erhöhtem Kapitalbedarf für die Beschaffung von Materialien errichtet werden.

Die Abschöpfungsstrategie wird sehr intensiv bei Computern, der Telekommunikations- und Konsumgüterindustrie mit technologischen Innovationen angewendet. So ist der Preis für neue Prozessoren der Firma Intel in der Einführungsphase auf den höchstmöglichen Preis festgelegt, der sich gegenüber Konkurrenzprodukten bezüglich der Produktneuheiten durchsetzen lässt. Bei jedem Absatzrückgang wird der Preis gesenkt, um die nächste Schicht preisbewusster Kunden zu gewinnen. Somit wird in jedem Marktsegment der größtmögliche Umsatz abgeschöpft.

3.2.3 Preisbildungsmethoden und Preisbildungsmechanismen

Es gibt mehrere Methoden, mit deren Hilfe Preise festgelegt werden. Teilweise haben sich diese bereits in unserem alltäglichen Leben bewährt, müssen aber im Einzelfall auf Ihre Eignung hin überprüft werden. In den folgenden Abschnitten sollten ausgewählte Methoden und Mechanismen der Preisbildung vorgestellt werden.

3.2.3.1 Auktionen

Vor ca. 2000 Jahren waren mathematische Modelle, mit deren Hilfe Märkte beschrieben werden können, noch nicht verfügbar und es etablierte sich mit Hilfe von Auktionen ein sehr gut funktionierender Markt, dessen prinzipielle Methoden und Mechanismen bis heute noch Gültigkeit besitzen und in weiten Bereichen des täglichen Lebens eingesetzt werden.

Die dominante Form des Bietens auf einem Markt ist die *Englische Auktion*. Hier beginnt ein Auktionator mit einem so genannten Vorbehaltspreis die Auktion. Dieser Vorbehaltspreis signalisiert den Interessenten den niedrigsten Preis, zu dem der Verkäufer prinzipiell bereit ist, das Gut abzugeben. Interessenten haben nun Gelegenheit, durch gegenseitiges Überbieten das Gut zu ersteigern. Hierfür muss nur im Vergleich zu den anderen, um das Gut konkurrierenden, Bietern eine größere Zahlungsbereitschaft signalisiert werden. Zur Effizienzsteigerung legt man ein so genanntes Vadium fest. Dies ist ein Betrag, um den ein Interessent das vorangegangene Gebot mindestens überbieten muss, damit der Auktionator dieses Gebot akzeptiert. Prinzipiell würden Auktionen auch ohne dieses Vadium funktionieren, jedoch beschleunigt dieses Vadium den Gesamtprozess wesentlich.

Eine andere Form der Auktion ist die *Holländische Auktion*. Hier beginnt der Auktionator mit einem sehr hohen Preis, den er so lange schrittweise herabsetzt, bis sich ein Kaufinteressent gefunden hat. Der Vorteil dieses Verfahrens im Vergleich zur Englischen Auktion ist, dass sie schneller abgewickelt werden kann, da schon der erste Bieter für das Produkt den Zuschlag erhält und der Vorgang somit abgeschlossen ist.

Eine weitere Auktion ist die *Auktion mit versiegelten Geboten*. Hier teilt jeder Bieter in einem geheimen Verfahren dem Auktionator seine detaillierte Zahlungsbereitschaft mit. Der Auktionator nimmt diese Gebote entgegen und vergleicht sie. Derjenige, der das höchste Gebot abgegeben hat, erwirbt somit das Gut zu dem im versiegelten Gebot gebotenen Preis.

Bei der *Vickery-Auktion* handelt es sich um eine Auktionsform, die sich nur leicht von der *Auktion mit versiegelten Geboten* abhebt. Hier wird zwar das Gut dem Meistbietenden zugeschlagen, jedoch zu dem Preis des zweithöchsten der versiegelt abgegebenen Gebote. Es wurde z.B. in [117] gezeigt, dass diese Art von Auktion im Wesentlichen dasselbe Ergebnis liefert als die Englische Auktion, allerdings sind hier keinerlei Iterationsschritte notwendig. Somit kommt diese Form der Auktion zu einem identischen Ergebnis bei reduziertem Zeitbedarf. Weitere Informationen zur Theorie der Auktionen sind in [90] zu finden.

3.2.3.2 Preisdifferenzierung

Von Preisdifferenzierung wird immer dann gesprochen, wenn ein Gut zu unterschiedlichen Preisen verkauft wird und sich diese Preisunterschiede nicht oder nicht gänzlich durch Kostenunterschiede begründen lassen. Auf unvollkommenen Märkten kann ein Anbieter durch Preisdifferenzierung seinen Gewinn erhöhen. Er spaltet dazu den Gesamtmarkt auf und verkauft auf den Teilmärkten gleiche Güter zu unterschiedlichen Preisen. Tabelle 2 gibt eine Übersicht über die unterschiedlichen Ausprägungen der Preisdifferenzierung.

Eine Preisdifferenzierung kann von einem einzelnen Anbieter nur insoweit durchgeführt werden, als er in der Lage ist, seinen Preis selbst festzulegen. Neben dem Monopolisten kann dies auch der Anbieter auf Polypol- und Oligopolmärkten innerhalb seines "monopolistischen Preisspielraumes". Diese Zusammenhänge wurden in [107] erstmals diskutiert. Dort führt Pigou drei Arten der Preisdifferenzierung ein.

Tabelle 2: Ausprägungen der Preisdifferenzierung

AUSPRÄGUNG	BEISPIELE
Räumliche Preisdifferenzierung	Niedrigerer Verkaufspreis inländischer Autos im Ausland; höhere Kraftstoffpreise bei Autobahntankstellen
Preisdifferenzierung nach Art der Abnehmer	Differenzierung der Stromtarife (teurerer Haushalts-, billigerer Industriestrom); niedrigere Eintrittspreise für Schüler und Studenten
Zeitliche Preisdifferenzierung	Zeitliche Preisstaffelung bei Miete von Ferienreisen; Winterpreise/Frühjahrspreise bei heimischem Obst;
Mengenmäßige Preisdifferenzierung	Gewährung von Mengenrabatten.
Preisdifferenzierung nach Verwendungszweck	Unterschied zwischen technischem Alkohol und Alkohol für den menschlichen Konsum bzw. Speisesalz und Industriesalz.

Bei der Preisdifferenzierung der ersten Art, wird jede Einheit an jenes Individuum verkauft, das die höchste Wertschätzung für das jeweilige Gut entgegenbringt. Dieser Verkauf erfolgt zu dem höchsten Preis, den dieses Individuum zu zahlen bereit ist. Im Idealfall bedeutet dies, dass der Produzent entlang der linearen Preis-Absatz-Funktion seine produzierten Mengeneinheiten verkaufen kann. Hier wurden die Annahmen getroffen, dass jeder Konsument die gleiche Menge des Gutes bezieht und gleichzeitig bei infinitesimal kleinen Preisabständen zwischen den Preisniveaus Nachfrage existiert.

Aus wirtschaftstheoretischer Sicht gibt es keine andere Möglichkeit, sowohl Konsument als auch Produzent optimaler zufrieden zu stellen. Der Gewinn des Produzenten kann nicht erhöht werden, und auch produzierte Menge ist so groß, dass der Markt vollständig bedient wird [107].

In der Praxis ist diese feingranulare Art der Preisdifferenzierung sehr schwer implementierbar, da der Aufwand zur Feststellung der individuellen Zahlungsbereitschaft so hoch ist, dass er nicht gerechtfertigt erscheint. Dies liegt unter anderem daran, dass unter Einsatz von Bündelungsmechanismen das Kosten/Nutzenverhältnis aus Sicht eines Betreibers optimiert werden kann. Die beiden verbleibenden Arten der Preisdifferenzierung behandeln solche Bündelungsarten.

Als Preisdifferenzierung zweiten Grades bezeichnet man jene Fälle, in denen der Preis von der gekauften Menge abhängt (Mengenrabatt).

Als Preisdiskriminierung dritten Grades bezeichnet man jene Fälle, in denen die Preise mit Merkmalen der Nachfrager variieren. Diese Art der Preisdifferenzierung kommt im täglichen Leben häufig vor z.B. wenn Fahr- oder Eintrittskarten an Schüler und Studierende günstiger verkauft werden als an "normale" Kunden.

3.2.4 Rolle des Preises

Wie bereits erwähnt, erfüllt der Preis verschiedene Funktionen. Die nachstehend genannten Preisfunktionen treffen in ihrer Gesamtheit nur für ideale Märkte zu. Auf Oligopol- und Monopolmärkten muss eine mehr oder minder starke Einschränkung vorgenommen werden.

- Gleichgewichtsfunktion [69]: Der Preis stellt das Gleichgewicht zwischen Angebot und Nachfrage auf dem Markt her.
- Signalfunktion [69]: Der Preis signalisiert den Knappheitsgrad eines Gutes. Dabei ist nicht die absolute Preishöhe relevant, vielmehr zeigen Preisänderungen an, wie sich die Knappheitsverhältnisse verschieben.
- Allokationsfunktion [69]: Der Preis lenkt das Angebot und damit den Einsatz der Produktionsfaktoren auf diejenigen Märkte, bei denen die größte Nachfrage herrscht und folglich der höchste Preis und damit der höchste Gewinn zu erzielen ist.
- Erziehungsfunktion [69]: Da im Falle des Idealen Marktes der Preis eine vom Anbieter nicht beeinflussbare Größe ist, wird er versuchen, möglichst sparsam mit den Produktionsfaktoren umzugehen, um seine Kosten niedrig zu halten und den Gewinn zu maximieren. Andererseits werden auch die Nachfrager bestrebt sein, die preisgünstigsten Einkaufsmöglichkeiten wahrzunehmen, um ihren Nutzen zu maximieren.
- Kompensationsfunktion [69]: Alle Kosten und Aufwendungen, die bei der Bereitstellung und dem Konsum von Gütern und Diensten anfallen, werden mit Hilfe der Preise erfasst und abgegolten.

Prinzipiell handelt es sich beim Internet und beim Mobiltelekommunikationsnetz um Systeme mit begrenzten Ressourcen. Mehrere Benutzer konkurrieren folglich um die limitiert zur Verfügung stehenden Ressourcen. Diese begrenzten Ressourcen können zum einen die verfügbare Bandbreite sein, die zwischen den konkurrierenden Benutzern geteilt werden muss, oder aber andere physikalische Beschränkungen, wie z. B. eine limitierte Kapazität von Einwahlknoten oder eine limitierte Anzahl von IP-Adressen. Beim Mobil Internet/Mobiltelekommunikation werden die Ressourcen vorwiegend durch das für die Kommunikation vorgesehene Frequenzband begrenzt. Sobald ein Kommunikationsnetz auf kommerzieller Basis betrieben wird, hat das Betreiberunternehmen verschiedene Ziele zu erfüllen. Diese Ziele können sich über mehrere Dimensionen erstrecken, die mit Hilfe einer geeigneten Preispolitik erreicht werden können.

Zunächst werden Preise dazu verwendet dem unternehmerischen Ziel der Profitmaximierung Rechnung zu tragen. Hier wird die vom Unternehmen erbrachte Leistung bewertet. Diese

Bewertung spiegelt sich direkt im Preis wieder. Als Bewertungsfaktoren spielen eingeschränkt die eigenen Kosten und vor allem die Marktverhältnisse eine wesentliche Rolle.

Um eine Bewertung eines Preismodells für einen Netzbetreiber auf qualitativer und quantitativer Ebene durchführen zu können, muss eine Analyse durchgeführt werden, welche die Kosten des Netzbetreibers entsprechend detailliert erfasst. Basierend auf dieser Kostenstruktur können dann Preise gebildet werden. Dieses Vorgehen ist zwar aus wirtschaftlicher Sicht sinnvoll, muss sich jedoch in der Regel den Marktverhältnissen den Unternehmenszielen unterordnen.

Preise haben auf vergleichbaren Märkten (z. B. Rohstoffmärkten), in denen ebenfalls Ressourcen limitiert zu Verfügung stehen, eine weitere Bedeutung. Sie werden zur Nachfragesteuerung eingesetzt, um die Nachfrage optimal an die eigenen Produktkapazitäten bzw. Ressourcenkapazitäten anzupassen.

Ähnlich wie bei den Unternehmen der Energieversorgung stehen die Ressourcen in der Telekommunikation nicht unbegrenzt zur Verfügung. Jedoch ist die zur Verfügung stehende Menge an Ressourcen im Falle der Energieversorger innerhalb einer gewissen Bandbreite statisch. Ähnlich verhält es sich auf dem Telekommunikationsmarkt. Hier ist ein gewisses Angebot an Ressourcen vorhanden und soll dann möglichst effektiv dem Endkunden verkauft werden. Das Angebot an dieser Ressource kann als absolut statisch betrachtet werden, denn ein Telekommunikationsnetz erfordert aufwendige Installationsarbeiten und komplexe Netzknoten, die neben anderen physikalischen Eigenschaften und regulatorischen Bedingungen die maximal zur Verfügung stehende Kapazität begrenzen. Leider ist aus Sicht der Netzbetreiber die Nachfrage nach dieser Ressource über den Tag verteilt nicht konstant, sondern es sind vorher-sagbare Schwankungen zu beobachten. Diese Schwankungen sorgen dafür, dass zu bestimmten Zeiten Überlastzustände auftreten können, da die Nachfrage das Angebot übersteigt. Zu anderen Zeiten wiederum ist ein Überangebot an Ressourcen festzustellen. Nun ist es aus Sicht der Anbieter von Ressourcen, die ein solch schwankendes Nachfrageverhalten antreffen, wünschenswert, die Nachfrage nach diesen zu verteilenden Ressourcen entsprechend auszugleichen. In diesem Zusammenhang ist festzuhalten, dass in der Telekommunikation unverbrauchte Ressourcen (im Gegensatz zu z.B. Rohstoffen) nicht zwischengespeichert werden können und bei Nichtallokation ersatzlos verfallen.

Somit ist ein Aspekt der Preispolitik eines Telekommunikationsbetreibers, die Nachfrage nach diesen limitiert zur Verfügung stehenden Ressourcen dahingehend zu steuern, dass sie zu einer möglichst hohen permanenten Netzauslastung unter der Randbedingung der Gewinnmaximierung führt. Geeigneterweise wird über die Preispolitik zusätzliche Nachfrage stimuliert.

Zusammengefasst lässt sich hier festhalten, dass die Rolle der Preisbildung als Unterstützung der technischen Netzplanung und Netzdimensionierung betrachtet werden kann. Dies versetzt dann den Netzanbieter in die vorteilhafte Lage, das Gut Bandbreite zu Marktpreisen unter Minimierung des Aufwandes anzubieten. Wegen der durch effektive Netzplanung minimierten so genannten Grenzkosten entsteht somit für ein Unternehmen ein Kostenvorteil, der sich langfristig positiv auf den Unternehmenserfolg durchschlagen sollte.

3.3 Marktorientierte Preisbildung

Eine Kernanforderung an eine geeignete Preisbildung ist der Bezug zum Wert eines Gutes für den Kunden. In diesem Zusammenhang soll ausdrücklich klargestellt werden, dass sich hier der Wert ausschließlich auf den Nutzen des Konsumenten bezieht, und nicht auf die Kosten

zur Leistungserstellung. Folgt man diesem Gedanken konsequent, dann ist eine Preisdifferenzierung, d.h. das identische Produkt wird verschiedenen Personen zu verschiedenen Preisen bei gleicher Bezugsmenge angeboten, sehr leicht nachvollziehbar, denn es kann davon ausgegangen werden, dass der Nutzen eines Produktes/Dienstes für verschiedenen Konsumenten unterschiedlich ist. Solch eine personalisierte Preisbildung erfordert, dass sehr genaue Kenntnisse über die Kunden vorhanden sind, da nicht davon auszugehen ist, dass Kundendetails über ihren individuellen Nutzen aktiv preisgeben. Somit ist aus Sicht eines Anbieters eines Gutes eine Strategie erforderlich, um dem Kunden ein entsprechendes Preiskonzept zu kommunizieren. Eine Methode dieser Kommunikation ist die Dienstbündelung (Service Bundling). Hier werden verschiedene elementare Dienste zu einem Bündel gruppiert und dann zu einem Preis dem Kunden angeboten. Dienstbündelung hat viele Vorteile ([89]), die nun kurz erläutert werden sollen.

- **Positionierung gegenüber der Konkurrenz:** Ein wichtiger Aspekt ist die Positionierung eines Produktes gegenüber der Konkurrenz. Vor allem wenn es sich um leicht austauschbare Produkte wie z.B. Netzzugang handelt, müssen Unternehmen versuchen, sich gegenüber der Konkurrenz abzuheben. Dies wird über spezielle Dienstbündel erreicht.
- **Kommunikation mit dem Konsumenten:** Ein wesentliches Ziel der Marktforschung ist, dem Konsumenten Informationen zu entlocken, die er in der Regel nicht offensichtlich preisgibt. Über das Anbieten verschiedener Dienstbündel können somit indirekt Benutzer kategorisiert werden, was zu weiteren Produkt- und Preisbildungsstrategien sehr wertvoll für ein Unternehmen sein kann.
- **Gewinnmaximierung:** Eine günstige Konfiguration eines Dienstbündels wirkt sich umsatzsteigernd aus. In der Regel werden Bündel zu einem günstigeren Preis als die Summe der Einzelprodukte angeboten. Ein Konsument, der sich nur für ein oder zwei Einzelprodukte eines Bündels interessiert, ist sehr oft dazu bereit, gegen entsprechend höheres Entgelt, das Bündel zu erwerben, obwohl er die restlichen Bestandteile eines Bündels nicht beziehen wollte. Der fiktive Preisaufschlag bewegt den Kunden jedoch, sich für das Bündel zu entscheiden. Dies führt bei geeigneter Zusammenstellung der Bündel zu einer Maximierung des Gewinns.

Ein Vorteil von Dienstbündelung, welcher auch im weiteren Verlauf dieser Arbeit noch relevant wird, ist eine Vereinfachung der Abrechnungsinfrastruktur. Ein Unternehmen kann durch Einführung von Dienstbündeln die Produktvielfalt erhalten, hat aber weniger Aufwand bei der Generierung der zur Endabrechnung relevanten Rohdaten, da sich diese in ihrer Granularität reduzieren.

3.4 Produktbündelung

Als Produktbündel bezeichnet man ein Paket von zwei oder mehreren Produkten, welche als Paket zu einem Preis angeboten werden. Ein typisches Beispiel für ein solches Produktbündel ist Microsoft Office. Hier werden Tabellenkalkulation, Textverarbeitung, Präsentationsprogramm und eine Datenbank zu einem Paket zusammengesetzt. Jedes dieser vier individuellen Produkte kann jedoch auch einzeln am Markt gekauft werden. Die Gründe für eine solche Bündelung können sehr vielfältig sein. Man kann als Hersteller z.B. versuchen, sozusagen im Sog eines Produktes mit hoher Marktdurchdringung über das Bündel auch weniger verbreitete Produkte am Markt zu etablieren, sollte die Preispolitik das Komplettbündel zu einem entsprechend attraktiven Preis anbieten.

Ein wesentliches Argument für die Definition solcher Preisbündel liegt im Verhalten der Konsumenten. Angenommen es gibt zwei Produkte (Produkt A und Produkt B) und zwei Arten

von Konsumenten (K1 und K2). K1 ist bereit für Produkt A 100 € und für Produkt B 80 € zu bezahlen. K2 hat genau entgegengesetzte Präferenzen. Somit müsste man für jedes Produkt 80 € verlangen, um den maximalen Umsatz zu generieren (320 €). Optimiert man jedoch den Umsatz basierend auf der angenommenen Situation am Markt, so müsste man ein Bündel zu 180 € am Markt anbieten und man hätte den Umsatz um 40 € auf 360 € erhöht. Dieses Beispiel zeigt, dass das Definieren von Dienstbündeln umsatzfördernd angewendet werden kann. In der obigen Beispielrechnung wurden die Kosten zur Leistungserstellung vernachlässigt. Dies ist generell nicht zulässig, jedoch bei Produkten wie z.B. der Softwareindustrie, oder auch im Bereich eines Netzbetreibers in erster Näherung problemlos machbar, da der Anteil der variablen Kosten entsprechend gering ist (vergleiche hierzu auch Kapitel 3.6).

3.5 Existierende Preismodelle in der Telekommunikation

Nachdem nun in den vorherigen Unterkapiteln dieses Kapitels die grundlegenden Mechanismen der Preispolitik, welche für den weiteren Verlauf dieser Arbeit relevant sind, vorgestellt wurden, soll nun eine Analyse der momentan existierenden Preismodelle für die drei Bereiche Festnetztelefon, Mobiltelefon und Internet vorgenommen werden. Ziel dieser Analyse ist nicht, die absolute Höhe der Preise zu evaluieren, sondern eine Übersicht über die Preisstruktur und die relevanten Parameter zu geben, welche grundsätzlich zur Preisbildung verwendet werden. Diese Analyse soll die aktuelle Diskrepanz zwischen den verschiedenen Medien aufzeigen anhand derer dann im weiteren Verlauf dieser Arbeit prinzipielle Probleme aufgezeigt werden, die ein Konvergieren der Netze mit sich bringt.

3.5.1 Preisstruktur im Festnetzbereich

Historisch war in Deutschland die Deutsche Bundespost / Deutsche Telekom Besitzer der Teilnehmeranschlussleitung. Der Aufbau dieser Infrastruktur wurde in einem reglementierten Marktumfeld geschaffen und es ist aus wirtschaftlichen Gesichtspunkten eher unwahrscheinlich, dass ein privater Investor hier flächendeckend eine zweite Infrastruktur aufbaut. Die ehemals staatlichen Telekoms hatten, wie in allen anderen Ländern Europas, auf ihren Märkten eine Monopolstellung, aus der sie sehr bequem Preise bilden konnten. Nach Öffnung des Marktes und gleichzeitiger Privatisierung der Telekom im Jahre 1996, versuchten andere Betreiber durch Mieten der Telekom-Kabelinfrastruktur im Netzanschlussbereich den Kunden einen eigenen Sprachtelefoniedienst anzubieten. Die Telekom versuchte über ein hohes Durchleitungsentgelt diese Wettbewerber vom Eintritt in diesen Markt abzuhalten. Die Regulierungsbehörde legte daraufhin die Preise für die Nutzung dieser Leitungsinfrastruktur fest [109]. Man kann den Markt inzwischen als Angebotsoligopol (vergl. Tabelle 1) bezeichnen. Hier haben sich prinzipiell zwei vorherrschende Preismodelle durchgesetzt. Dies sind der Festvertrag, und ein so genanntes call-by-call Verfahren. Diese beiden Modelle werden nun näher erläutert.

3.5.1.1 Festvertrag

Prinzipiell besteht ein Festnetzvertrag aus einer festen Preiskomponente, der monatlichen Anschlussgebühr, und einer variablen Komponente, die über die Parameter Tageszeit, Sprechdauer, Wochentag und Entfernung eine variable Preiskomponente ermittelt. Prinzipiell wird der Ressourcenverbrauch je nach Komposition der oben genannten Parameter in einen Takt überführt, der dann auf einen variablen Preis abgebildet wird. Es gibt hier sowohl sekunden-genaue Abrechnung, als auch eine Abrechnung in 6-Sekunden bzw. Minuten-Intervallen. Es sollte hier noch erwähnt werden, dass sich diese Preise auf die Dauer beziehen, in der eine Telefonverbindung aufgebaut ist. Jegliche Parameter, die sich am Informationsinhalt, der Informationsmenge usw. orientieren, sind nicht Bestandteil des Preismodells.

3.5.1.2 Call-by-Call im Festnetz

Call-by-Call-Festnetztelefonie ist ein Dienst, der durch den regulatorischen Eingriff der Regulierungsbehörde für Post und Telekommunikation ermöglicht wurde. Hier mieten Betreiber ohne eigenes Netz für die Dauer eines Gespräches die Netzanschlussleitung von der Deutschen Telekom. Bei dem in diesem Bereich eingesetzten Preismodell wird von einem Festpreisanteil abgesehen und nur ein variabler Preisanteil erhoben. Auch hier sind die weiteren Parameter, die als Differenzierungskriterium für das Preismodell eingesetzt werden, Entfernung, Tageszeit und Verbindungsdauer und ggf. der Wochentag. In der Regel gibt es eine grobe Einteilung in Entfernungsbereiche, für die jeweils ein tageszeitabhängiger Preis festgelegt wird. Hierzu werden Zeitzonen gebildet, die eine Arbeitszeit und Freizeit/Nacht trennen. Weiter wird zwischen Arbeitstagen und Wochenende unterschieden. Für Auslandsgespräche wird pauschal zwischen verschiedenen Zielländergruppen unterschieden. Weiter sind in der Vergangenheit Produktbündel aufgetreten, welche so konstruiert sind, dass über einen erhöhten Grundbetrag z.B. an Wochenenden jegliche Kommunikation frei ist.

3.5.1.3 Sonderdienste im Telefonnetz

Beim traditionellen Sprachtelefonienetz wurde durch Einführung des Intelligenten Netzes (IN) die Möglichkeit geschaffen, bestimmte Nummern für Sonderdienste zu reservieren. Bei diesen Sonderdiensten teilen sich der eigentliche Inhalte-Anbieter, auch Content Provider genannt, mit dem Netzbetreiber die dann erhöhten Gebühren auf. Somit hat der Netzbetreiber implizit die Möglichkeit, indirekt am Inhalt der Nachrichten zu partizipieren. Vor allem im Bereich „Erwachsenen-Inhalte“ trägt dieser Dienst sehr stark zum Umsatz sowohl der Netzanbieter, als auch der Inhalte-Anbieter bei.

3.5.2 Preisstruktur im Mobiltelefonbereich

Da das Mobiltelefonnetz historisch auf dem herkömmlichen Sprachtelefonienetz aufbaut, ist es nicht verwunderlich, dass sich das im Mobiltelefonbereich verwendete Preisschema sehr stark dem traditionellen, bei der Festnetztelefonie angewandten Schema ähnelt. Auch hier kann zwischen einem Preisschema, das eine variable Gebührenkomponente und eine Festpreiskomponente enthält, und zwischen einem Preisschema, welches lediglich aus einer reinen variablen Preiskomponente besteht, unterschieden werden.

3.5.2.1 Festvertrag

Der Festvertrag wird in der Regel als Dienstbündel angeboten. Bestandteil des Bündels ist typischerweise ein Endgerät. Die Verträge haben in der Regel eine zeitlich befristete Laufzeit und bestehen aus einem variablen Preisanteil und einem Festpreisanteil, der sich in seiner Höhe dem der Festnetztelefonie orientiert. Der variable Preisanteil beinhaltet nicht mehr den Parameter Entfernung. Ruft man einen mobilen Teilnehmer an, hat man vor Anruf keine Information über den tatsächlichen Aufenthalt des Angerufenen. Anstatt des Parameters Entfernung erfolgt die Preisfestsetzung nach Differenzierung des Zielnetzes. Zusätzlich gibt es eine Preiskomponente Roaming, bei der für den Empfang eines Gespräches bezahlt werden muss, sollte sich der Benutzer nicht im Heimatnetz aufhalten (Empfang eines Gespräches im Ausland). Weiter gibt es Sonderdienste, wie z.B. das versenden einer Kurznachricht (SMS) für die ein spezieller Preis zu entrichten ist, der sich wiederum am Zielnetz orientiert. Bisher konnte man an der Rufnummervorwahl erkennen, in welchem Netz sich der Angerufene befindet. Die Möglichkeit vor Verbindungsaufbau über die Kosten des Anrufes informiert zu werden, wird durch die Rufnummerportabilitätsverordnung teilweise außer Kraft gesetzt (siehe hierzu § 43 Absatz 5 des Telekommunikationsgesetzes). Rufnummernportabilität ermöglicht einem mobilen Kunden bei Wechsel des Netzbetreibers seine Rufnummer behalten zu dürfen. Somit geht der Netzidentifikator absehbar verloren, da aus der Vorwahl nicht mehr direkt auf das Netz des Angerufenen zu schließen ist und somit der Anrufer keinerlei Anhaltspunkte mehr

über den Preis für die beabsichtigte Verbindung erhält. Es wird zwar ein Dienst eingeführt, bei dem ein Anrufer dies vorab kostenlos in Erfahrung bringen kann, jedoch ist die Benutzerakzeptanz dieses Dienstes offen und es bleibt abzuwarten, ob die Rufnummernportabilität eine Auswirkung auf die etablierten Parameter der Preismodelle haben wird.

3.5.2.2 Prepaid-Vertrag

So genannte Prepaid-Verträge haben sich in den letzten Jahren sehr stark verbreitet. Kernkonzept beim Prepaid-Vertrag ist, ein Mobiltelefon mit einem Gesprächsguthaben zu versehen, welches konsumiert werden kann. Diese Endgeräteeigenschaft wird auf der SIM-Karte (vergl. Kapitel 2.2.1.1) hinterlegt. Somit wird bei der Autorisierung ein Teilnehmer als Prepaid-Kunde identifiziert und das verbleibende Guthaben wird überprüft. Es entstehen für den mobilen Teilnehmer keine Fixkosten. Die Gebührenstruktur eines Prepaid-Vertrages unterscheidet sich nicht wesentlich von der des Festvertrages. Es sind lediglich signifikant höhere Gebühren zu beobachten.

Datentarife.

	DataConnect ^{cl}	CombiCard Data ²⁰⁾ Verfügbar nur bei bestehendem T-Mobile Laufzeitvertrag.
Einmaliger Bereitstellungspreis	entfällt	entfällt
Mindestlaufzeit des Vertrages ³⁾	Keine ¹⁶⁾ / 24 Monate	24 Monate
Monatlicher Paketpreis ¹⁾	-	6,00
Monatlicher Grundpreis im 60/1-Sekunden-Takt ¹⁾	5,00*	entfällt ²⁰⁾
Monatlicher Optionspreis für die voreingestellte Datenoption T-Mobile Data 10 ¹⁾	10,00	-
Monatliches Inklusivvolumen ^{b),1)}	10 MB	5 MB
Preis pro MB (nach Verbrauch des Inklusivvolumens) ^{b)}	1,90	3,90
Verbindungspreise Inland pro Minute von T-Mobile ¹⁾ (Für Fax-Datenverbindungen, 9,6 Kbit/s)		
▪ zum deutschen Festnetz, zu T-Mobile ¹⁾ , zur Mobilbox	0,22	0,40
▪ zu Vodafone D2, E-Plus, O ₂	0,50	0,60
Verbindungspreise Inland pro Minute von T-Mobile ¹⁾ (für Sprachverbindungen von T-Mobile ¹⁾)		
▪ zum deutschen Festnetz, zu T-Mobile ¹⁾ , zur Mobilbox	Sprachverbindungen nicht möglich	0,40
▪ zu Vodafone D2, E-Plus, O ₂		0,60
Versand von SMS (Inland) in alle deutschen Mobilfunknetze, je SMS	0,19	0,19

Alle Preise in € inklusive Mehrwertsteuer.

Abbildung 24: Mobile Datendienste (Quelle: T-Mobile)

3.5.2.3 Sonderdienste in der Mobilkommunikation

Im Bereich des Mobilfunks kommen mehrere Sonderdienste zum Einsatz. Ein wesentlicher Dienst ist die ins Netz integrierten Sprachbox, für deren Abfrage ebenfalls eine Pauschalgebühr erhoben wird, in jüngster Zeit tritt dieser Dienst jedoch auch als Bestandteil eines Bündels auf. Ein weiterer, nicht zu vernachlässigender Dienst ist das Versenden von Kurznachrichten (SMS) und der Übergang einer SMS zur E-Mail. Bei der SMS ist es einem Benutzer erlaubt, max. 180 Zeichen zu übertragen, für die dann eine Pauschalgebühr erhoben wird. Sollte ein potentieller Empfänger einer SMS nicht aktiv am Netz registriert sein, wird diese Nachricht entsprechend lange im Netz zwischengespeichert und bei Neuregistrierung dem Ziel übermittelt. Somit handelt es sich hierbei um einen asynchronen Dienst.

3.5.2.4 Preisstruktur für Mobile Datendienste

Wie in Kapitel 2.2 erwähnt kann die GPRS-Technologie als erste Mobilfunktechnologie betrachtet werden, welche Daten und Sprache auf einer mobilen Netzplattform verbindet. Abbildung 24 zeigt exemplarisch ein von T-Mobile angebotenes Preismodell. Ähnliche Preismodelle existieren von anderen Mobilfunkanbietern (z.B. Vodafone, e-plus, O2). Es wird deutlich, dass hier ein erster Versuch unternommen wurde, ein Preismodell mit einer Volumenkomponente (vergleiche Kapitel 3.5.2.3) und einer Verbindungskomponente zu definieren. Für die Nutzung der Datendienste wird ein volumenabhängiger Preis erhoben, der sich an keinem weiteren Parameter orientiert. Zusätzlich wird eine Stundennutzungskomponente erhoben, die konzeptionell identisch mit den verbindungsorientierten Parametern ist, wie sie aus der Festnetz- und Mobiltelefonie bekannt sind.

3.5.3 Preisstruktur im Internet

Dienste im Internet sind am Rande des Netzes platziert und so genannte Dienstanbieter (Service Provider) bieten diese Dienste transparent zur Netzinfrastruktur an. Der Netzanbieter hat hier wenige Möglichkeiten an den Geschäftsmodellen der Dienstanbieter zu partizipieren und generiert lediglich aus dem Transport von Daten Umsatz. Somit handelt es sich beim Internet um eine so genannte "offene Netzplattform". Diese Tatsache hat vermutlich wesentlich zum Erfolg des Internets beigetragen, jedoch hat sich im Internet noch kein befriedigendes Geschäftsmodell etabliert, da historisch die Entwicklung dieser Netzplattform von überwiegend akademischen Einrichtungen vorangetrieben wurde, welche jegliche kommerzielle Denkweise während der frühen Entwicklungsphase nicht ausreichend berücksichtigten.

3.5.3.1 Internetzugang für Private Konsumenten

Den meisten privaten Internetbenutzern steht bisher nur die Telefonleitung als Anschlussmedium zum Internet zur Verfügung. Hier wird mit Hilfe eines Modems oder auch mittels DSL-Technologie eine Verbindung zu einem Internet-Einwahlknoten hergestellt. Somit kann bei der Analyse des Preismodells das angewendete Schema nicht vollständig von dem der Festnetztelefonie entkoppelt betrachtet werden.

Als Gebührenmodell werden hier momentan zwei verschiedene Schemata eingesetzt. Zum einen ist die so genannte Flat-Rate ein Dienst, welcher sich zunehmender Beliebtheit erfreut. Abbildung 25 gibt eine exemplarische Übersicht über verschiedene Internet Tarife über eine Telefonleitung am Beispiel der Deutschen Telekom/T-Online. Aus der Abbildung ist ersichtlich, dass das Preismodell aus einer festen Komponente, einer so genannten Grundgebühr, und einer variablen, rein zeitabhängigen Komponente besteht.

	GRUND- GEBÜHR [€]	VARIABLE PREISANTEIL [CENT / MINUTE]	SONSTIGE BESCHRÄNKUNGEN
T-Online by night	9,90	1,6	23:00 Uhr – 9:00 Uhr
T-Online by day	14,90	1,6	9:00 Uhr – 17:00 Uhr
T-Online eco	8	2,9	
Call by Call	0	3,9	
T-Online surftime 30	29	0	
T-Online surftime 60	55	0	
T-Online surftime 90	79	0	
T-Online surftime 120	99	0	
TDSL Flat	49	0	+ TDSL Kosten

Abbildung 25: T-Online Tarife (Quelle T-Mobile)

Eine weitere Variante des Preismodells enthält lediglich einen volumenabhängigen Pauschal tariff. Das Kombinationsmodell Festkostenanteil/variabler Kostenanteil basiert historisch auf der verbindungsorientierten Systematik der Festnetztelefonie. Die Parameter "30 – 120" in den „T-Online Surftime Tarifen“ bezeichnen die Dauer in Minuten, die in der monatlichen Grundgebühr enthalten ist. Es handelt sich hierbei um eine Flat-Rate mit Zeitbegrenzung.

3.5.3.2 Internet für akademische Einrichtungen

Das Preismodell für einen Zugang direkt ins Internet, d.h. nicht über ein verbindungsorientiertes Medium, unterscheidet sich grundsätzlich von den bisher vorgestellten Preismodellen. Leider kommen bislang nur sehr wenige Endkunden in den Genuss, direkt an das Internet angeschlossen werden zu können.

Abbildung 26 zeigt exemplarisch das Gebührenmodell des Forschungsnetzes des Landes Baden-Württemberg (BelWü) für einen solchen Fall. Es wird deutlich, dass das Preismodell im Wesentlichen aus einem Festkostenanteil besteht, bei dem das Element Preisdifferenzierung eingesetzt wird. Dieser Festpreis staffelt sich dann mit der Geschwindigkeit des Netzzuganges. Weiter wird hier noch eine volumenbasierte Komponente hinzugefügt.

Produkt	Preis in EUR			
	Schule	Bibliothek	Landeseinrichtung	öffentl. Einrichtung
monatliche Nutzungspauschale für eine ISDN-Wählverbindung	0.00	0.00	29.00	58.00
monatliche Nutzungspauschale für eine Festverbindung bis 128KBit	0.00	0.00	220.40	348.00
monatliche Nutzungspauschale für eine Festverbindung bis 512KBit	0.00	0.00	348.00	580.00
monatliche Nutzungspauschale für eine Festverbindung bis 2MBit	0.00	0.00	493.00	812.00
monatliche Grundgebühr für eine Festverbindung ab 10MBit	0.00	0.00	348.00	522.00
Volumengebühr pro GByte Daten und Monat	0.70	0.00	1.16	2.32

Abbildung 26: Internet Anschlussgebühr des regionalen Forschungsnetzes BelWü (Quelle: Belwue)

3.6 Kostenstruktur eines Netzbetreibers

Nach dem in den bisherigen Abschnitten die grundlegende Preismethodik auf dem aktuellen Telekommunikationsmarkt analysiert wurde, soll in diesem Abschnitt ein allgemeiner Überblick über die Kostenstruktur der Netzbetreiber auf diesem Sektor gegeben werden. Fokus hierbei ist nicht die detaillierte Berechnung der absoluten Höhe von Betriebskosten, sondern um einen groben Überblick über Struktur und die Zusammensetzung der Kosten eines Netzbetreibers, welche im weiteren Verlauf dieser Arbeit qualitative Bewertungsgrundlagen liefern sollen. Generell ist bei Kosten zwischen variablen Kosten und Fixkosten zu unterscheiden. Variable Kosten sind Kosten, die sich mit der Menge der Produkte oder Dienstleistungen verändern: Sie sind abhängig vom Beschäftigungsgrad und sinken bzw. steigen mit ab bzw. zunehmender Erzeugnismenge. Hierbei kann man zwischen proportionalen variablen Kosten, progressiven variablen Kosten und degressiven variablen Kosten unterscheiden. Fixkosten sind Kosten, welche von der Erzeugnismenge unabhängig sind. Im Folgenden soll auf diese

Kostenarten näher eingegangen werden. Zur Definition der Kostenarten wird auf [69] verwiesen.

3.6.1 Fixkosten

Die in diesem Kapitel erwähnten Fixkosten geben eine Übersicht über die im Kontext dieser Arbeit relevanten Fixkosten eines Mobilfunk-Netzbetreibers.

3.6.1.1 Lizenzkosten

Lizenzen im Mobiltelekommunikationsbereich werden vom Staat vergeben, um einem Unternehmen die zweckgebundene und mit Auflagen versehene exklusive und kommerzielle Nutzung eines Frequenzspektrums zu erlauben. In der Regel ist dieses Recht auf einen bestimmten Zeitraum begrenzt. Somit wird über diese Politik ein geschlossener Markt etabliert. Bei der UMTS-Auktion in Deutschland wurde die Vergabe der Lizenz mit Auflagen verbunden. Konkret muss ein nach dem UMTS-Standard kompatibles Netz flächendeckend in der Bundesrepublik Deutschland aufgebaut und betrieben werden. Die Netzabdeckung (als zeitlicher Stufenplan) wurde vorgegeben. Diese Lizenzgebühr kann als reiner Fixkostenblock betrachtet werden, da sich deren Höhe nicht an irgendwelchen variablen Parametern orientiert. Das Entgelt für die Lizenzen wurde in den meisten Ländern, darunter auch Deutschland, sofort fällig und es fallen bei einem angenommenen Zinssatz von 5 % für die Telekommunikationsanbieter in Deutschland jährlich ca. 400 Millionen € Zinskosten fiktiv an. Setzt man einen Kredit zu den angenommenen Bedingungen voraus, entspräche dies einer täglichen Zinsrate von 1,1 Millionen €, die seit August 2000 entrichtet werden müssten. Für den Betrieb von WLAN-Hotspots ist hingegen keine besondere Lizenz erforderlich.

3.6.1.2 Installationskosten

Netzinstallationskosten beinhalten zunächst Kosten für die Netzausrüstung, welche für den Aufbau eines Netzes erforderlich sind. Für ein elementares Netz der 3G-Technologie mit deutschlandweiter Flächendeckung rechnet man hier mit ca. 2,5 Milliarden €.

Die Kosten dieser Kategorie werden in der Regel über Kredite finanziert und sind kurzfristig betrachtet nicht gebrauchssensitiv. Es kann jedoch bei zahlreicher Nutzung des Netzes davon ausgegangen werden, dass sie mit der Nutzung des Netzes steigen werden. Dies liegt vor allem daran, dass in der Anfangsphase eines Netzausbaus die Kapazitäten knapp bemessen werden (vergleiche Kapitel 3.2.2). Steigt die Akzeptanz des Netzes an, reichen diese Kapazitäten nicht mehr aus und es muss ein Netzausbau erfolgen, will man die Nachfrage weiterhin bedienen können. Wird also das Netz stark benutzt, erfolgt hier eine Kostenanpassung.

3.6.1.3 Netzinfrasturkturkosten

Netzinfrasturkturkosten beinhalten jegliche mit der Netzinfrasturktur zusammenhängenden Kosten, die von der Nutzung unabhängig sind. Dies sind Kosten wie Miete der Antennenanlagen, Software und technische Ausrüstung.

3.6.2 Variable Kosten

In diesem Kapitel werden die Variablen Kosten detailliert vorgestellt, welche im Kontext dieser Arbeit relevanten sind.

3.6.2.1 Benutzerverwaltungskosten

Um einen Benutzer an das Netz anzuschließen, fallen mehrere Arten von Kosten an. Zunächst gibt es einmalige Kosten, wie das Aufnehmen der benutzerspezifischen Daten. Wesentliche Bestandteile dieser Daten sind Bankverbindung, Benutzername, Benutzeranschrift usw. Diese Kosten entstehen im Rahmen eines so genannten Customer-Care und Vertriebsprozesses.

Weiter entstehen hier benutzerspezifische Verwaltungskosten wie z. B. Gesprächsdatenerfassung und Datenhaltung der Abrechnungsdaten, die später zur Rechnungserstellung, dem so genannten Billing, verwendet werden.

3.6.2.2 Reklamationen

Reklamationskosten sind somit indirekte Kosten, die aber in erster Näherung mit der absoluten Anzahl der Benutzer zunehmen.

3.6.2.3 Rechnungserstellung

Da aus Sicht der Benutzer erwartet wird, die konsumierten Ressourcen bei der aus der Sprachtelefonie abgeleiteten Dienste einzeln nachzuweisen und abzurechnen, steigt der Aufwand für diese Rechnungserstellung mit absoluter Zahl der Benutzer an. Zunächst sind hier die Portokosten für die Versendung der Rechnungen zu erwähnen, was bei einem Kundenstamm von T-Mobile (ca. 28 Mio.) monatliche Fixkosten von ca. 12 Mio. € darstellt. Weiter müssen die Rohdaten erfasst, verarbeitet und den Benutzern zugewiesen werden.

3.6.3 Weitere Kosten

Die im Folgenden aufgeführten Kosten stellen teilweise einen signifikanten Kostenfaktor für einen Netzbetreiber dar, sind jedoch für den weiteren Verlauf dieser Arbeit von nachrangiger Rolle und werden nur der Vollständigkeit halber aufgeführt.

3.6.3.1 Netzverbindungskosten

Netzverbindungskosten beinhalten hauptsächlich die Kosten, die aus Vereinbarungen mit anderen Netzanbietern entstehen, um den Benutzern des eigenen Netzes weltweite Erreichbarkeit zu ermöglichen. Diese Kosten werden auch Peering-Kosten genannt und sind in der Regel kapazitätsabhängig.

3.6.3.2 Betriebs und Wartungskosten

Betriebskosten setzen sich aus Personalkosten und Ausrüstungskosten zusammen, welche nicht zu den Installationskosten hinzugezählt werden.

Wartungskosten sind Kosten, welche nur entstehen, weil der Kunde das Netz gebraucht. Bei der Telekommunikationsindustrie hängen diese nicht direkt von der Nutzungsintensität ab.

3.6.3.3 Personalkosten

Personalkosten für einem Netzbetreiber wird von ca. 125000 € pro Jahr und Mitarbeiter ausgegangen.

3.6.3.4 Kapitalkosten

Zur Bestimmung der Kosten für die Bereitstellung von Netzelementen stellen die Kapitalkosten des eingesetzten Produktivkapitals einen wesentlichen Kostenfaktor dar. Die hohe Kapitalintensität des Betriebs von Telekommunikationsnetzen macht diesen Kostenblock zum wichtigsten Faktor der Kosten des Leistungsangebotes.

3.6.3.5 Externe Kosten

Externe Kosten treten immer dann auf, wenn begrenzt zur Verfügung stehende Güter derartig verteilt werden, dass nicht jeder Nachfrager geeignet bedient werden kann und denjenigen Benutzern, die nicht bei der Ressourcenverteilung berücksichtigt werden konnten, dadurch fiktive Kosten entstehen, welche sich hauptsächlich in Form von Wartezeit und den Kosten, die sich aus dieser Wartezeit ergeben, ausdrücken.

3.6.3.6 Kalkulatorische Kapitalverzinsung

Zinskosten auf das in den Anlagegütern gebundene Kapital werden von den Unternehmen üblicherweise kalkulatorisch angesetzt. Hier von der Prämisse ausgegangen wird, dass das gesamte betriebsnotwendige Vermögen aus dem insgesamt eingesetzten Eigen und Fremdkapital finanziert wird. Für das eingesetzte Eigenkapital entstehen dem Unternehmen zwar keine Zinsaufwendungen, es ist jedoch zu berücksichtigen, dass das Unternehmen zumindest die Rendite einer alternativen Investitionsmöglichkeit erwirtschaften sollte, damit die Eigenkapitalgeber nicht abwandern. Auch bei diesen Kosten handelt es sich um nicht gebrauchssensitive Kosten, die als reine Fixkosten zu betrachten sind.

3.7 Kostenmodell

Die Kosten eines Netzbetreibers setzen sich, wie in den vorhergehenden Abschnitten erwähnt, aus mehreren Faktoren zusammen, auf die im Rahmen dieser Arbeit nicht ausführlich eingegangen werden kann. Dies sind neben Finanzierungs-, Wartungs- und Betriebskosten vor allem die Kosten der Benutzerverwaltung. Hierfür wurde in [69] ein abstraktes Kostenmodell vorgestellt, welches sich auf die für den weiteren Verlauf dieser Arbeit relevanten Kosten, die in direkten Zusammenhang mit einem Preismodell zu betrachten sind, fokussiert. Abbildung 27 zeigt dieses generische Kostenmodell eines Netzbetreibers.

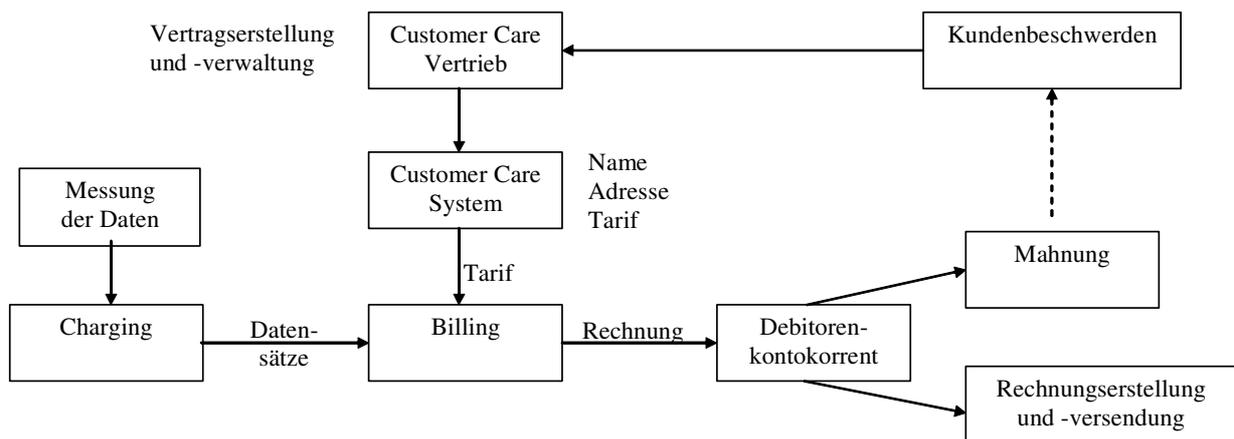


Abbildung 27: Kostenstruktur eines Preissystems

Um einen Benutzer im Netz zu registrieren, fallen mehrere Arten von Kosten an. Zunächst sind hier einmalige Kosten, wie das Aufnehmen der benutzerspezifischen Daten, zu erwähnen. Wesentliche Bestandteile dieser Daten sind Bankverbindung, Benutzername, Benutzeranschrift usw. Diese Kosten entstehen im Rahmen eines so genannten Customer-Care- und Vertriebsprozesses. Weiter entstehen hier benutzerspezifische Verwaltungskosten wie z. B. Gesprächsdatenerfassung und Datenhaltung der Abrechnungsdaten, die später zur Rechnungserstellung, dem so genannten Billing, verwendet werden.

Parallel zu diesen Aufwänden muss die technische Voraussetzung geschaffen werden, Nutzungsdaten mittels einer geeigneten Messtechnik erfassen zu können. Die Kosten dieser Nutzungsdatenerfassung hängen sehr stark vom implementierten Preismodell ab. Bei einem gebrauchorientierten Preismodell sind die Kosten für die Nutzungsdatenerfassung sehr hoch, da gegebenenfalls jedem Benutzer jede einzelne Verbindung bzw. Aktivität im Netz einzeln ausgewiesen werden muss. Die Datenmengen, welche hier anfallen, sind sehr groß. Weiter stellt das Verwalten dieser Daten einen nicht zu unterschätzenden Kostenblock dar. Konkret müssen die Rohdaten konsolidiert, mittels des Charging-Prozesses aufbereitet, und einem Billingprozess zugeführt werden.

Sowohl Messung/Erfassung der Daten, als auch der sich daran anschließende Charging-Prozess, stellen sehr hohe Anforderungen an die erforderlichen IT-Systeme. Der hierfür aufzubringende Aufwand verhält sich in erster Näherung proportional zu der zu erfassenden Datenmenge. Bei einem mobilen Netz, in dem Benutzer zwischen administrativen Netzen wechseln und gegebenenfalls über Netzgrenzen hinweg Kommunikationsszenarien aufrechterhalten, entstehen zusätzlich Kosten, welche das Zusammenführen dieser Messdaten aus den verschiedenen Teilnetzen erforderlich macht. In diesem Zusammenhang ist zu erwähnen, dass ein Netzbetreiber nicht sowohl ein Preismodell frei definieren als auch einem Kunden globales Roaming anbieten kann, da der Roaming-Partner ggf. nicht die erforderliche Infrastruktur zur Nutzungsdatenerfassung bereitstellen kann.

Der Billingprozess verknüpft nun die Daten aus dem Customer-Care System mit den aus dem Messprozess erfassten Daten und führt sie der Rechnungserstellung zu. Konkret bezeichnet man mit Billing den Prozess, bei dem die reinen Rohdatensätze, welche im Charging Prozess aufbereitet wurden, geeignet mit den benutzerindividuellen Tarifen verknüpft werden. Resultat des Billingprozesses ist eine Rechnung für einen Benutzer mit Festvertrag bzw. eine Guthabenreduktion eines Prepaid-Benutzers. Wesentlicher Kostenfaktor ist hier die Datenmenge der zu erfassenden Daten. Natürlich kann diese reduziert werden, indem weniger technische Parameter erfasst werden. Nicht zu unterschätzen ist hier jedoch auch der Zeitbedarf für den Gesamtprozess. Je nach Komplexität der erforderlichen Datenverarbeitung und Anzahl der Kunden ist der Zeitbedarf für die letztendliche Rechnungserstellung und Übermittlung zum Kunden wesentlich. Aber nicht nur die Anzahl der zu messenden Daten ist ein wesentlicher Kostenfaktor, sondern auch die Tatsache, dass bestimmte technische Parameter einfacher und weniger aufwändig messbar sind, als andere. Generell kann man festhalten, dass beim Billing eine sehr hohe Rechenleistung erforderlich ist, die mit der Komplexität der Tarife und der Anzahl der technischen Parameter, die im Rahmen des Messprozess erfasst werden müssen, sehr stark steigt. Beim Billing wird in der Regel Software eingesetzt, für die Lizenzkosten bezahlt werden müssen. Diese Kosten orientieren sich an der absoluten Anzahl der Kunden. Die erforderliche Rechenleistung für den Billingprozess kann direkt in Kosten für die dafür erforderliche Infrastruktur transferiert werden. Die Kosten für Messung, Charging und Billing drücken sich weiter sowohl in Personalbedarf für Wartung, Pflege und Betrieb der Infrastruktur, als auch in Ausrüstung der technischen Einrichtungen aus. Ein Problem ist hier, dass die Daten schnellstmöglich verarbeitet und dem nachgelagerten Prozess zur Verfügung gestellt werden müssen, da es aus Sicht der Liquiditätssituation eines Netzbetreibers erforderlich ist, die Zahlungseingänge so schnell wie möglich zu erhalten. Hierzu müssen die Rechnungen entsprechend schnell bereitgestellt werden. Im Falle eines Prepaid-Kunden ist hier eine Abrechnung in Echtzeit notwendig, wie im weiteren Verlauf dieser Arbeit noch ausführlicher dargestellt wird. Dies gilt ebenso in bestimmten Bereichen des Festnetzes (z.B. für Endsysteme im Hotelzimmer).

Im Debitorenkontokorrent werden die Daten aus dem Billing zu Rechnungen weiterverarbeitet. Hier fallen Kosten für informationstechnische Einrichtungen an. Je nach Preismodell variieren dann die Kosten für die Rechnungserstellung. Gleichzeitig müssen Zahlungseingänge überprüft und Mahnungen verwaltet werden. Basierend auf der Rechnung bzw. Mahnung führen Kundenbeschwerden zu relativ hohen Personalkosten. Im weiteren Verlauf dieser Arbeit wird noch gezeigt, dass diese Kosten in ihrer absoluten Höhe vom Preismodell abhängig sind, da bestimmte Preismodelle tendenziell mehr Angriffsfläche für Kundenbeschwerden bieten als andere. Ein weiterer wesentlicher Teil der Benutzerverwaltungskosten ist das Überprüfen der korrekten Zahlungseingänge.

3.8 Zusammenfassung und Ausblick

Kosten und Preise spielten in diesem Kapitel eine zentrale Rolle. Es wurde aufgezeigt, dass diese beiden Parameter in einem kommerziellen Umfeld nicht so stark voneinander abhängen, wie es oft im technischen Umfeld angenommen wird.

Um in einem kommerziellen Umfeld in der entsprechenden Marktform als Unternehmen erfolgreich zu sein, müssen Mechanismen implementiert werden, welche zunächst zu einer harten Entkopplung von Kosten und Preisen führen. Ein Unternehmen orientiert sich bei der Festsetzung der Preise ausschließlich am Markt und versucht seine Produkte gewinnmaximierend und im Rahmen einer Unternehmensstrategie zu platzieren. Abgesehen von dem Aspekt, dass ein Unternehmen ein strategisches Unternehmensziel verfolgt, das eine kurzfristige und eine langfristige Perspektive beinhaltet, ist ein Unternehmen dann erfolgreich, wenn es einen geeigneten Markt für die eigenen Produkte schafft. Die Schaffung eines Marktes erfordert das erfolgreiche Zusammenarbeiten der vier Disziplinen des Marketings (Preispolitik, Produktpolitik, Distributionspolitik und Kommunikationspolitik).

Im Rahmen dieser Arbeit spielen Kommunikationspolitik und Distributionspolitik eine untergeordnete Rolle und wurden deshalb nicht behandelt. Aspekte von Produkt und vor allem Preispolitik müssen jedoch beachtet werden, um Anforderungen an die kommerzielle Netzinfrastruktur ableiten zu können, wie im weiteren Verlauf dieser Arbeit noch zu sehen sein wird.

Um nun den Unternehmenserfolg zu maximieren, müssen sehr wohl die Kosten betrachtet werden. Dies ist jedoch nicht in jeder Marktform gleich relevant. Tritt jedoch Konkurrenz am Markt auf, ist es erforderlich die Kosten zu minimieren, um nicht in Gefahr zu geraten, an Handlungsspielraum gegenüber einem Wettbewerber zu verlieren. Somit müssen die Kosten zur Leistungserstellung permanent und unabhängig vom Preis der für ein Produkt verlangt wird, minimiert werden.

Betrachtet man sich die Art und Weise der Preisbildung in den existierenden Telekommunikationsmärkten, so kann man festhalten, dass die Konzepte Dienstbündelung und Preisdifferenzierung sehr weit verbreitet sind und es wird allgemein (und besonders im Rahmen dieser Arbeit) davon ausgegangen, dass diese Mechanismen ebenfalls in einem kommerziellen Mobilem Internet zu berücksichtigen sind. In jedem Fall wird aus diesen grundlegenden Betrachtungen die Anforderung an ein kommerzielles Mobiles Internet abgeleitet, sehr flexibel und dynamisch benutzerspezifische Dienste skalierbar anbieten zu können. Dies ist erforderlich, um als Betreiber schnell auf einem Markt reagieren zu können, um ggf. auf Sonderaktionen eines Wettbewerbers, wie z.B. das zeitbegrenzte Anbieten neuer Leistungsbündel am Markt, entsprechende Konkurrenzangebote platzieren zu können.

Weiter soll an dieser Stelle nochmals explizit festgehalten werden, dass jeglicher Dienst im kommerziellen Umfeld einem Benutzer angeboten wird und somit benutzerorientiert ist. Dies erscheint auf den ersten Blick als triviale Aussage, es ist jedoch festzustellen, dass das Internet diese benutzerorientierte Sichtweise noch nicht hinreichend verinnerlicht hat. Im Internet ist jegliche Leistungserstellung dienstorientiert.

Nach dem nun in den beiden bisherigen Kapitel eine breite Basis geschaffen wurde, die sowohl technische als auch nichttechnische Aspekte beinhaltet, werden im nächsten Kapitel die vorhandenen Strategien, Konzepte und Mechanismen vorgestellt, die ein künftiges Mobiles Netz in ein kommerzielles Netz transferieren sollen.

4 Strategien, Konzepte und Mechanismen zur Kommerzialisierung des Mobilten Internet

4.1 Einleitung

In diesem Kapitel werden Strategien, Konzepte und Mechanismen vorgestellt, welche entweder aktuell im technischen Einsatz, oder in der wissenschaftlichen Diskussion sind, um ihren einzelnen Beitrag zur Kommerzialisierung des Mobilten Internet leisten zu können. Da der Begriff kommerzielles Mobiles Internet bisher noch nicht definiert wurde, wird am Ende dieses Kapitels eine exemplarische Architektur eines kommerziellen, Mobilten Internets vorgestellt.

Die Strategien, Konzepte und Mechanismen zur Kommerzialisierung des Internets können in zwei Bereiche eingeteilt werden. Dies sind zunächst jene, welche aktuell (vorwiegend im 2G/3G-Umfeld) eingesetzt werden bzw. kurz vor der Einführung stehen und andere, welche eher als theoretische Modelle existieren und deren Realisierung in der Praxis bisher aussteht. Im Verlauf dieses Kapitels soll für beide Bereiche gezeigt werden, an welcher Stelle im Sinne einer Kommerzialisierung des Internets Verbesserungen notwendig sind, die dann im Rahmen dieser Arbeit dargestellt werden. Es kann an dieser Stelle schon vorweggenommen werden, dass nicht alle 2G/3G-Mechanismen nahtlos in ein kommerzielles Internet integriert werden können, weil einige Lösungsansätze auf leitungsvermittelnden Architekturprinzipien beruhen. Weiter sind die im „historischen“ Internet verfügbaren, d.h. auf dem paketvermittelnden Internet aufsetzenden Mechanismen, Strategien und Konzepte nicht für ein kommerzielles Mobiles Internet konzipiert worden.

Im weiteren Verlauf dieses Kapitels wird zunächst ein Überblick über Mobilität im Kontext dieser Arbeit gegeben. Dann werden die Internet-spezifischen, rein technischen Mechanismen vorgestellt, welche für den weiteren Verlauf dieser Arbeit von elementarer Bedeutung sind. Im Anschluss daran werden einige auserwählte Konzepte aus der 2G/3G-Umgebung vorgestellt. Und zum Schluss wird ein Überblick über ausgewählte Vorschläge zur Preisbildung im Internet gegeben.

4.2 Mobilitätsaspekte

Mobilität in einem Kommunikationsnetz beinhaltet drei Kernaufgaben, die von der Mobilitätsverwaltung erbracht werden müssen. Zunächst müssen die mobilen Instanzen erreichbar gemacht werden. Hierzu gehört vor allem das Ermitteln der aktuellen Adresse des mobilen Teilnehmers. Außerdem muss einem mobilen Teilnehmer an jedem beliebigen Ort Zugang zu den Telekommunikationsdiensten ermöglicht werden. Schließlich müssen Netzzugangswchsel von einem Kommunikationsnetz erkannt werden und die erforderlichen Mechanismen initiiert werden, damit aktuelle Verbindungen während dieses Wechsels aufrechterhalten werden. Diese so genannte Verbindungsweitergabe (Handover) sollte in einer Weise vollzogen werden, dass sie für den Teilnehmer nicht wahrgenommen werden.

Um nun in einem kommerziellen Internet diese Funktionen zu erfüllen, muss zunächst der Begriff Mobilität genauer spezifiziert werden. Hierzu werden drei verschiedene Ausprägungen der Mobilität im Folgenden näher erläutert.

4.2.1 Endgerätemobilität

Aktuelle Kommunikationsnetze, welche Mobilität unterstützen, beschränken sich in der Regel auf Endgerätemobilität (Terminal Mobility) [120]. Dieser Begriff steht für Endsysteme, die in der Lage sind, sich an einem Netz zu registrieren, um dort eine Kommunikation aufnehmen

bzw. eine existierende Kommunikation fortsetzen zu können. Das Endsystem wird fest einem Benutzer zugeordnet und die für diesen Benutzer getroffenen Vereinbarungen bezüglich der konkreten Dienstleistung werden fest auf dieses Endsystem abgebildet. Dem Endgerät wird erlaubt, sich über die Grenzen einer Netzelement hinweg zu bewegen. Die Identität des Endgerätes bleibt hierbei unverändert. Generell ist festzuhalten, dass bei dieser Art von Mobilität zwischen Netzen innerhalb eines administrativen Bereiches und zwischen verschiedenen administrativen Bereichen zu unterscheiden ist. Dies stellt aus Sicht des Benutzers ggf. keinen wahrnehmbaren Unterschied dar, ist jedoch aus operativer bzw. technischer und administrativer Sicht völlig unterschiedlich. Auf diese Aspekte wird im weiteren Verlauf dieser Arbeit noch näher eingegangen.

4.2.2 Benutzermobilität

Das starre Koppeln eines Benutzers an ein Endgerät, wie bei der Endgerätemobilität implementiert, bringt einige Einschränkungen mit sich. Ein Benutzer ist permanent darauf angewiesen, „sein“ Endgerät entsprechend bereit zu halten, da ohne dieses kein Netzzugang erlaubt wird. Im Rahmen der Benutzermobilität (User Mobility) erfolgt eine Entkopplung dieser starren Beziehung zwischen Gerät und Benutzer [81]. Das Abbilden eines Benutzers auf sein aktuell verwendetes Endgerät erfolgt dynamisch während einer Authentisierungsphase. Hier werden dann die Parameter, die z.B. ein Benutzer im Rahmen eines Vertrages mit einem Netzbetreiber vereinbart hat, auf das aktuell verwendete Gerät abgebildet. Diese Zuordnung ist dann semi-statisch und kann jederzeit aufgehoben werden um z.B. einem anderen Benutzer zeitweilig ein Endgerät zu überlassen.

4.2.3 Dienstmobilität

Dienstmobilität definiert die Fähigkeit des Netzes, Endbenutzern, unabhängig von ihrem Aufenthaltsort, abonnierte Dienste zur Verfügung zu stellen. Konkret bedeutet dies, dass ein Benutzer immer auf „seine“ Dienste zugreifen kann, ungeachtet von welchem Ort aus ein Dienst genutzt wird [33]. Es soll in diesem Zusammenhang ausdrücklich erwähnt werden, dass der Dienst hier aus Sicht des Benutzers logisch gleich bleibt, wegen der Mobilität des Benutzers jedoch permanent von einer anderen Instanz im Netz erbracht werden kann. Dienstmobilität liegt beispielsweise dann vor, wenn ein mobiler Benutzer von jedem Punkt der Welt Zugriff auf die eigenen E-Mails hat, jedoch ortsbezogen zum versenden der Mails verschiedene Mail-relays verwendet, welche aus netztopologischer Sicht basierend auf dem aktuellen Aufenthaltsort geeigneterweise den Dienst erbringen.

4.2.4 Definition und Begriffe

Bevor nun ein detaillierter Überblick über die aktuellen Strategien, Konzepte und Mechanismen gegeben wird, werden zunächst einige relevante Begriffe definiert.

4.2.4.1 Auswärtiges Netz und Heimatnetz

Der Netzbetreiber, der ein vertragliches Verhältnis mit dem Endkunden unterhält, ist der Betreiber des Heimatnetzes. Alle anderen Betreiber, die in den Bereitstellungsprozess eines Dienstes eingebunden sind, werden als auswärtige Anbieter bezeichnet. Somit sind Heimatnetzbetreiber und Betreiber des auswärtigen Netzes Rollen, die ein jeder Betreiber in einem speziellen Szenario einnehmen kann. Das Netzelement, das die jeweilige Betreiberinstanz repräsentiert, ist der AAA-Server, der analog zur bereits erwähnten Rollenverteilung die Rolle eines AAA.f (AAA-Server im auswärtigen Netz) bzw. eines AAA.h (AAA-Server im Heimatnetz) einnimmt [50]. Somit unterstützt ein AAA-Server stets die Funktionalität eines AAA.h-Servers und des AAA.f-Servers.

4.2.4.2 Dienst

Ein Dienst ist eine Leistung, die von einem Dienstanbieter erbracht wird. Im Gegensatz zu Gütern sind Dienste nicht beweglich und Dienstherstellung (Produktion) und Dienstkonsum erfolgen zeitgleich. Dienste können sehr vielfältig sein. Zum Beispiel kann eine zweistündige Videokonferenz, die Audio, Video, ein gemeinsames Zeichenbrett (Shared Whiteboard) und einer garantierten Leistung von 1000000 Paketen innerhalb der entsprechenden Zeit mit einer vereinbarten Priorität eine Beschreibung für einen Dienst sein.

4.2.4.3 Sitzung

Eine Sitzung (Session) entspricht der Bereitstellung eines Dienstes in einem bestimmten Zeitraum. Beginn und Ende der Sitzung ergeben sich aus der detaillierten Beschreibung des Dienstes. Beispiele für eine Sitzung sind das Herunterladen einer bestimmten Datei, eine Audio/Videokonferenz oder aber das Erfassen bestimmter Nutzungsdaten. Auch hier kann eine Verkettung verschiedener Transaktionen als Sitzung bezeichnet werden. Eine Sitzung kann somit aus verschiedenen Untersitzungen oder Transaktionen bestehen und durchläuft folgende drei Phasen: Sitzungsaufbau, Sitzungslaufzeit und Sitzungsabbau. Eine Sitzung ist durch eine Sitzungs-ID eindeutig identifiziert.

4.2.4.4 Ober- und Untersitzung

Eine Sitzung kann einer Ober- oder Untersitzung entsprechen. Man spricht von einer Obersitzung, wenn sie aus einer oder mehreren Untersitzungen besteht, welche Bestandteil der Obersitzung sind. Alle Sitzungen bilden zusammen einen Sitzungsgraphen. Im Falle eines hierarchischen Aufbaus lässt sich ein Sitzungsbaum darstellen, der auf oberster Hierarchieebene mit der Obersitzung beginnt. In bestimmten Fällen ist es vorteilhaft, die Identifikation einer Obersitzung von der Identifikation einer Untersitzung abzuleiten. Aus Gründen der Vertraulichkeit sollte jedoch das detaillierte Wissen über eine Untersitzung nicht dazu führen, dass die Details einer Obersitzung ebenso bekannt sind.

4.2.4.5 Peer-Sitzung

Eine Peer-Sitzung ist eine Sitzung, welche auf derselben hierarchischen Ebene agiert, wie eine weitere Sitzung. Diese Peer-Sitzungen können miteinander interagieren.

4.2.4.6 Transaktion

Eine Transaktion besteht aus Anforderungs- und Antwortnachrichten, die zwischen verschiedenen Instanzen ausgetauscht werden. Zwischen diesen beiden Nachrichten erfolgen in der Regel weitere Verarbeitungsprozesse, welche in der Regel nicht in Echtzeit abgearbeitet werden und sehr komplex sein können. Transaktionen unterscheiden sich von den Sitzungen darin, dass kein kontinuierlicher Datenfluss stattfinden muss. Beispiele für Transaktionen sind z.B. die Autorisierung eines Dienstes oder das Kaufen eines Buches. Wird das Buch in einem Online-Shop gekauft, wäre die Transaktion nach Bezahlung durch den Kunden und Übergabe des Buches abgeschlossen. Somit können Transaktionen Teil einer Sitzung sein. Eine Authentisierungs- und/oder Autorisierungstransaktion wird oft dazu verwendet, um vor Beginn einer initiierten Sitzung zu bestimmen, ob diese Sitzung rechtmäßig ablaufen kann.

4.2.4.7 Nutzungsdatenerfassung

Als Nutzungsdatenerfassung (Metering) bezeichnet man die Rohdatenerfassung von Daten für einen bestimmten Zweck [50]. Im Rahmen dieser Arbeit werden die Nutzungsdaten ausschließlich im Kontext einer späteren Abrechnung behandelt.

4.2.4.8 Aufbereitung der Nutzungsdaten

Die Aufbereitung der Nutzungsdaten (Accounting) bezeichnet das Sammeln und Aufarbeiten der aus der Nutzungsdatenerfassung gewonnenen Rohdaten in ein für den nachgelagerten Prozess erforderliches Format [50]. Hier werden in der Regel Nutzungsberichte (Accounting Records) erzeugt, welche dann die für die Nutzungsdatenerfassung und -verarbeitung erforderlichen Informationen entsprechend weiterverarbeitet werden.

4.2.4.9 Accounting-Sitzung

Eine Accounting-Sitzung besteht aus einer Folge von Accounting Records, die als Verkettung von einzelnen Nachrichten oder als Verkettung von mehreren Accounting-Transaktionen bezeichnet werden können.

4.2.4.10 Charging

Charging bezeichnet die Zusammenführung der aufbereiteten Nutzungsdaten mit den benutzerspezifischen Preisinformationen, welche in der Regel aus dem Vertrag zwischen Betreiber und Benutzer hervorgehen [77].

4.2.4.11 Billing

Billing bezeichnet den Prozess, der dem Charging nachgelagert ist und letztendlich zur Rechnungserstellung kommt [77].

4.2.4.12 Auditierung

Auditierung beschreibt den Prozess, der die verfügbaren Daten dahingehend untersucht, dass gewährleistet werden kann, dass ein vereinbarter Dienst entsprechend bereitgestellt wurde und die Leistungsparameter der Übereinkunft bzw. des Vertrages entsprechend bereitgestellt wurden [50]. Auditierung sollte unabhängig von einer Instanz erbracht werden, welche am Dienstbereitstellungsprozess beteiligt ist.

4.2.4.13 Auditierungs-Transaktion

Auditierungs-Transaktionen beschreiben einen Austausch von Nachrichten, die im Rahmen des Auditierungsprozesses benötigt werden. Dies beinhaltet das Zwischenspeichern von Auditierungsinformationen.

4.2.4.14 AAA-Transaktion

AAA-Nachrichten (vergl. Kapitel 4.3) und AAA-Transaktionen können als eine Art Signalisierung für die Kontrolle einer Sitzung bezeichnet werden. AAA-Transaktionen müssen eindeutig identifizierbar sein. Zwischen folgenden AAA-Transaktionen ist zu unterscheiden:

- **Authentisierungs-Transaktion:**
Eine Authentisierungs-Transaktion beschreibt den Austausch von Anforderung/Antwort (Request/Response) Nachrichten, um die Authentisierung durchzuführen.
- **Autorisierungs-Transaktion:**
Eine Autorisierungs-Transaktion beschreibt den Austausch von Anforderung/Antwort Nachrichten um die Autorisierung durchzuführen. Da eine Autorisierung in der Regel an eine bestimmte Zeitdauer gebunden ist, kann während einer Session eine Erneuerung dieser Autorisierung in Form einer Re-Autorisierungs-Transaktion notwendig sein
- **Accounting-Transaktion:**
Eine Accounting-Transaktion beschreibt den Austausch von Anforderung/Antwort Nachrichten um das Accounting durchzuführen. Accounting kann in Form von

Accounting-Transaktionen durchgeführt werden, die über den detaillierten Ressourcenverbrauch berichten. Accounting-Transaktionen können entweder permanent während der gesamten Laufzeit einer Sitzung auftreten, oder aber nur während des Beginn und/oder Ende einer Sitzung.

4.2.4.15 Binding

Binding beschreibt den Prozess, der eine Verbindung zwischen verschiedenen Nachrichten, Transaktionen und Sitzungen herstellt, die logisch zu einem Prozess gehören. Im Rahmen dieser Arbeit liegt hierbei der Schwerpunkt auf dem Accounting- oder Auditierungs-Prozess.

4.2.4.16 Accounting-Bericht

Ein Accounting-Bericht (Accounting Record) beinhaltet alle relevanten Informationen, die während eines Accounting-Prozesses generiert werden.

4.2.4.17 Charging-Bericht

Ein Charging-Bericht (Charging Record) beinhaltet alle relevanten Informationen, die während eines Charging-Prozesses generiert werden

4.2.4.18 Billing-Bericht

Ein Billing-Bericht (Billing Record) beinhaltet alle relevanten Informationen, die während eines Billing-Prozesses generiert werden

4.3 Authentisierung, Autorisierung und Accounting

In der Internet-Gemeinde wurde erkannt, dass bestimmte Mechanismen, welche vor allem zur Kommerzialisierung des Netzes erforderlich sind, nicht von Beginn an bei der Spezifikation berücksichtigt wurden. Einer diese offensichtlichen Mängel sind Authentisierung, Autorisierung und Accounting. Als Ergebnis dieser Erkenntnis wurde unter dem Akronym AAA innerhalb der IETF und unter dem Akronym AAAArch innerhalb der IRTF je eine Standardisierungsgruppe gebildet, welche prozedural und architektonisch Authentisierung, Autorisierung und Accounting standardisiert ([78] und [25]).

Generell kann man festhalten, dass die Funktionalitäten, welche AAA bietet, die Grenze definiert, an der die betriebswirtschaftlichen Konzepte eines Netzbetreibers auf die technologischen Konzepte eines Netzbetreibers stoßen und dass ohne diese standardisierte, logische Schnittstelle keinerlei kommerziellen Mechanismen im Internet skalierbar implementiert werden können. Somit ist eine leistungsfähige AAA-Architektur eine notwendige Voraussetzung um Wertschöpfungsketten auf höherer Ebene implementieren zu können.

Im Rahmen dieser Arbeit wird die Thematik Authentisierung, Autorisierung und Accounting ausschließlich im Sinne der Aktivitäten innerhalb der IETF und IRTF verwendet, welche nun in diesem Kapitel vorgestellt werden. Bevor die detaillierte AAA-Architektur vorgestellt wird, soll zunächst eine Übersicht über die prinzipiellen AAA-Konzepte und einige mit AAA in engem Zusammenhang stehenden Mechanismen gegeben werden.

4.3.1 Authentisierung

Authentisierung ist ein Mechanismus, der die Identität eines Benutzers zweifelsfrei feststellen soll. Somit stellt eine grundlegende Rolle in der Netzwerksicherheit die Authentizität eines Benutzers dar. Nur wenn die Authentizität eines Benutzers sichergestellt ist, können Maßnahmen wie Autorisierung und Accounting entsprechend eingesetzt werden. Als die Authentisierung unterstützende Maßnahme können auch Eigenschaften des Telefonnetzes ausgenutzt werden, z.B. der automatische Rückruf (Call Back) oder die Rufnummernübermittlung als

spezieller Dienst im Telefonnetz. Authentisierung basiert oftmals auf einer Überprüfung eines Dialogs, bei dem ein sicheres Kennwort überprüft wird. Wesentliche Prinzipien der Authentisierung werden in den nächsten Abschnitten kurz vorgestellt.

4.3.1.1 Kennwortübertragung

Die Authentisierung nach einzelnen Benutzernamen und zugehörigen Kennwörtern stellt die Mindestanforderung für benutzerbezogenes Autorisieren und Abrechnen dar. Prinzipiell kann eine Kennwortübertragung sowohl verschlüsselt als auch unverschlüsselt erfolgen. Ein Problem bei der unverschlüsselten Übertragung von Kennwörtern ist, dass durch Mithören auf den Übertragungsleitungen das Kennwort von Unbefugten gelesen und missbraucht werden kann. Das Problem der unsicheren Übertragungswege bzw. die unverschlüsselte Übertragung von Kennwörtern kann auf verschiedenste Weise gelöst werden, ist jedoch immer mit erhöhtem Aufwand in der Authentisierungsphase verbunden. Die fehlende bzw. nichteinheitliche Standardisierung ist hierbei ein Problem. Eine einfache Möglichkeit ist ein so genanntes Frage/Antwort (Challenge/Response) System. Hier wird eine Kombination aus einer zufällig generierten Zeichenfolge und deren Verschlüsselung genutzt. Jeder Benutzer muss hierzu auf einem Authentisierungsserver einen Schlüssel (Kennwort) in Klartext hinterlegen. An den Benutzer wird in der Authentisierungsphase eine zufällig generierte Zeichenfolge (Challenge) gesendet. Diese Zeichenfolge wird vom Empfänger mit Hilfe des Kennworts verschlüsselt und zurückgesendet. Die verschlüsselte Zeichenfolge (Response) wird mit Hilfe des hinterlegtem Schlüssels und der ursprünglich gesendeten Zeichenfolge verglichen.

Aus Sicht der Sicherheit bereitet bei diesem Verfahren die Schlüsselaufbewahrung Probleme, da dies meist in Klartext erfolgt. Verfahren mit Einmalpasswörtern umgehen diese Schlüsselproblematik. Bei diesen Verfahren ist es nicht notwendig, dass ein Kennwort in Klartext auf dem Authentisierungsserver gespeichert wird, oder dass es bei einer Kennwortänderung in Klartext übertragen werden muss. Mit dem Einsatz von Einmalpasswörtern kann zwar das Abhören der Passwörter verhindert werden, nicht jedoch die freiwillige oder fahrlässige Weitergabe an Unbefugte.

4.3.1.2 Authentisierung durch Tokensysteme

Die bisher beschriebenen Verfahren zur Authentisierung gehen davon aus, dass der Benutzer ein Kennwort (Credential) besitzt. Das Problem hierbei besteht darin, dass nicht immer zweifelsfrei sichergestellt werden kann, ob und wie ein Kennwort ggf. ungewollt preisgegeben wurde. Eine wesentlich sicherere Zugriffsmethode wird mit Hilfe von so genannten Tokensystemen erreicht. Bei den meisten dieser Verfahren benötigt man zur Authentisierung ein Kennwort und ein Token. Tokensysteme bestehen aus einem Authentisierungsserver und einem Token für die Benutzer. Zur Authentisierung muss man im Besitz eines Kennworts (z.B. Personal Identification Number – kurz PIN) und eines Token (z.B. Crypto-Card, Smart-Card, etc.) sein. Dem Benutzer wird in der Authentisierungsphase eine Zeichenfolge (Challenge) zugesendet, die Antwort (Response) kann nur mit Hilfe des Tokens und der PIN berechnet werden. Replay-Attacken werden durch zufällig erzeugte Zeichenfolgen im Authentisierungssystem verhindert. Bei Verfahren mit Zeitsynchronisation kann auf ein Challenge/Response-Verfahren verzichtet werden, allerdings eröffnet sich während eines kurzen Zeitfensters (ca. 60s) die Möglichkeit von Replay-Attacken. Sollte ein Token verloren gehen, ist durch die PIN noch gewährleistet, dass die Sicherheit nicht automatisch durch den Verlust des Token allein gefährdet ist.

4.3.1.3 Passwort-Authentisierungsprotokoll

Die Authentisierung nach dem Passwort-Authentisierungsprotokoll (Password Authentication Protokoll – PAP) entspricht einer einfachen Authentisierung mit Benutzername und Kenn-

wort, wobei das Kennwort im Klartext übertragen wird. Die Kommunikationsgegenstelle sendet wiederholt Pakete, welche den Benutzernamen und das Kennwort enthalten, bis entweder eine Empfangsbestätigung empfangen, oder die Verbindung von der Gegenseite abgebrochen wird. Auf die Probleme bei dieser Art der Authentisierung wurde bereits in Kapitel unter 4.3.1.1 hingewiesen.

4.3.1.4 Challenge-Handshake-Authentisierungsprotokoll

Beim Authentisierungsverfahren nach dem Challenge-Handshake-Authentisierungsprotokoll (CHAP) erfolgt eine verschlüsselte Übertragung des Kennworts, welche auf nicht umkehrbaren Hash-Funktionen basiert, die auf eine Kombination der Zufallsfolge (Challenge) und des Kennworts angewendet werden. Die so verschlüsselte Antwort wird vom entfernten Rechner an den Authentifizierungsserver zurückgesendet (Response), welcher nach Anwendung der gleichen mathematischen Operation auf die Eingangswerte (Challenge, Kennwort) die Ergebnisse miteinander vergleicht. In der Regel wird als Hash-Funktion hauptsächlich der so genannte MD5-Algorithmus verwendet [113], kann aber ggf. im Einzelfall ausgehandelt werden.

4.3.2 Autorisierung

Unter Autorisierung versteht man die vom jeweiligen Benutzer abhängige dynamische Verwaltung eines Ressourcenverbrauches. In der Regel folgt hieraus eine benutzerspezifische Einschränkung, die sich nach vielfältigen Regeln richten kann. Eine solche Einschränkung wird z.B. über ein benutzerspezifisches Profil kommuniziert, welches in einer allgemeinen Benutzerdatenbank gespeichert wird. Mit der Autorisierung werden im Wesentlichen zwei Ziele verfolgt: zum einen kann der Zugriff auf Ressourcen im Netzwerk eingeschränkt werden, und zum anderen kann die Komplexität für bestimmte Benutzergruppen geeignet verwaltet werden. Mit dem Konzept der Gruppierung von Benutzern zu Benutzergruppen und der anschließenden Definition von Gruppenprofilen wird die Zusammenfassung ganzer Benutzerklassen ermöglicht, was zu einer Verringerung des erforderlichen Verwaltungsaufwandes führt.

4.3.3 Aufbereitung der Nutzungsdaten

Das Aufbereiten der Nutzungsdaten zum Zweck der späteren Rechnungserstellung (Accounting) ist ein komplexer Prozess, bei dem erfasste Rohdaten über die Ressourcennutzung einem Benutzer zugeordnet werden. Der Gesamtprozess muss sicher und fehlerfrei funktionieren, da in der Regel eine hohe Sensibilität der Verbraucher gegen Fehler im Markt zu beobachten ist.

4.3.4 Charging

Charging für Kommunikationsdienste ist im Internet traditionell unzureichend betrachtet worden. Dies liegt hauptsächlich an der ursprünglich nichtkommerziellen Nutzung des Internets, welches durch zentrale Finanzierungsmodelle aufgebaut wurde. In jüngster Zeit wurden Mechanismen implementiert, welche eine volumenbasierte Abrechnung eines Benutzers im Festnetz ermöglicht. Bedingt durch die Kommerzialisierung des zukünftigen Mobilens Internets werden Protokolle und Ansätze erforderlich, die in der Lage sind, Überlastsituationen über Nachfragesteuerung zu vermeiden und somit dieses Mobile Internet kommerziell selbsttragend zu finanzieren. Zur Lösung dieser Probleme sind technische Hilfsmittel und Protokolle für Charging erforderlich, die eine Ermittlung der angefallenen Beträge erlauben, welche dem Verbraucher der Ressourcen in Rechnung gestellt werden müssen.

4.3.5 Auditierung

Der Begriff Auditierung (Auditing) wird sehr weit gefächert verwendet. Im Rahmen dieser Arbeit bezeichnet man Auditierung als eine unabhängige Untersuchung der Nutzungsdaten-

bereichte, oder anderer erfasster Nutzungsdaten, welche je nach Zweck der Auditierung als Nachweis für eine Dienstleistung bzw. einen Dienstverbrauch entsprechend der getroffenen Vereinbarungen der Vertragsparteien dienen. Auditierung ist besonders im Zusammenhang mit finanziellen Transaktionen eine sehr wichtige Funktion, welche als Kerngrundlage für jegliche Kommerzialisierung in einem entsprechenden Rechts/Wirtschaftssystem betrachtet werden kann, da hier sichergestellt werden muss, dass eine Aktion nach deren Abschluss nicht von einem Vertragspartner zu Unrecht angefochten werden kann. Diese wird in der Fachliteratur als Nichtabstreitbarkeit (non-repudiation) bezeichnet. Die ISO hat in diesem Zusammenhang ein Modell entwickelt, welches in [58], [59] und [60] beschrieben ist.

4.3.6 Nutzungsdatenerfassung

Netzbetreiber sind sehr daran interessiert, Detailinformationen über die Beschaffenheit und den exakten Weg von Datenpaketen durch ihr Netz zu bekommen. Diese Verkehrsdaten sind notwendig, um eventuelle Abnormalitäten schnell erfassen zu können, um Daten für Preisabrechnung zu generieren und um für die Kapazitätsplanung des Netzes relevante Informationen zu generieren. Die einfachste Möglichkeit ist es, Pakete, die einen bestimmten Messpunkt passieren, entsprechend zu protokollieren. Da oftmals nur die Informationen notwendig sind, die im Kopf des IP-Paketes stehen, genügt es, die Protokolle auf Basis von Schicht 3 zu erstellen. Im Anschluss an diese Datenerfassung (Metering) wird der in Kapitel 4.3.3 beschriebene Prozess gestartet, der diese Daten geeignet analysiert. Der Vorteil von diesem Vorgehen ist, dass die reine Protokollierung mit angemessenem Aufwand sehr leistungsfähig abgewickelt werden kann. Zusätzlich ist es möglich, sehr feingranulare Information zu generieren, da jedes Paket einzeln erfasst wird. So kann man so das Verhalten zwischen zwei verschiedenen Messpunkten sehr präzise vergleichen. Ein Beispiel für ein solches Vorgehen ist in [64] beschrieben.

Ein Nachteil dieses Vorgehens ist, dass die Datenmenge, setzt man diese Methode in größerem Stil ein, sehr groß wird. Dies erschwert ein effektives Weiterverarbeiten der Daten. Ein weiterer Ansatz für Verkehrsmessungen in einem paketorientierten Netz ist das Erfassen so genannter *Flows* anstatt der einzelnen Pakete. Der Begriff eines Flows wurde innerhalb der IETF in [16] eingeführt und kann auf verschiedene Arten interpretiert werden. Hier werden Pakete einem Flow zugeordnet wenn sie folgendem 5-Tupel entsprechen (Protokoll, IP-Quell und -Zieladresse, Quell- und Ziel-Portnummer). Ein Flow ist unidirektional definiert. Tritt innerhalb eines definierten Zeitintervalls kein weiteres Paket auf, das einem bestimmten Flow zugeordnet werden kann, so wird dieser Flow als beendet definiert. Dieses Konzept wird z.B. von den Produkten der Firma CISCO [28] eingesetzt. Ein weiteres etabliertes System ist das in der IETF entwickelte RMON-System [121]. In Kapitel 4.4.5 wird auf eine weitere etablierte Methode detaillierter eingegangen.

4.4 IETF AAA-Architektur und angrenzende Standards

Die IETF AAA-Arbeitsgruppe hat sich zum Ziel gesetzt, Anforderungen an AAA im Zusammenhang mit dem Netzzugang zu definieren. Basierend auf diesen Anforderungen definiert innerhalb der IETF die AAAArch-Arbeitsgruppe eine generische Architektur [78] und innerhalb der IETF die AAA-Arbeitsgruppe die entsprechenden Protokolle.

So wurde von der IETF das Diameter-Protokoll [25] standardisiert, welches in Kapitel 4.4.1 noch näher vorgestellt wird, und von der IETF eine generische Architektur entworfen, welche den definierten Anforderungen genügt. Randbedingungen waren die Unterstützung von IPv6, die Unterstützung von Accounting unabhängig der Netztechnologie und eine Kompatibilität zum aktuell für diesen Zweck eingesetzten Radius-Protokoll [111]. Diameter wurde mit der Vorgabe entwickelt, aufgetretene Nachteile von Radius zu verbessern. Dies ist vor allem ein

gesicherter Transport zwischen den Komponenten. Anstatt UDP wird bei Diameter SCTP [123] zur Verwendung vorgeschlagen. Weiter wurden bei Diameter Sicherheitsmechanismen nach [73] implementiert und es ist erstmals möglich, Nachrichten aus dem Netz zu initiieren. Dies ist relevant, wenn z.B. ein Prepaid-Kunde sein verbleibendes Restguthaben aufbraucht, das Netz dies erkennt und dieses Ereignis zu den entsprechenden Netzknoten kommuniziert.

Die generische AAA-Architektur ist in Abbildung 28 dargestellt. Kernkonzept dieser Architektur ist die Annahme, dass jeder Benutzer bei einem Netzbetreiber „zu Hause“ ist und dort unter Vertrag steht. Somit wird das Internet aus Sicht eines Benutzers logisch in zwei Teile geteilt, das Heimatnetz und den Rest. Über Roaming-Abkommen kommunizieren AAA-Server, welche je nach Situation die Rolle eines AAA-Servers im Heimatnetz (AAA.h) und die Rolle eines AAA-Servers in einem auswärtigen Netz (AAA.f) einnehmen können. Ein Benutzer muss sich in jedem Fall beim lokalen AAA-Server anmelden, welcher dann im Auftrag dieses Benutzers beim entsprechenden AAA.h ermittelt und eine Freigabe beim Heimatnetz beantragt. Diese Freigabe, oder auch Registrierung genannt, ist grundsätzlich ein an definiertes Zeitintervall gebunden.

Obwohl AAA ursprünglich nicht zur Mobilitätsverwaltung vorgesehen wurde, ist es prinzipiell geeignet, die IP-Mobilitätsverwaltung in einem kommerzialisiertem Netz zu unterstützen. Mit dieser Architektur, bei der äquivalente Systemkomponenten in der 2G-Systemarchitektur vorzufinden sind (vergleiche Kapitel 2.2.1.1), hat die IETF einen ersten Schritt in Richtung kommerzialisiertes Internet vollzogen.

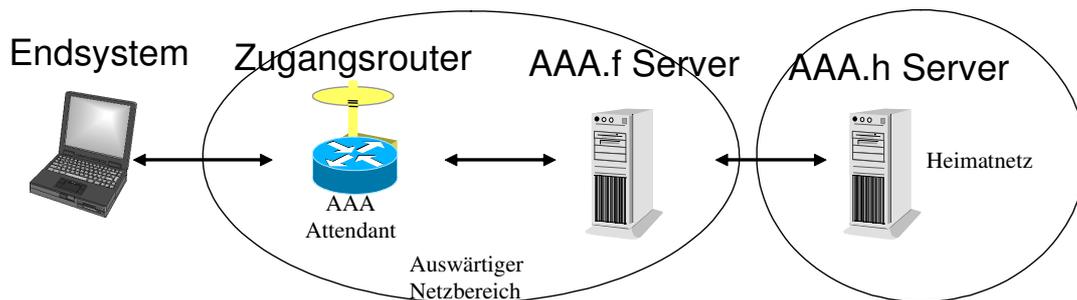


Abbildung 28: Generische AAA-Architektur

Das Konzept eines „heimatlosen Benutzers“ wurde hier nicht verwirklicht, was vor allem nicht im Interesse der Netzbetreiber ist. Netzbetreiber setzen hier auf ein Konzept, welches einen Benutzer vertraglich an einen Betreiber bindet. Dies ermöglicht, dass ein Kunde ein Vertrauensverhältnis zu einem Betreiber aufbauen kann, was in einer kommerziellen Umgebung aus Sicht eines Netzbetreibers von zentraler Bedeutung ist. Ein heimatloser Benutzer würde hier direkt mit mehreren Betreibern einen Vertrag abschließen müssen und würde, je nach Aufenthaltsort/Bewegungsverhalten, zahlreiche vertragliche Beziehungen dynamisch eingehen müssen. In den folgenden Abschnitten sollen nun die detaillierten Mechanismen und Protokolle vorgestellt werden, welche dieses grundlegende Szenario unterstützen.

4.4.1 Diameter

Diameter ist ein Protokoll, welches auf seinem Vorgänger Radius aufbauend die grundlegenden Funktionen der AAA-Architektur unterstützt. Diameter verwendet im Vergleich zu Radius leicht modifizierte bzw. erweiterte Protokoll-PDUs, was die gemeinsame Verwendung von Diameter und Radius vereinfacht. Diameter berücksichtigt die Anforderungen von verschiedenen IETF-Arbeitsgruppen und anderen Standardisierungsgremien. Dies sind:

- IP Routing for Wireless/Mobile Hosts (MOBILEIP)

- Network Access Server Requirements (NASREQ)
- Roaming Operations (ROAMOPS)
- Telecommunications Industry Association (TIA)

4.4.1.1 Diameter-Protokollarchitektur

Diameter ist ein so genanntes Peerto-Peer-Protokoll. Hierbei handelt es sich um ein Basisprotokoll, das sehr leicht erweitert werden kann, um flexibel AAA-Dienste mit heterogenen Netzzugangstechnologien zusammenbinden zu können. Die Diameter-Gesamtarchitektur ist modular aufgebaut.

Kernziel von Diameter ist die Koordination von Accounting und Netzzugang. Wie in Abbildung 29 gezeigt, ist Diameter in zwei Teile geteilt. Dies sind das Diameter-Basisprotokoll (Diameter Base Protocol) und eine Vielzahl von Diameter-Anwendungen (Diameter Applications).

Das Diameter-Basisprotokoll [25] stellt ein Kerngerüst an Funktionalität zur Verfügung, welche ein AAA-Protokoll nach [7] benötigt. Das Basisprotokoll kann entweder nur für elementares Accounting verwendet werden, oder aber es kann sehr flexibel mit spezifischen Diameter-Anwendungen zusammenarbeiten. Diese Diameter-Anwendungen definieren sehr anwendungsspezifische Dateneinheiten und Funktionen, welche durch Diameter eine große Flexibilität erlangen. Neben den Diameter-Anwendungen sind auch beim Basisprotokoll Erweiterungsmöglichkeiten vorgesehen.

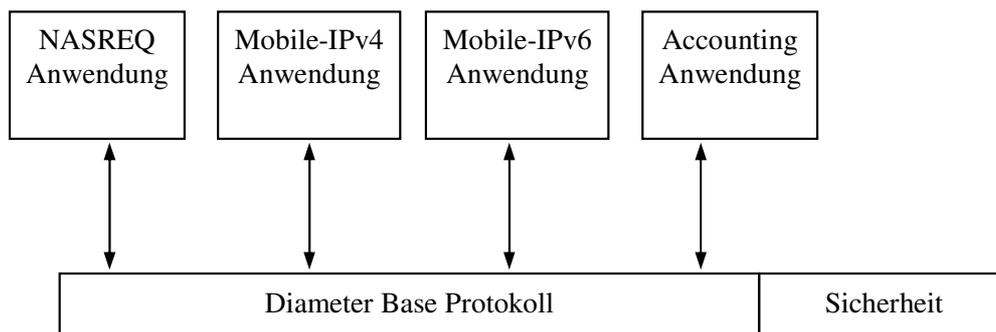


Abbildung 29: Diameter-Protokollarchitektur

4.4.1.2 Diameter-Basisprotokoll

Das Diameter-Basisprotokoll bietet die gesamte Funktionalität an, die erforderlich ist, um elementare AAA-Dienste anbieten zu können. Folgende Funktionen werden erfüllt:

- Transport der Dienstprimitive, der so genannten AVPs (Attribute Value Pairs)
- Vereinbarung der geeigneten Kommunikationsparameter (Capabilities Negotiation)
- Fehlermanagement
- Erweiterbarkeit durch hinzufügen neuer Befehle und AVPs wie in [7] gefordert.
- Basisdienste wie das Verwalten von Sitzungen

Jegliche Daten werden in Form eines AVP vom Protokoll transportiert. Diese AVPs werden entweder vom Basisprotokoll selbst, oder aber von Anwendungen verwendet. Einige dieser AVPs werden vom Protokoll selbst benötigt. AVPs werden im Diameter-Protokoll eingesetzt um folgende Funktionen zu erfüllen:

- Transport von Authentisierungsinformationen. Diese Informationen dienen dazu, einen Benutzer an einem Diameter-Server anzumelden.

- Transport von dienstspezifischen Autorisierungsinformationen die den entsprechenden Instanzen die notwendigen Informationen bereitstellen, um zu entscheiden, ob eine Ressourcenfreigabe erfolgen kann.
- Austausch von Informationen, die den Ressourcenverbrauch belegen. Dies kann sowohl zu Accounting, als auch zur Kapazitätsplanung verwendet werden.
- Weiterleiten von Diameter-Nachrichten innerhalb eines Diameter-Server-Verbundes.

Das Diameter-Basisprotokoll bevorzugt SCTP als Transportprotokoll (siehe hierzu auch Kapitel 2.1.3.3). Der Grund hierfür ist die Anforderung an das Protokoll, sowohl echtzeitfähig, als auch verlässlich zu sein, was von SCTP geleistet wird. In jedem Fall muss Diameter jedoch auf einer gesicherten Transportverbindung aufsetzen.

4.4.1.3 Diameter-Nachrichtenformat

Diameter-Nachrichten bestehen aus einem Diameter-Kopf, an den sich eine flexible Anzahl von Diameter-AVPs anschließen. Der Diameter-Kopf ist ähnlich aufgebaut, wie der IP-Kopf und ist in Abbildung 30 abgebildet.

Das Versionsfeld identifiziert die aktuelle Version des Protokolls. Hat eine Anforderungsnachricht (*Request*) das R-Feld gesetzt, ist bei der entsprechenden Antwort dieses Feld nicht mehr gesetzt. Das P-Feld signalisiert, ob die Antwort lokal verarbeitet werden muss, oder ob sie ggf. einem anderen AAA-Server weitergesendet werden darf. Das E-Feld signalisiert einen Fehler. Die restlichen Felder (r-Felder) sind für zukünftige Funktionen reserviert. Das Kommando-Feld legt den Typ der Diameter-Nachricht fest und identifiziert den speziellen Diameter-Befehl. Hierbei handelt es sich um ein 24 bit Feld, welches entweder mit Standard-IETF-Befehlen, oder aber mit proprietären Befehlen versehen werden kann. Mittels zentraler Standardisierung der Befehle wird verhindert, dass Inkompatibilitäten durch eine unkoordinierte Verwendung der proprietären Befehle verursacht werden.

Version								Nachrichtlänge	
R	P	E	r	r	r	r	r	Kommando	
Hersteller ID									
Hop-zu-Hop Identifizierung									
Ende-zu-Ende Identifizierung									
AVPs....									

Abbildung 30: Diameter-Kopf

4.4.1.4 Diameter-Dienstprimitive

Nach dem allgemeinen Format des Kopfes folgen die Dienstprimitive, die so genannten AVPs. Diese signalisieren spezifische AAA-Informationen. Neben Authentisierung, Autorisierung und Accounting sind dies Routing Informationen, sicherheitsspezifische Informationen und Informationen, die konfigurationsspezifisch sind. Jedes AVP besteht aus einem AVP-Kopf, welcher in Abbildung 31 dargestellt ist, und einem AVP-Rumpf.

Ein AVP ist ein Informationspaar, welches von der IANA standardisiert wird und aus zwei Teilen besteht. Ein Teil des Paares beschreibt das Attribut, also einen spezifischen Befehl, und die zweite Hälfte des Paares beschreibt den dazugehörigen Wert oder Inhalt. Das Attribut wird eindeutig durch AVP-Code und Vendor-ID-Feld festgelegt. Die ersten 256 AVP-Befehle oder Codes sind auch hier für Rückkompatibilität mit Radius reserviert. Diameterspezifische Codes sind somit größer als 256.

AVP-Kode									
V	M	P	r	r	r	r	r	AVP-Länge	
Hersteller ID (optional)									
Daten									

Abbildung 31: Diameter AVP-Kopf

Die genaue Verwendung von AVPs soll mittels folgendem Beispiel basierend auf Kapitel 2.1.4.1 verdeutlicht werden. Eine Anforderung eines mobilen Endsystems mit gesetztem R-Feld und dem Befehl 260 beinhaltet neben weiteren AVPs die Heimatadresse des Mobilten Knoten, die Adresse des HA, mit dem der MN Verbindung aufnehmen möchte, den Benutzernamen und die Anforderung zur Erzeugung eines abgesicherten Schlüssels, der zur weiteren Kommunikation verwendet wird. Die Antwort des AAA.h-Servers mit dem Befehl 262 beinhaltet dann die angeforderten Informationen in AVPs verpackt, oder ggf. eine negative Antwort.

AVP-Felder beherbergen Information über die Art und Weise, wie das Attribut eines Befehls beim Empfänger behandelt werden soll. Das V-Feld erfordert eine negative Antwort. Das M-Feld zeigt an, ob ein Proxy-Router erforderlich ist. Das P-Feld signalisiert, dass das AVP mit einer digitalen Unterschrift signiert worden ist. Auch hier sind die r-Felder für zukünftige Verwendung reserviert.

4.4.1.5 Diameter-Sitzungs-IDs

Diameter beinhaltet Sitzungs-IDs, um das Binding zwischen verschiedenen AAA-Transaktionen zu gewährleisten. Dies geschieht mit Hilfe so genannter Sitzungs-ID AVPs, welche definiert wurden, um mobile Benutzer zu authentisieren und zu autorisieren. Die Sitzungs-ID ist dann für die gesamte Laufzeit einer Sitzung gültig und endet, wenn dies explizit signalisiert wird.

Das Sitzungs-ID-AVP (AVP Code 263) ist in [25] spezifiziert und identifiziert eindeutig eine Sitzung. Alle Nachrichten, die zu einer spezifischen Sitzung zugeordnet werden können, müssen mit einer eindeutigen und einheitlichen Sitzungs-ID gekennzeichnet werden. Das Format einer Diameter-Sitzungs-ID ist in Abbildung 32 dargestellt.

<client FQDN>[:<port>];<high 32 bits>;<low 32 bits>[:<optional value>]

Abbildung 32: Format einer Diameter-Sitzungs-ID

4.4.1.6 Fehlerbehandlung in Diameter

Man kann zwischen zwei verschiedenen Klassen von Diameter-Fehlermeldungen unterscheiden. Dies sind Protokollfehler und Anwendungsfehler. Protokollfehler deuten auf einen Fehler im Basisprotokoll hin und werden routenabschnittsweise behandelt. Anwendungsfehler deuten auf einen Fehler in der Diameter-Anwendung hin. Sobald eine Nachricht einen Fehler identifiziert hat, wird eine Antwort erwidert, bei der das Fehlerfeld gesetzt ist. Sollte der Fehler nicht auf der Übertragung, sondern in einer Instanz des Diameter Servers erkannt werden, so wird dann direkt sowohl Fehler-Feld, als auch das geeignete AVP zur Kommunikationsgegenstelle gesendet. Tabelle 3 zeigt die verschiedenen Klassen von Fehler-AVPs. Es handelt sich hierbei um 32 bit Felder, welche von der IANA verwaltet werden

Tabelle 3: Diameter Fehler Codes

Code	Klasse	Beschreibung
1xxx	Informell	Zusätzliche Maßnahmen erforderlich um Vorgang abzuschließen
2xxx	Erfolg	Anforderung erfolgreich verarbeitet
3xxx	Protokoll Fehler	Fehler
4xxx	Vorübergehender Fehler	Anforderung fehlgeschlagen, weiterer Versuch folgt
5xxx	Ständiger Fehler	Anforderung endgültig fehlgeschlagen

4.4.1.7 Diameter-Anwendungen

Das Diameter-Protokoll wurde sehr flexibel konzipiert, um als Basisprotokoll für sehr vielfältige Anwendungen eingesetzt werden zu können. Diese Flexibilität zeichnet sich vor allem aus durch:

- Das Definieren neuer AVP-Werte, die anwendungsspezifisch sind und anwendungsspezifische Nachrichten enthalten.
- Das Festlegen neuer AVPs für den Fall dass es keine existierenden AVPs gibt.
- Das Erstellen neuer Authentisierungs- und Autorisierungsanwendungen.
- Das Diameter-Basisprotokoll ist bereits für Aufbereitung der Nutzungsdaten und die damit zusammenhängende Weiterleitung der Berichte ausgelegt. Für Authentisierung und Autorisierung ist in jedem Fall eine spezielle Anwendung erforderlich. Bei einem Diameter-Szenario, das nicht in eines der existierenden Lösungsschemata passt, ist der Entwurf einer neuen Diameter-Anwendung erforderlich.
- Das Erstellen neuer Anwendungen zur Aufbereitung der Nutzungsdaten, da die bereitgestellte Funktionalität ggf. nicht ausreicht, um komplexere Szenarien abzubilden. So wurde z.B. eine Diameter-Anwendung für SIP in [43] vorgestellt.
- Anwendungsbezogene Authentisierungsprozeduren.
- Unterstützung von Authentisierungsmethoden wie z.B., das Extensible Authentication Protocol (EAP) [10].

Durch die Empfehlung jeweils existierende AVP-Attribute und AVPs zu verwenden, wird der gesamte Standardisierungsprozess und die Interoperabilität zwischen verschiedenen Anwendungen vereinfacht. Dies macht Diameter zu einem sehr flexibel einsetzbaren Protokoll. Um sicherzustellen, dass das Diameter-Protokoll sowohl allgemein einsetzbar ist, als auch konkreten Anforderungen genügt, wurde innerhalb der IETF AAA-Arbeitsgruppe der Schwerpunkt auf verschiedene Diameter-Anwendungen gelegt. Dies sind:

- Diameter Mobile IPv4 Application
- Diameter Extensible Authentication Protocol (EAP) Application
- Diameter Network Access Server Application
- Diameter Credit-Control Application
- Diameter Session Initiation Protocol (SIP) Application

Für den weiteren Verlauf dieser Arbeit sind vor allem die drei letztgenannten Anwendungen von Interesse und somit soll auf diese nun näher eingegangen werden.

Diameter NASREQ-Anwendung

Die NASREQ-Anwendung [24] definiert eine Diameter-Anwendung, bei der ein Diameter-Server mittels einem PPP/SLIP-Aufbau den Zugang zu Netzressourcen steuert ([119], [114]). Die Anwendung besteht aus drei Teilen:

- Traditionelle Authentisierung, die sowohl das Passwort-Authentisierungsprotokoll (PAP), als auch das Challenge-Handshake-Authentisierungsprotokoll CHAP unterstützt.
- Erweiterbare Authentisierung, die durch die Unterstützung des EAP gewährleistet wird.
- Autorisierung und Accounting für Netzzugangsserver NAS (Network Access Server).

Diameter Mobile IPv6-Anwendung

Die Diameter Mobile IP-Anwendung [79] erweitert das Diameter-Basisprotokoll. Im Rahmen dieser Erweiterung wird dem Diameter-Server ermöglicht, AAA-Funktionen für Mobile IP-Dienste, die einem mobilen Endsystem zur Verfügung gestellt werden, auszuführen. In Kombination mit den Mechanismen, die die Kommunikation zwischen administrativen Domänen regeln und die das Basisprotokoll unterstützt, ermöglicht diese Anwendungen mobilen Knoten Dienste im administrativen Bereich dritter Betreiber (Foreign Service Provider) zu beziehen. Die Mobile IP-Anwendung definiert darüber hinaus, wie zwischen diesen beiden administrativen Netzbereichen bzw. zwischen Diameter-Servern dieser administrativen Netzbereiche Accounting-Daten ausgetauscht werden. Die Vorteile von Diameter in Verbindung mit Mobile IP-Anwendungen sind:

- Bessere Skalierbarkeit der Sicherheitsbeziehungen.
- Mobilitätsverwaltung zwischen administrativen Bereichsgrenzen eines Betreibers hinweg.
- Dynamische Zuweisung des Heimatagenten.

Um, wie bei Mobile IP vorgesehen, eine Sicherheitsbeziehung zwischen dem mobilen Endsystem und dem Heimatagenten zu unterhalten, ist vorgesehen, dass der Diameter-Server als Schlüsselverteilungszentrum agiert (Key Distribution Centre – KDC). Hier werden dynamisch Registrierungsschlüssel erzeugt und an die mobilen Instanzen verteilt, um die Registrierungsnachrichten für Mobile IP abzusichern. Es ist einem Heimat-AAA-Server erlaubt, einen Vertraulichkeitsbeweis (Shard Secret) zu erzeugen, ohne dass dieses dem auswärtigen AAA-Server bekannt gemacht wird.

Im Einklang mit dem von der IETF ROAMOPS-Arbeitsgruppe definierten Roaming-Modell [5], verwenden die AAA-Server die Benutzeridentität wie in NAI [6] vorgesehen, welches vom Diameter-Basisprotokoll für die Verkehrslenkung der AAA-Nachrichten zwischen administrativen Netzen eingesetzt wird. Dies erlaubt einem mobilen Endsystem eine statische Heimadresse bzw. einen statischen Heimatagenten während der mobilen Sitzung und darüber hinaus beizubehalten. Weiter definiert die Mobile IP-Anwendung den speziellen Nachrichtenaustausch, welcher erforderlich ist, wenn ein sich in einem auswärtigen Netz aufhaltendes Endsystem Dienste anfordert.

4.4.2 Die AAAArch-Arbeitsgruppe der IRTF

Die IETF AAA-Arbeitsgruppe bearbeitet formal dieselben Fragestellungen, wie die IRTF AAAArch-Arbeitsgruppe, welche im Jahre 1999 gegründet wurde und im Herbst 2004 die Arbeit eingestellt hat. Generell ist der Blickwinkel dieser beiden Gruppen völlig unterschiedlich. Während die IETF AAA-Arbeitsgruppe kurzfristige Anforderungen an ein AAA-Protokoll ableitete, auf derer Basis das Diameter-Protokoll entstanden ist, hat sich gezeigt, dass zusätzlich eine generische Architektur für AAA erforderlich ist. Diese wurde von der IRTF AAAArch-Gruppe entworfen und unterstützt Mobilität im Sinne von Mobile IP. Schwerpunkt sind hier neben der Architektur ein Fördern verschiedener AAA-Server und die Definition einer Schnittstelle, welche spezifischen Anwendungsmodulen Zugriff auf AAA-Funktionen

verschafft. Bei der Gründung der AAAArch-Arbeitsgruppe wurden folgende Arbeitsschwerpunkte bzw. Ziele vereinbart:

- Ein Operieren von AAA-Servern über Betreiber Grenzen hinweg.
- Eine offene, erweiterbare Architektur, die schrittweise auf zukünftige Internetdienste ausgedehnt werden kann.
- Ein Konzept, welches AAA-Transaktionen nicht auf Netzbetreiber beschränkt, sondern jeder Art von Inhaltsanbietern mit einbezieht.
- Ein Sitzungskonzept, das anwendungsunabhängig arbeitet.
- Die Unterstützung von Sicherheitsmechanismen.
- Weltweite Skalierbarkeit.

Erste Referenzimplementierungen dieser Architektur unter Verwendung von Diameter sind z.B. im EU Projekt Moby Dick [39], [50] entstanden. Siehe hierzu auch Kapitel 4.8.

4.4.3 PANA

PANA (Protocol for carrying Authentication for Network Access) ist ein Protokoll, das Authentisierung für Netzzugang regelt. PANA wurde von der IETF PANA-Arbeitsgruppe [102] spezifiziert. Das Ziel dieser Gruppe ist es, ein Protokoll zu definieren, das es den Benutzern ermöglicht, sich selbst beim Zugangnetz auf Basis von IP-Protokollen von jeglichen Endsystemen aus zu authentisieren. PANA ermöglicht somit eine Entkopplung von Endgerät und Benutzer. Durch die Entkopplung ist es notwendig, dem Konsumieren eines Dienstes eine Authentisierung voranzustellen, welchen dem Autorisierungsprozess vorgelagert ist. Diese Authentisierung erfordert ein Protokoll, das die Flexibilität besitzt, verschiedene Authentisierungsarten flexibel zu unterstützen. Dies beinhaltet eine dynamische Auswahl der Anbieter verschiedener Dienste, die ein mobiler Endbenutzer ortsunabhängig auswählen kann. Da es auf Schicht 2 kein einheitliches Authentisierungsprotokoll gibt, welches technologieübergreifend eingesetzt werden kann, wurden in jüngerer Zeit sehr viele Lösungen eingeführt, die auf verschiedenen Schichten eine proprietäre Authentisierung durchführen. Beispiel hierfür ist das Setzen des Registrationsbits bei Mobile IP, das ein Registrieren auf Basis des SIP-Protokolls oder die Autorisierung auf Anwendungsebene, die z.B. eine Passwordeingabe zur Freischaltung von Webseiten erfordert.

All diese Authentisierungsmechanismen greifen viel zu kurz und stellen keine umfassende und schichtenübergreifende Autorisierung dar, welche den harten zeitkritischen Anforderungen von mobilen Szenarien ausreichend Rechnung tragen.

Der zentrale Gedanke von PANA ist das Bereitstellen eines Protokolls, welches mobilen Benutzern ermöglicht, sich auf Basis des IP-Protokolls authentisieren zu können. Weiter sieht PANA vor, dass eine Zusammenarbeit mit der netzseitigen AAA-Infrastruktur erfolgen kann, ohne die Diameterspezifischen Protokolldirektiven im Detail kennen zu müssen. PANA zielt somit auf den im Rahmen von AAA nicht spezifizierten Netzabschnitt zwischen dem Endsystem und dem AAA-Attendant ab und kann weder als neues Authentisierungsverfahren, noch als Schlüsselverteilungsmechanismus bezeichnet werden. PANA transportiert in Vereinbarung mit EAP (siehe hierzu auch Kapitel 4.4.4), die Authentisierungsnachrichten auf der letzten Meile zwischen Endsystem und Netz. Um alle Anforderungen von EAP erfüllen zu können, ist jedoch eine Erweiterung von EAP erforderlich. Diese Aufgabe wird jedoch nicht innerhalb der PANA-Arbeitsgruppe behandelt.

Die PANA-Arbeitsgruppe arbeitet unter der Annahme, dass sich in jedem (mobilen) Endgerät ein PANA Client (*PaC*) befindet, der bereits mit einer IP-Adresse konfiguriert ist, bevor der

PANA-Authentisierungsprozess einsetzt. Dies erfordert eine geeignete Zusammenarbeit mit Mobile IP, welches die IP-Adresse dynamisch bereitstellt. Mittels dieser IP-Adresse muss es dem mobilen Endgerät ermöglicht sein, spezifische Parameter zu einem PANA-Authentisierungs-Agenten (*PAA*) übermitteln zu können. Nach einer erfolgreichen Authentisierung können weiterführende Operationen angestoßen werden. Dies kann z.B. das Benutzen einer anderen IP-Adresse sein, oder aber das Erneuern von netzseitigen Filterregeln, die nach erfolgreicher Autorisierung entsprechende Rechte der IP-Adresse und somit dem entsprechenden Endgerät bereitstellen. In der IETF PANA-Arbeitsgruppe wird aktuell davon ausgegangen, dass die Fähigkeit EAP-Rümpfe geeignet zu transportieren, ausreichend ist. Es wird jedoch nicht ausgeschlossen, dass EAP ggf. an PANA angepasst werden muss, um zu einem effizienten und einheitlichen Authentisierungskonzept zu kommen. PANA gibt ausdrücklich keine Ressourcen nach erfolgreicher Authentisierung frei, um z.B. Ressourcenzugriffe von nichtautorisierten Teilnehmern zu verhindern. Der Kommunikationspartner des *PaC* ist der *PAA*. Dieser arbeitet nicht als ein Akteur, der explizit Ressourcen freigibt, sondern kommuniziert das Ergebnis eines Authentisierungsvorganges zu einer Komponente, die für die Ressourcenverwaltung (hier Ressourcenfreigabe) zuständig ist.

PANA ermöglicht ein Authentisieren ohne spezielle Kenntnis der darunter liegenden Schicht 2 Protokolle und eignet sich somit prinzipiell für konkurrierende Systeme (multiaccess), als auch für Punkt-zu-Punkt Systeme und unterstützt verschiedene Authentisierungs-Methoden, dynamische Betreiberwahl und Roaming.

Zusammenfassend lässt sich festhalten, dass PANA sehr leicht in die AAA-Infrastruktur eingebettet werden kann. Im Folgenden werden die beiden bereits kurz erwähnten PANA-Komponenten *PAA* und *PaC* kurz vorgestellt.

4.4.3.1 PANA-Authentisierungsagent

Der PANA-Authentisierungsagent (*PAA*) ist für die Authentisierung der Passwörter zuständig, welche vom PANA Client bereitgestellt werden. Zusätzlich ist er für die Bereitstellung der Ressourcen verantwortlich, die im Anschluss an eine erfolgreiche Authentisierung erfolgt. In diesem Zusammenhang soll erwähnt werden, dass PANA keine Benutzer, sondern Netz-schnittstellen authentisiert. Somit ist es einzelnen Benutzern möglich, sich gleichzeitig über mehrere Netz-schnittstellen am Netz anzumelden, da PANA die Authentisierung basierend auf dem 2-Tupel Benutzer-ID/Netz-schnittstellen-ID Nutzer durchführt.

4.4.3.2 PANA Client

Der PANA Client (*PaC*) ist ein Prozess, der auf einem Endsystem operiert, um einen Netzzugang bei einem PANA-Authentisierungsagenten zu erlangen. Ein *PaC* ist somit einem Endgerät zugewiesen und hat geeigneten Zugang zu den für die Authentisierung erforderlichen Passwörtern.

Weitere in den PANA-Prozess eingebundene Komponenten sind der Zugangsrouten und eine Policy-Instanz, die die Entscheidungen des PANA-Authentisierungsagenten in geeignete Netzkonfigurationsbefehle umwandelt.

4.4.4 EAP

EAP (Extensible Authentication Protocol) ist ein allgemeines Protokoll zur Authentisierung, welches mehrere Authentisierungsmechanismen unterstützt. EAP ist in [10] spezifiziert.

Die grundlegende Funktionsweise von EAP basiert auf dem Austausch von EAP-Anforderungen/Antwortnachrichten zwischen Client und Authentisierungsinstanz. EAP-Nachrichten

bestehen aus einem Code-Feld, das aktuell 4 EAP-Nachrichten spezifiziert (*Anforderung, Antwort, Erfolg, Fehler*), einem ID-Feld, das dazu angedacht ist, Anforderung und Antwort miteinander abzugleichen und einem Längenfeld, welches die Länge des Gesamtpakets definiert.

Der Protokollablauf beginnt in der Regel mit dem Versenden einer EAP-Anforderungsnachricht, welche die Authentisierungsinstanz an den Klienten versendet. Der Klient wiederum beantwortet diese Anforderung und überträgt entweder die Passwörter/Berechtigungsnachweise direkt, oder schlägt eine spezielle Authentisierungsmethode vor. Der Protokollablauf wird mit dem Versenden einer *Erfolg*-Nachricht bzw. eine *Fehler*-Nachricht abgeschlossen.

4.4.5 RTFM

4.4.5.1 Architektur

In der IETF wurde Echtzeit-Flussmessung in der Real Time Flow Measurement (RTFM)-Arbeitsgruppe behandelt. Diese Arbeitsgruppe ruht zurzeit, da die Standards und Referenzimplementierungen verfügbar sind und aktuell keine neuen Anforderungen an die Arbeitsgruppe adressiert sind.

Ziel dieser Arbeitsgruppe war es, eine Architektur für Flow-Messung zu erarbeiten und zu standardisieren. Diese Architektur definiert drei verschiedene Einheiten:

- Meter
- Meter-Reader
- Meter-Manager

Kernelement der IETF RTFM-Gruppe ist die Definition der so genannten *Flows*, welche virtuell die unabhängig voneinander über das Netz transportierten Pakete in einen logischen Zusammenhang bringen, um auf diese Weise eine fiktive „Schicht 3 Verbindung“ zu definieren.

Als Ergebnis dieser Aktivität entstand sowohl eine Architektur, als auch eine Referenzimplementierung, mit der IP Flows effektiv gemessen werden können. Abbildung 33 stellt die Kernkomponenten der IETF RTFM-Architektur zusammenfassend dar.

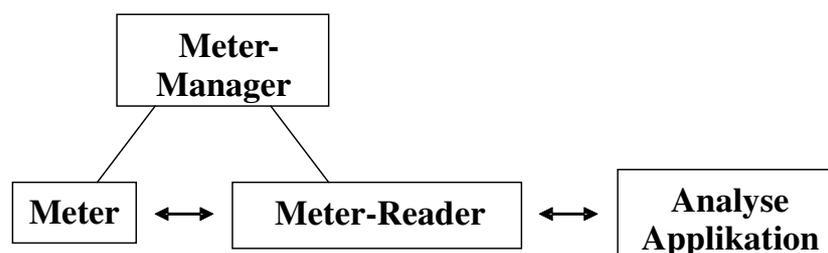


Abbildung 33: IETF RTFM-Architektur

Die Architektur hat die Eigenschaft, dass ein Flow-Modell auf verschiedenen Netzschichten (Schicht 3, Schicht 4 oder auf Anwendungsebene) definiert werden kann. Die Flow-Attribute sollten jedoch in einer Weise definiert werden, dass sie unabhängig von dem jeweiligen Schicht 3-Protokoll Allgemeingültigkeit besitzen [15].

Innerhalb des Modells wird die Schicht 3-Instanz Meter genannt. Der Meter zählt die spezifizierten Attribute, wie z.B. Summe aller Pakete und Byte, und weist diese einer zählbaren Instanz zu, die über Attribute wie IP-Quelladresse wiederum diese Daten in ein zählbares For-

mat bringt, damit eine virtuelle Schicht 3-Verbindung definiert werden kann. Der Meter verfügt über einen Satz an Regeln, die definieren, welche Flows für den jeweiligen Messprozess von Bedeutung sind und welche Daten entsprechend vernachlässigt werden können.

Der Meter-Reader sammelt die Daten, welche vom Meter bereitgestellt werden. Somit muss ihm die Identität des Meters bekannt gemacht werden. Dies beinhaltet auch eine eindeutige Adressierung des Meters, da diese beiden Instanzen logisch getrennt im Netz platziert werden. Der Meter-Reader muss weiter darüber informiert werden, zu welchen Zeitpunkten er die Daten aus dem Meter liest. Diese Daten werden vom Meter-Reader lokal abgespeichert um in einem nachgelagerten Prozess detailliert analysiert werden zu können. Für weitere Informationen über diese Architektur wird auf [19] verwiesen.

Der Meter-Manager konfiguriert sowohl Meter als auch Meter-Reader.

4.4.5.2 Referenzimplementierung

Eine freie Implementierung der IETF RTFM-Architektur ist NetTraMet und NeMac [20]. NetTraMet ist die Implementierung der RTFM-Meter. NetMac implementiert die Funktion von Meter-Manager und Meter-Reader. Abbildung 34 zeigt die Zuordnung der einzelnen Netzschichten der NeTraMet-Referenzimplementierung in einer TCP/IP-Umgebung.

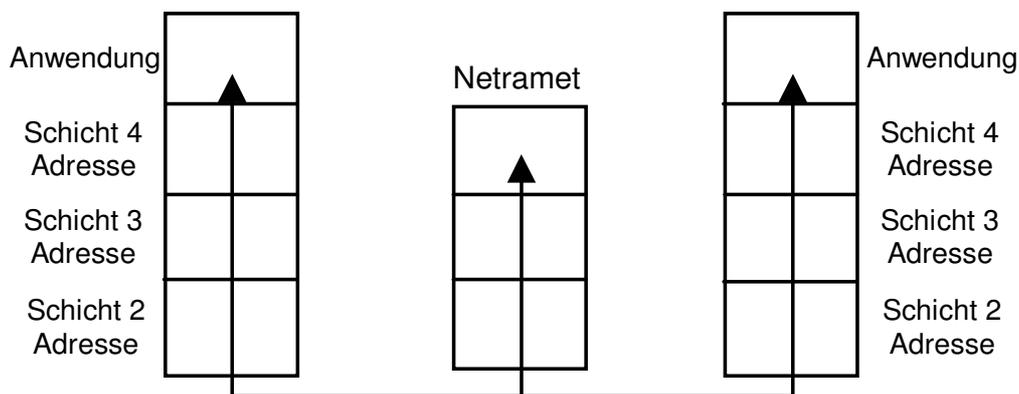


Abbildung 34: NeTraMet-Messkonzept

Die prinzipielle Arbeitsweise dieser beiden Implementierungen soll im Folgenden kurz beschrieben werden: Der Meter schreibt zunächst je nach seiner Konfiguration die Daten in seinen Speicher. Hier kann über die Konfiguration eine bestimmte Vorauswahl an Daten, die für den jeweiligen Messvorgang von Interesse sind, getroffen werden. Die Regeln, mit denen diese Vorauswahl getroffen wird, können sich auf folgende Parameter beziehen:

- Protokoll
- Portnummer
- DSCP
- IP-Quell- und/oder Zieladresse.

Die NeMac-Implementierung kommuniziert mit NetTraMet über das SNMP-Protokoll. Eine entsprechende MIB [21] wird periodisch ausgelesen. Das Leseintervall, in der entsprechenden Literatur auch SNMP Request-Intervall genannt, kann manuell mit einem kleinsten Intervall von einer Sekunde konfiguriert werden. Das Ergebnis in Form einer ASCII-Datei, bzw. in eine SQL-Datenbank exportiert, ist exemplarisch in Abbildung 35 gezeigt.

NeTraMets großer Vorteil ist, dass die jeweiligen Flows frei spezifiziert werden können [14]. Die für diese Spezifikation notwendigen Regeln werden in einem so genannten Regelsatz (ruleset) definiert.

date	time	flowindex	sourc	sourcepeeraddress	destpeeraddress	sourcectra	sourcectra	desttra	firsttime	topdus	fro
2002-12-04	17:25:48	7	2	2001:638:202:16:210:4BFF:FE3E:8309	2001:638:202:10:240:96FF:FE36:2BAD	58	129	0	1612	26	0
2002-12-04	17:25:53	4	2	FE80::204E2FFF:FE3A:D258	FF02::1	58	134	0	495	40	0
2002-12-04	17:25:53	5	2	FE80::204E2FFF:FE0E:C99B	FF02::1	58	134	0	554	37	0
2002-12-04	17:25:53	6	2	2001:638:202:10:240:96FF:FE36:2BAD	2001:638:202:16:210:4BFF:FE3E:8309	58	128	0	1612	30	0
2002-12-04	17:25:53	7	2	2001:638:202:16:210:4BFF:FE3E:8309	2001:638:202:10:240:96FF:FE36:2BAD	58	129	0	1612	30	0
2002-12-04	17:25:56	4	2	FE80::204E2FFF:FE3A:D258	FF02::1	58	134	0	495	44	0
2002-12-04	17:25:56	5	2	FE80::204E2FFF:FE0E:C99B	FF02::1	58	134	0	554	40	0
2002-12-04	17:25:56	6	2	2001:638:202:10:240:96FF:FE36:2BAD	2001:638:202:16:210:4BFF:FE3E:8309	58	128	0	1612	33	0
2002-12-04	17:25:56	7	2	2001:638:202:16:210:4BFF:FE3E:8309	2001:638:202:10:240:96FF:FE36:2BAD	58	129	0	1612	33	0
2002-12-04	17:26:00	4	2	FE80::204E2FFF:FE3A:D258	FF02::1	58	134	0	495	47	0
2002-12-04	17:26:00	5	2	FE80::204E2FFF:FE0E:C99B	FF02::1	58	134	0	554	44	0
2002-12-04	17:26:00	6	2	2001:638:202:10:240:96FF:FE36:2BAD	2001:638:202:16:210:4BFF:FE3E:8309	58	128	0	1612	37	0
2002-12-04	17:26:00	7	2	2001:638:202:16:210:4BFF:FE3E:8309	2001:638:202:10:240:96FF:FE36:2BAD	58	129	0	1612	37	0
2002-12-04	17:26:04	4	2	FE80::204E2FFF:FE3A:D258	FF02::1	58	134	0	495	51	0
2002-12-04	17:26:04	5	2	FE80::204E2FFF:FE0E:C99B	FF02::1	58	134	0	554	48	0
2002-12-04	17:26:04	6	2	2001:638:202:10:240:96FF:FE36:2BAD	2001:638:202:16:210:4BFF:FE3E:8309	58	128	0	1612	41	0
2002-12-04	17:26:04	7	2	2001:638:202:16:210:4BFF:FE3E:8309	2001:638:202:10:240:96FF:FE36:2BAD	58	129	0	1612	41	0
2002-12-04	17:26:04	12	2	2001:638:202:10:240:96FF:FE36:2BAD	2001:638:202:16:210:4BFF:FE3E:8309	6	32779	22	5611	1	1
2002-12-04	17:26:08	4	2	FE80::204E2FFF:FE3A:D258	FF02::1	58	134	0	495	55	0
2002-12-04	17:26:08	5	2	FE80::204E2FFF:FE0E:C99B	FF02::1	58	134	0	554	52	0
2002-12-04	17:26:08	6	2	2001:638:202:10:240:96FF:FE36:2BAD	2001:638:202:16:210:4BFF:FE3E:8309	58	128	0	1612	45	0
2002-12-04	17:26:08	7	2	2001:638:202:16:210:4BFF:FE3E:8309	2001:638:202:10:240:96FF:FE36:2BAD	58	129	0	1612	45	0
2002-12-04	17:26:08	12	2	2001:638:202:10:240:96FF:FE36:2BAD	2001:638:202:16:210:4BFF:FE3E:8309	6	32779	22	5611	2	2
2002-12-04	17:26:12	4	2	FE80::204E2FFF:FE3A:D258	FF02::1	58	134	0	495	59	0

Abbildung 35: NeTraMet-Ausgabe in SQL-Datenbank

4.4.6 Policy-based Accounting

Tseby, Carle und Zander haben in [26] ein Konzept zur Nutzungsdatenerfassung vorgeschlagen, welches von der IETF von der IETF als *Experimental Standard* akzeptiert wurde. Kerngedanke des Policy-based Accounting-Konzeptes ist es, mit einer generischen Architektur auf die Anforderungen verschiedener Dienstanbieter einzugehen, diese sind:

- Größe des Dienstanbieters
- QoS-Bereitstellungskonzept
- Dienstklassen
- Charging-Modelle

Die vorgeschlagene Policy-based Accounting-Architektur basiert auf IETF AAA-Mechanismen nach Kapitel 4.4 und sieht Regeln vor, nach denen der Prozess der Weiterverarbeitung der Nutzungsdaten effektiv vollzogen werden kann.

In diesem Vorschlag wird zwischen vier verschiedenen Schemata unterschieden, welche eine regelbasierte Nutzungsdatenerfassung und deren Weiterverarbeitung beschreiben. Dies sind:

- Integriertes Verfahren: Hier wird der Prozess der Nutzungsdatenerfassung und nachgelagerter Weiterverarbeitung als mit dem eigentlichen Dienst gekoppelt betrachtet. Man kann dies als dienstspezifisches Vorgehen beschreiben.
- Diskretes Verfahren: Hier wird der Prozess der Nutzungsdatenerfassung und nachgelagerter Weiterverarbeitung völlig vom Dienst entkoppelt. Dieses Konzept eignet sich prinzipiell, um den gesamten Prozess auslagern zu können.
- Verfahren, welche sich auf einen administrativen Bereich beschränken: Hier beschränken sich die Nutzungsdatenerfassung und der Austausch von den entsprechenden Nutzungsberichten auf den Bereich einer administrativen Organisation.
- Verfahren, welches betreiberübergreifend ist: Hier ist konzeptionell vorgesehen, dass die Nutzungsdatenberichte zwischen verschiedenen Administrativen Bereichen ausgetauscht werden. Dies ist z.B. im klassischen Roaming bei der Mobilkommunikation sehr verbreitet.

4.5 Internet Protocol Data Record

Internet Protocol Data Record (IPDR) ist ein offenes Konsortium von verschiedenen Diensteanbietern, Herstellern, Systemintegratoren und kommerziellen Herstellern von Abrechnungsoftware, das sich das Ziel gesetzt hat, den Austausch von benutzerspezifischen Daten z.B. zum Zweck der Abrechnung zu standardisieren. Bis jetzt wurden im Rahmen dieser Initiative Datenstrukturen vorgeschlagen, welche Benutzerdaten zur Abrechnung spezifizieren. Darüber hinaus wurden Bibliotheken vorgeschlagen, welche diese Strukturen implementieren und es wurde ein Repository von Dienstspezifikationen definiert.

Ein Ergebnis der IPDR-Initiative ist die Network Data Management Usage (NDMU) Specification [30]. In ihrer aktuellen Version 3 wurde ein generisches Format errichtet, das sowohl heutigen Anforderungen an ein Accounting Format genügt, als auch den Anspruch hat, zukünftigen Anforderungen zu genügen. Es unterstützt zahlreiche zukünftige Dienste, wie z.B. Voice over IP, Streaming Video, IP VPN (Virtual Private Network) und auch jegliche 3G-Dienste. Durch die Einbeziehung von XML (Extensible Markup Language) wird ein schnelles und problemloses Interagieren mit Lösungen erreicht, die sich von der IPDR-Spezifikation unterscheiden. Durch die Möglichkeit, herstellereigene Erweiterungen direkt vorzunehmen, können sowohl Rückkompatibilitätskriterien, als auch sehr spezielle herstellereigene Lösungen realisiert werden. Im Folgenden sollen nun einige ausgewählte Spezifikationen näher vorgestellt werden.

4.5.1.1 IPDR DOCSIS

Die IPDR DOCSIS (Data Over Cable Service Interface Specification) Spezifikation beschreibt die Schnittstellenanforderung an Kabelmodems (CM), welche auf Kabelfernsehtznetzen bidirektionale Horchgeschwindigkeits-Datenzugänge bereitstellen. Diese Entwicklung ist mit dem Kontext dieser Arbeit insofern vergleichbar, als dass ein existierendes Übertragungsmedium mit konkurrierendem Medienzugriff dem Internet zugänglich gemacht wird, und sich nun die Frage einer geeigneten Abrechnung stellt.

Auch bei DOCSIS bekommt der Benutzer, ähnlich wie bei Mobile IP, dynamisch eine IP-Adresse zugewiesen. Somit kann ein Benutzer nicht statisch über einen längeren Zeitraum anhand einer statischen IP-Adresse identifiziert werden. Im Gegensatz zur Mobile IP & AAA-Lösung werden jedoch bei DOCSIS die Benutzer mittels ihrer Schicht 2-Adresse eindeutig identifiziert bzw. Accounting-Berichte werden anhand der Schicht 2-Adresse verwaltet.

Mittels eines geeigneten Konzeptes zur Nutzungsdatenerfassung wird zum Zweck der Auditierung die Leistungserstellung überprüft. Diese Überprüfung geschieht anhand von vereinbarten SLAs, welche die zu erbringende Dienstgüte beschreiben, die ein Betreiber einem Benutzer im Rahmen eines Vertrages anbietet. In der Regel ist ein SLA nicht frei verhandelbar, sondern der Betreiber bietet dem Benutzer an, zwischen einigen vordefinierten Paketen auszuwählen.

Diese Pakete sind Service Bündel im Sinne von Kapitel 3.4, die dann zu differenzierten Preisen angeboten werden können. Ein solches Paket bestimmt nun die Parameter anhand von einer eindeutigen Kennung (*Service Flow ID – SFID*) und einem Dienstklassennamen (*Service Class Name – SCN*) und stellt in IPDR DOCSIS die kleinste Einheit der Nutzungsdaten dar, welche im Erfassungsprozess generiert wird.

IPDR DOCSIS sieht vor, dass der Benutzer das gewählte Dienstpaket ändern kann. Somit ist es möglich, einem Benutzer verschiedene IPDRs zuzuordnen.

Ein wesentlicher Parameter in IPDR ist die benutzte Bandbreite, welche in beiden Richtungen erfasst wird. Je nach vereinbartem SLA ist die Kabelinfrastruktur gefordert, Pakete ggf. zu verzögern bzw. zu verwerfen, sollten die vereinbarten SLA-Obergrenzen bzgl. maximaler Bandbreite überschritten werden. Diese Daten werden in IPDR DOCSIS ebenfalls mit dem Ziel erfasst, sie als „Akquisitionsargument“ zur Neudefinition der SLAs mit dem Kunden einzusetzen.

Prinzipieller Ablauf

Die DOCSIS 1.1 RFI-Spezifikation [31] stellt einen Mechanismus bereit, mit dem sich ein Kabelmodem (*CM*) mit der netzseitigen Gegenstelle, dem so genannten Kabelmodem Terminierungssystem (*CMTS*) beim Einschalten registrieren und konfigurieren kann. Diese Konfiguration beinhaltet neben dieser Authentisierung auch eine Autorisierung und die anschließende Konfiguration der entsprechenden Dienstgüteparameter. Diese Konfiguration besteht im Wesentlichen aus einer Klassifikation der die Schnittstelle zur Kabelinfrastruktur passierenden Pakete in einen Dienstfluss. Ein Dienstfluss ist hier eine in unidirektionaler Richtung ablaufende Sequenz von Paketen, welcher gemäß SLA dieselbe Dienstgüte zugewiesen werden muss. *CM* und *CMTS* führen nun Verkehrsformungs-, Priorisierungs- und Auswahlverfahren gemäß SLA für die entsprechende Nutzungsklasse durch.

Die grundlegenden Mechanismen für diese Dienstgüterebereitstellung bestehen aus 4 elementaren Funktionen. Diese sind zunächst die Vorkonfiguration und Registrierung von *CM* und *CMTS* (QoS-Parameter). Weiter muss eine Klassifikation der Pakete erfolgen, welche es ermöglicht, Pakete einem entsprechenden Fluss zuzuordnen. Im Anschluss daran werden die verschiedenen Flüsse zu Dienstklassen gruppiert, damit höher liegende Anwendungen auf konsistente Art und Weise eine dedizierte Dienstgüte für eine Anwendung anfordern können.

Die einheitliche Verwendung der *SCNs* ist mittels einer MIB (Management Information Base) sichergestellt. In DOCSIS wird sichergestellt, dass die *SCNs* innerhalb eines administrativen Bereiches eindeutig und unverwechselbar sind. Hier werden Namen wie z.B. *PrimaryUp*, *GoldUp*, *VoiceDn*, oder *BronzeDn* verwendet. Weiter gibt es den Begriff eines Dienstpakets, welches ein SLA zwischen einem Betreiber und den Endbenutzers definiert. Dieses Paket gruppiert verschiedene Dienstklassen unter Verwendung der *SCNs*. Hier sind Dienstpakete wie z.B. Gold, Silber oder Bronze denkbar.

Jeder aktive Dienstfluss wird durch eine 32 bit *SFID* identifiziert. Diese wird vom *CMTS* explizit für einen Fluss einem *CM* zugewiesen. Die *SFID* ist jedoch nicht ausreichend, um die Abrechnung geeignet durchführen zu können, da die Abrechnungsinfrastruktur nicht in der Lage ist, Parameter wie z.B. Bandbreite eindeutig einer *SFID* zuzuweisen. Die Kombination von *SCN* und *SFID* ermöglicht hier eine eindeutige Zuordnung der Nutzungsdaten mit den Nutzerdaten im Abrechnungssystem. Das Abrechnungssystem kann dann die Abrechnungseinheiten entsprechend jedem einzelnen dienstspezifischen Zähler zuweisen. Somit erhält das Abrechnungssystem dienstklassenspezifische Zählerstände basierend auf den Dienstklassen. Sollten einem Fluss verschiedene *SFIDs* bei einheitlichem *SCN* zugewiesen worden sein, generiert das Abrechnungssystem verschiedene Abrechnungssätze.

Anforderungen an die Abrechnungsberichte

Das *CMTS* stellt dem Abrechnungssystem vorformatierte Abrechnungsberichte nach Anfrage zur Verfügung. Dies geschieht für alle am *CMTS* angeschlossenen Benutzer separat. Hier ist das kleinste Intervall, zwischen zwei Abrechnungsberichten, die zum *CMTS* gesendet werden, auf 15 Minuten festgelegt. Zur Bewerkstelligung dieses Prozesses wurden in IPDR folgende Anforderungen definiert:

- Der Benutzer muss mittels seines Endgerätenamens und der IP-Adresse identifiziert werden. Weiter wird der absolute Zeitpunkt, in der der Abrechnungsbericht entstanden ist, und der Zeitpunkt, zu dem sich das Gerät erstmals am Netz angemeldet hat, erfasst.
- Der Abrechnungsbericht beinhaltet die Schicht 2-MAC-Adresse und alle IP-Adressen, die das Endsystem verwendet hat.
- Der Abrechnungsbericht muss für jeden aktiven und im Abrechnungsintervall terminierten Dienstfluss einen Eintrag haben. Diese Dienstflüsse werden durch *SFID* und *SCN*, welche zur Sitzung verwendet wurden, identifiziert. Eine Unterscheidung zwischen aktiven und terminierten Dienstflüssen ist erforderlich.
- Jedes *CMTS* generiert zunächst interne Abrechnungsberichte, welche nach einem Formatierungsprozess dem externen Abrechnungssystem bereitgestellt werden. Das Löschen dieser internen Berichte von außen ist nicht möglich und wird frühestens nach Abholung der aufbereiteten Berichte durch einen internen Prozess ermöglicht. Hierfür wird das SNMP-Protokoll mit einer entsprechenden MIB, die so genannte DOCSIS-QOS-MIB, vorgesehen.
- Die Flussrichtung der Dienstflüsse muss ohne Bezug auf den *SCN* jederzeit nachvollziehbar sein. Die Zahl der Pakete und Octets ist jeweils pro Flussrichtung zu erfassen. Weiter ist die Zahl der Pakete in „downstream“ Richtung, welche wegen SLA-Verstoß verzögert bzw. verworfen wurden, zu erfassen.
- In IPDR DOCSIS müssen jegliche Zähler 64 bit lang sein und werden vom *CMTS* verwaltet. Nach jeglicher Neuinitialisierung müssen diese auf null gesetzt werden. Diese Neuinitialisierung wird mittels der *sysUpTime* erfasst (vergleiche Kapitel 4.5.1.2).
- Um die Formatierung zu vereinfachen, wird ein XML-Format für die Abrechnungsberichte vorgeschlagen.
- Es wird in IPDR DOCSIS darauf hingewiesen, dass das Datenvolumen der Abrechnungssysteme erheblich ist; für den effektiveren Transfer der Berichte wird eine Reduzierung der Datenmenge mittels einer Datenkomprimierung gemäß [38] vorgeschrieben. Hier wird von einem Verhältnis von 30:1 ausgegangen.
- Um die Übertragung der Abrechnungsberichte zu verbessern, ist eine TCP-Verbindung mit entsprechend hoher Durchsatzrate erforderlich. Zur Übertragung wird FTP zwingend vorgeschlagen.
- Zur Sicherstellung der Datenkonsistenz, holt sich das Abrechnungssystem explizit die Daten vom *CM* und informiert explizit nach erfolgreichem Übertrag das *CM*, um ihm eine Freigabe des belegten Speichers zu erlauben. Nach dieser Freigabe wird im *CM* ein neuer Abrechnungsbericht erzeugt.
- Die Übertragung der Abrechnungsberichte erfolgt verschlüsselt und schließt einen Authentisierungsprozess mit ein. Die Vertraulichkeitsbeweise werden nur zu Abrechnungszwecken verwendet.

4.5.1.2 Elementare Attribute der Nutzungsdaten

Ein DOCSIS IPDR besteht aus verschiedenen Bestandteilen, die nun im Detail vorgestellt werden.

IPDR Information

- *IPDRcreationTime*: Ein generischer IPDR wird mit einem Entstehungsdatum versehen. Dieses Datum ist weltweit einheitlich (standardisierte Zeitzone GMT in DOCSIS 1.1 IPDRs).
- *seqNum*: Optional kann noch eine laufende Nummer vergeben werden

CMTS Information

Ein DOCSIS 1.1 IPDR beinhaltet die folgenden Elemente, welche das Endsystem (Kabelmodem) eindeutig identifiziert. Dies sind:

- *CMTShostName*
- *CMTSipAddress*: in DOCSIS aktuell eine IPv4-Adresse
- *CMTSsysUpTime*: Die Zeit, bei der das Endsystem aktiv ans Netz angeschlossen wurde.

Subscriber Information

Ein DOCSIS 1.1 IPDR beinhaltet folgende Elemente, welche den Benutzer eines Endsystems (Kabelmodems) einheitlich identifizieren. Dies sind:

- *SubscriberId*: Die Subscriber-ID besitzt weltweite Gültigkeit und wird durch die MAC-Adresse des CM beschrieben. Sie entspricht also der Ethernetadresse der kabelseitigen Schnittstelle der Kabelmodems.
- *CMipAddress*: Die IP-Adresse der Endsystems wird dazu verwendet, das Kabelmodem zu Verwaltungszwecken jederzeit aus dem Netz heraus erreichen zu können. Diese Adresse wird dem Kabelmodem automatisch zugewiesen um z.B. über SNMP weitere Verwaltungsaufgaben durchführen zu können.
- *CPEipAddress*: Diese Adresse wurde dem Endsystem für das Reportingintervall zur Verwendung zugewiesen.

Service Flow Information

Eine DOCSIS IPDR beinhaltet folgende Elemente, welche den Dienstfluss identifizieren und die Nutzungsdaten für den entsprechenden Dienstfluss repräsentieren:

- *SFtype*: Ein Dienstfluss kann zwei verschiedene Zustände einnehmen. Diese sind entweder *Interim* oder *Stop*. Ein Fluss vom Typ *Interim* beschreibt einen aktiven Dienstfluss. Ein Fluss vom Typ *Stop* berichtet über einen beendeten Dienstfluss. In diesem Fall gibt es nur diesen einen Nutzungsbericht. Erstreckt sich eine Sitzung über mehrere Berichtsintervalle, dann ist der letzte dieser Berichte vom Typ *Stop*, die vorhergehenden vom Typ *Interim*.
- *SFID*: Ein Dienstfluss ist innerhalb der *CMTS* eindeutig als solcher durch seine *Service Flow ID (SFID)* identifizierbar. Diese Eindeutigkeit wird dadurch gewährleistet, dass die MAC-Adresse Bestandteil dieser ID ist. Im Detail besteht die Service Flow ID aus einer MAC ID, einer eindeutigen Identifizierung der MAC-Schnittstelle und einem SFID-Teil, der sich relativ zur MAC-Schnittstelle verhält.
- *ServiceClassName (SCN)*: Der *SCN* wird vom *CMTS* einer Dienstsitzung zugewiesen. Dieser externe Name wird mit den QoS-Parametern verbunden. Diese QoS-Parameter definieren, wie die einer Sitzung zugeordneten Pakete behandelt werden, um das vereinbarte SLA einzuhalten. Innerhalb DOCSIS 1.1 wurden Dienstklassen vorgeschlagen, welche zunächst abstrakt definiert wurden. In diesem Zusammenhang ist zu erwähnen, dass die Verwendung eines *SCNs* innerhalb der DOCSIS-Schnittstelle zwischen *CM* und *CMTS* optional ist, jedoch für eine eindeutige Abrechnung, die dienstspezifisch vorgesehen ist, empfohlen ist.
- *SFdirection*: Dieser Parameter gibt die Flussrichtung des Dienstflusses an. Hierbei kann es sich folglich entweder um einen so genannten upstream Fluss, also in das Netz gerichtet, oder einen downstream Fluss, also aus dem Netz kommend, handeln.
- *OctetsPassed*: Dieser Parameter gibt die Anzahl der erfassten Oktets an, welche einem Dienstfluss zugeordnet werden konnten.

- *PktsPassed*: Dieser Parameter gibt die Anzahl der Pakete an, welche einem Dienstfluss eindeutig zugeordnet werden konnten.
- *SLAdropPkts*: Dieser Parameter gibt die Anzahl der verworfenen Pakete an, die einem Dienstfluss eindeutig zugeordnet werden konnten. Die Verwerfungen werden im Rahmen von IPDR DOCSIS nur in "downstream" Richtung erfasst.
- *SLAdelayPkts*: Dieser Parameter gibt den aktuellen Zählerstand der Pakete wieder, die wegen eines Verstoßes gegen das SLA verzögert weitergegeben wurden. Je nach Preismodell ist dieser Zähler jeweils pro Flussrichtung verschieden. Im Rahmen von IPDR wird nur die so genannte „downstream“ Richtung erfasst.

4.5.1.3 IPDR VoIP

Als Voice over IP (VoIP) bezeichnet man die Sprachkommunikation zwischen zwei oder mehreren Parteien über eine internetbasierte Verbindung. Ein Anruf ist von einem Anrufer initiiert und wird von einem oder mehreren Gegenstellen empfangen.

Im Rahmen der IPDR VoIP-Spezifikation [32] werden Nutzungsreports zu einer zentralen Netzinstanz weitergeleitet. Bei diesem Vorgang werden zusätzlich weitere Parameter, wie z.B. die Dienstgüte, verteilt erfasst und an einem zentralen Punkt gegeneinander abgestimmt. Diese per VoIP-Anruf erfassten Informationen werden dann mit dem SLA, also den aus dem Benutzervertrag abgeleiteten Dienstgüteparametern, abgeglichen. Erst nach diesem Abgleich ist es möglich, die entsprechenden Abrechnungs-Berichte (vergleiche Kapitel 4.2.4.16 und Kapitel 4.2.4.17) zu erstellen und an das Abrechnungssystem weiterzuleiten. Bei IPDR VoIP ist vorgesehen, nach jedem Ende einer VoIP-Sitzung (unabhängig davon, ob sie natürlich oder anderweitig beendet wurde) einen IPDR zu erzeugen. Optional ist weiter vorgesehen, bei entsprechend länger andauernden Sitzungen auch zwischendurch IPDRs zu erzeugen, die sich an bestimmten Ereignissen orientieren.

4.6 Abrechnungsinfrastruktur in der ITU

Die ITU (International Telecommunication Union) ist die Institution der Vereinten Nationen, welche Telekommunikationsstandards verabschiedet. So arbeitet innerhalb der ITU ein permanentes Organ an verschiedenen Standards und Empfehlungen, welche auch die Nutzungsdatenerfassung und das Verwalten der erfassten Daten in einem Telekommunikationsnetz umfasst. Konkret wurde in der Q.825-Empfehlung ([64]) die Funktionalität und das Informationsmodell der Nutzungsdatenerfassung in einem Telekommunikationsnetz spezifiziert.

Diese Spezifikation wurde hauptsächlich für Analogtelefonie, ISDN und IN-Dienste erstellt, welche zum Zweck der Nutzungsdatenverwaltung von verbindungsorientierten Sitzungen in diesen Netzen für die spätere Rechnungserstellung vorangetrieben.

Die Q.825-Empfehlung beinhaltet die Spezifikation eines Datenformats der Nutzungsdatenerfassung und nachgelagerten Weiterverarbeitung, welche in so genannten *Call Detail Records* beschrieben wird.

Abbildung 36 illustriert den in der ITU-T Q.825-Empfehlung spezifizierten Gesamtprozess. In der Empfehlung wird bereits auf die enorme Datenmenge, welche bei der Erfassung der Daten entsteht, hingewiesen. Hier wird bereits in der Komponente *CDR generation control* die Möglichkeit einer intelligenten Konfiguration der zu erfassenden Parameter vorgeschlagen. Für die Übertragung der erfassten Daten ist vorgesehen, sowohl Echtzeitereignisse, als auch Berichte, welche ein definiertes Zeitintervall zusammenfassen (bulk), zu unterstützen.

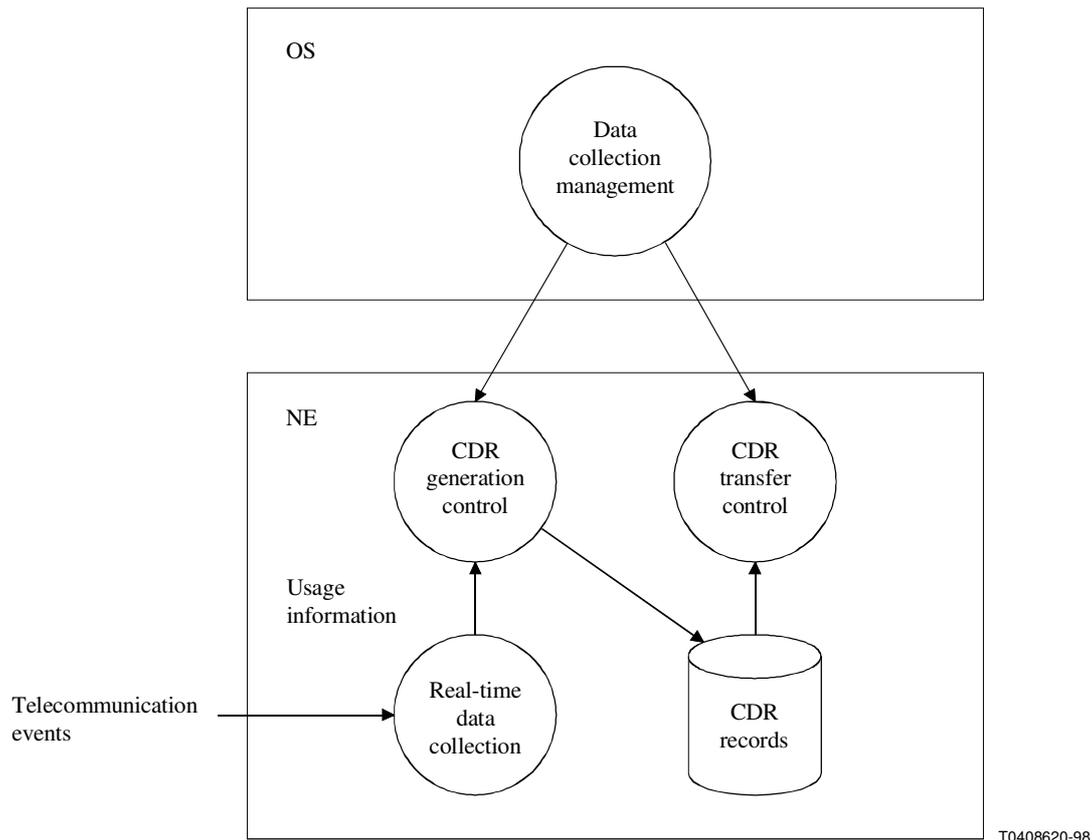


Abbildung 36: Nutzungsdatenerfassung und nachgelagerte Weiterverarbeitung aus [64]

4.7 2G und 3G-ChargingProtokolle

In diesem Unterabschnitt soll ein Überblick gegeben werden, wie in den in Kapitel 2.2 vorgestellten Architekturen die Abrechnung implementiert ist.

Aktuell ist für Charging im 3G-Umfeld das GTP Charging-Protokoll vorgesehen [64]. Hier werden Nutzungsdatenberichte so genannte *Charging Data Records* oder auch *Call Detail Record (CDRs)* erzeugt, die von der zentralen Instanz (*CDF*) gesammelt werden. In diesem Zusammenhang sei hier nochmals auf Abbildung 16 verwiesen.

4.7.1 GPRS-Nutzdatenerfassung

Bei GPRS wird mit Hilfe des GTP-Protokolls das Erfassen der Nutzungsdatenberichte unterstützt. Die *CDF* wird jedoch nur dann eingesetzt, wenn die Meter-Instanz und die Accounting-Instanz nicht am identischen physikalischen Ort stattfinden. Oftmals befinden sich im 3G-Netz der Meter auf dem *SGSN* und der Accounting-Prozess auf dem *GGSN*.

Ursprünglich wurde *GTP* für den Charging-Prozess für *GPRS*-Diensten entwickelt, bei dem die Pakete im *GPRS*-Netz getunnelt werden. Ein modifizierter Standard *GTP'* basiert auf *GTP* und wurde mit einigen Erweiterungen versehen. *GTP'* arbeitet auf der GA-Schnittstelle (vergl. Abbildung 16), setzt jedoch kein *GPRS*-Kernnetz voraus und kann prinzipiell auch auf anderen Netzarchitekturen implementiert werden. *GTP'* beinhaltet die folgenden Funktionen:

- Einen *CDR*-Übertragungsmechanismus zwischen den *GPRS*-Knoten welche die *CDRs* erzeugen und dem Charging-Gateway.
- Weitervermittlung eines *CDRs* zu einem anderen Charging-Gateway falls erforderlich.

- Die Fähigkeit Kommunikationsfehler zwischen Netzelementen, welche die *CDRs* verarbeiten zu erkennen.
- Die Fähigkeit eines *CDR*-Knotens den eigenen Zustand zu signalisieren. Hier ist es z.B. notwendig, nach einem Ausfall eines *CDR*-Knoten benachbarten Knoten diesen Ausfall anzuzeigen.

4.7.2 3GPP-Abrechnungssystem

Innerhalb von 3GPP wird eine Abrechnungsinfrastruktur definiert, deren Kernkonzept das Verwalten von *CDRs* festlegt. Er wurde speziell für Abrechnung von UMTS-Diensten entwickelt.

Die Nutzungsdaten, welche vom System erzeugt werden, werden neben dem Zweck der Rechnungserstellung des Endbenutzers auch zur Abrechnung zwischen den verschiedenen Betreibern und zur statistischen Analyse und zur Bearbeitung von Kundenbeschwerden (Auditierung) verwendet.

4.7.2.1 3GPP Charging-Architektur

Abbildung 37 zeigt die Architektur der Abrechnungsinfrastruktur von 3GPP. Die Datenerfassung erfolgt für jedes einzelne Endsystem von entweder dem bedienenden *MSC*, dem *SGSN* oder dem *GGSN*. Die Datensätze, welche der Betreiber zur letztendlichen Erzeugung der Rechnung verwendet, sind nicht spezifiziert. Weiter behandelt der Standard keine Zeitintervalle, zwischen zwei Rechnungen.

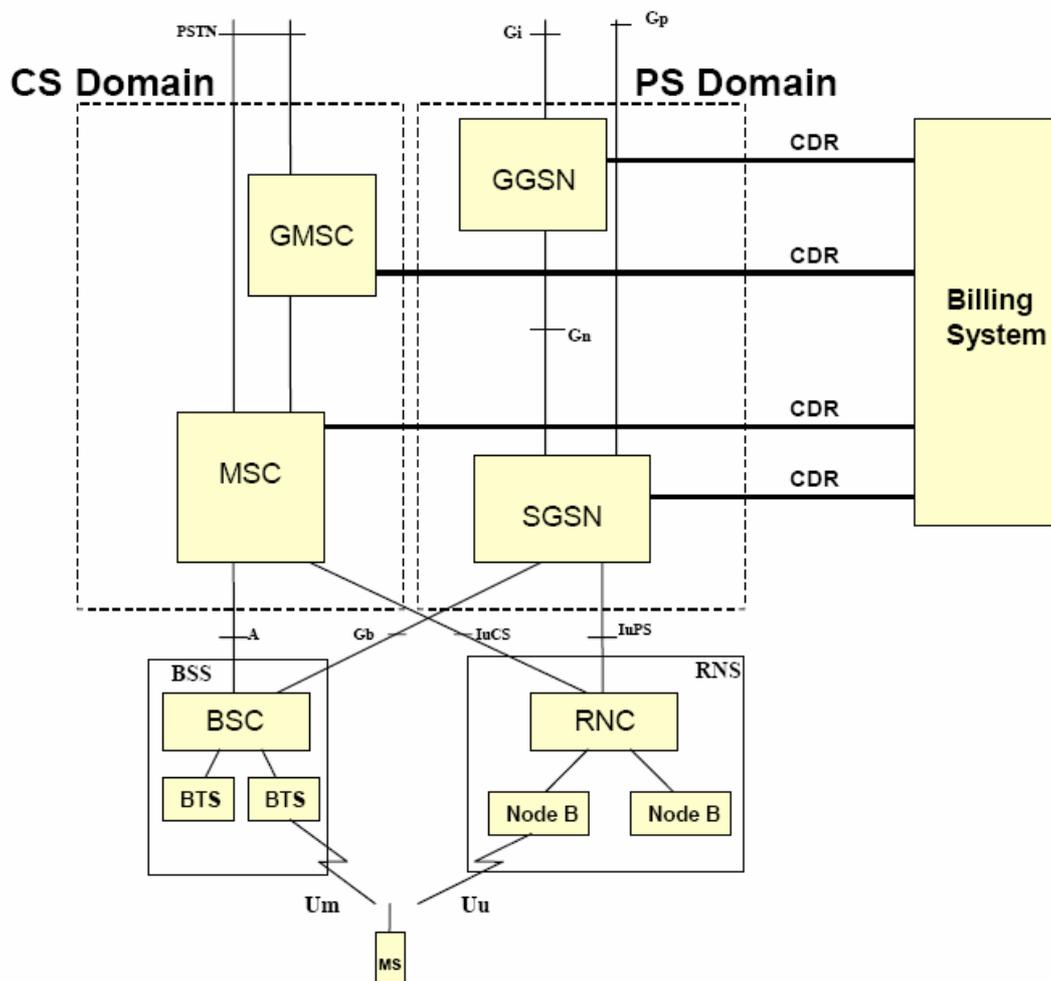


Abbildung 37: 3GPP Charging-Architektur (Quelle: 3GPP Konsortium)

4.7.3 OSA/PARLAY

Zukünftige Mobile Netze werden im Vergleich zu den aktuell verfügbaren Netzen weitaus vielfältigere Dienste bereitstellen. Diese zukünftigen Dienste, wie z.B. das synchrone und asynchrone Übertragen von Multimedia-Inhalten in Verbindung mit ortsbezogenen Diensten, werden einem Netzbetreiber neue kommerzielle Möglichkeiten eröffnen, welche als Kernmotivation für Investitionen in diese zukünftige Netzarchitektur betrachtet werden können.

Um diese neue Netzarchitektur auf ein kommerziell erfolgreiches Niveau zu heben, wird allgemein davon ausgegangen, dass die Einbindung so genannter Drittanbieter unverzichtbar ist. Dies erfordert jedoch neue Abrechnungs- und Benutzerverwaltungsmechanismen und Konzepte.

Die Service-Plattform des IN, wie momentan in den 2G und 3G-Netzen eingesetzt, hat nicht die Offenheit und Transparenz, um eine Dienstplattform für ein mobiles heterogenes Netz bereitzustellen. Eine wesentliche Schwachstelle der 2G-Netzarchitektur ist, dass jeder Dienst in ein standardisiertes Format transferiert werden muss, bevor er im Netz angeboten werden kann. Da dies ein sehr spezifisches Wissen und Verständnis der Netzschichten erfordert, ist es für Drittanbieter sehr schwer möglich, sich dieses Marktsegment zu erschließen. Aus diesem Grund wurden *Application Programming Interfaces* (APIs) spezifiziert, die katalytisch bei der Einführung dieser neuen Dienste wirken sollen. Heute kann man feststellen, dass sich diese Hoffnung nicht erfüllt hat.

Diese offenen APIs sind z.B.: JAIN [70] und Parlay [103]. Hier wurde durch die Abstraktion der Dienstdefinition von den telekommunikationsspezifischen Details ein Geschäftsmodell auf den Weg gebracht, das in 3G-Netzen kurzfristig und sehr flexibel ein Einführen dieser Dienste ermöglichen soll.

In 3GPP wurde als Folge von JAIN und Parlay die *Open Service Access (OSA)*-Architektur spezifiziert [3]. *OSA* beinhaltet eine Sammlung von verschiedener APIs, mit deren Hilfe eben Drittanbietern von Diensten Zugang zum Telekommunikationsmarkt verschafft werden kann. *OSA* bezieht sich auf Parlay, jedoch beinhaltet Parlay einige weitere APIs, die nicht Bestandteil von *OSA* sind. Eine wesentliche Eigenschaft von Parlay ist die Technologieunabhängigkeit. Somit kann Parlay unabhängig von der Schicht 2-Technologie eingesetzt werden.

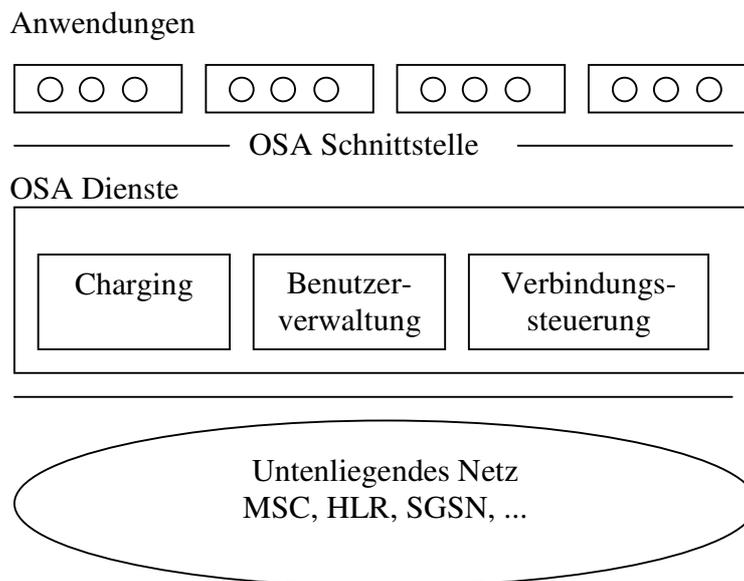


Abbildung 38: OSA-Übersicht

Neben der Sammlung von APIs definiert die *OSA*-Spezifikation darüber hinaus zwei wesentliche Konzepte. Dies sind das so genannte *Virtual Home Environment (VHE)* und das Basis-konzept der *OSA*. Benutzer von Endsystemen wird hier ermöglicht, personalisierte Dienste zu jeder Zeit und an jedem Ort konsumieren zu können, unabhängig von der Netztechnologie, die den Zugang zum Netz bereitstellt. Um diese Architektur auch für zukünftige Anwendungen offen zu halten, ist ein flexibles Gerippe zur Bereitstellung von Diensten nötig. In Abbildung 38 ist eine Übersicht über die *OSA*-Architektur dargestellt. Diese beinhaltet sowohl die Gesamtarchitektur, als auch die Funktionalität und Mechanismen, um dritten Diensteanbietern Zugang zum Netz zu verschaffen.

Die Funktionalität, d.h. die Dienste die *OSA* anbietet, sind neben Charging, Benutzerverwaltung und Verbindungssteuerung noch Benutzerermittlung, Benutzerstatus, Sitzungskontrolle für Daten und eine Charakterisierung der Beschaffenheit der Endsysteme. Diese werden nun im Einzelnen kurz vorgestellt.

4.7.3.1 OSA-Rahmenwerk

Wenn dritte Anbieter auf die Kernfunktionalität eines Netzes zugreifen, sind Sicherheit und Datenintegrität wesentliche Aspekte, welche berücksichtigt werden müssen. Nur erfolgreich authentifizierte Anwendungen sollten auf diese tiefer liegenden Funktionen zugreifen dürfen und der Gebrauch dieser Funktionen sollte zusätzlich streng kontrolliert werden. Beide Anforderungen werden von *OSA* erfüllt. Das *OSA*-Rahmenwerk verwaltet alle sicherheitskritischen Funktionen. Dies beinhaltet zunächst die Authentisierung der Anwendungen und die daran anschließende Bereitstellung von Diensten, die die Anwendungen gebrauchen möchten. Zusätzlich wird eine Integritätsverwaltung unterstützt, welche sowohl Verfügbarkeit der Dienste, als auch die Leistungsfähigkeit der Dienste berücksichtigt.

4.7.3.2 Verbindungskontrolle

Die *OSA*-Verbindungskontrolle überwacht und kontrolliert den Verbindungsaufbau über sowohl leitungsvermittelnde, als auch paketvermittelnde Netztechnologien. In diesem Zusammenhang werden vor allem Anwendungen über eventuell eingehende Ereignisse (Beginn bzw. Ende einer Verbindung) unterrichtet, damit sowohl die Verkehrslenkung, als auch das Charging zu Abrechnungszwecken gesteuert werden kann. Zur Unterstützung von mobilen Szenarien ist eine Interaktion zwischen dem mobilen Kommunikationsnetz und dem Festnetz erforderlich. Hierfür wurde das *CAMEL*-Protokoll standardisiert. Für weiterführende Informationen zu *CAMEL* wird auf [4] verwiesen.

4.7.3.3 Benutzerlokation

Viele zukünftige Anwendungen werden ortsbezogen sein und dem Anwender eine ortsbezogene Personalisierung der Anwendung ermöglichen. Die *OSA*-Ortserkennung stellt eine Schnittstelle bereit, über die der aktuelle Aufenthaltsort eines Benutzers ermittelt werden kann. In *OSA* kann eine Anwendung den Aufenthaltsort eines bestimmten Benutzers zu einer beliebigen Zeit anfordern. Dies kann auch periodisch erfolgen.

4.7.3.4 Charging

Der *OSA* Charging-Dienst definiert Chargingoptionen für verbindungsorientiertes Charging ebenso, wie für so genannten inhaltebasiertes Charging. Die Grundfunktionen können sowohl für Prepaid, also auch für Postpaid-Abrechnungsschemata verwendet werden und bestehen aus folgenden zwei wesentlichen Vorgängen:

- Überprüfung des Kreditlimits eines Benutzers und

- Reservierung einer Gebühr bzw. Reduzierung des verbleibenden Budgets eines Benutzers.

4.7.3.5 Benutzerstatus

Der *OSA*-Benutzerstatus ermöglicht Anwendungen zu überprüfen, ob z.B. das Endsystem eines Benutzers am Netz registriert ist.

4.7.3.6 Sitzungskontrolle

Die *OSA*-Sitzungskontrolle beinhaltet sowohl die Steuerung und Verwaltung der Verkehrslenkung, als auch eine Ereignisverwaltung.

4.7.3.7 Endsystem-Charakterisierung

Anwendungen können die charakteristischen Eigenschaften des Endgerätes ausnutzen, um die benutzerspezifischen Informationen geeignet darstellen zu können. Dies sind, z.B. die Display-Eigenschaften des Endgerätes (Größe, Farbe, Auflösung usw.).

4.7.3.8 Präsenz und Erreichbarkeitsmanagement

Die Hauptziele des Präsenz und Erreichbarkeitsmanagements (*PAM*) ist die Vereinfachung von Anwendungsentwicklungen mit der Idee, dadurch eine große Anzahl von Anwendungen, die von so genannten Drittanbietern angeboten und ggf. betrieben werden, ins Netz zu bekommen. *PAM* zielt hierbei auf die Verwaltung dieser Anwendungen ab und standardisiert die notwendigen Verwaltungsabläufe, indem Erreichbarkeits- und Präsenzinformationen über Netzbetreibergrenzen hinweg ausgetauscht werden. Kernidee von *PAM* ist eine Vereinheitlichung der Begriffe Identität, Präsenz und Erreichbarkeit, die dann netzunabhängig, protokollunabhängig, technologieunabhängig und anwendungsunabhängig gültig sind.

4.7.3.9 Policy Management

Da davon auszugehen ist, dass *OSA*-Dienste zunehmend Mechanismen wie Zugangskontrolle und Benutzermanagement einsetzen, um den Zugang zu Ressourcen und Diensten zu verwalten, die ggf. von einem Drittanbieter dem Kunden angeboten werden, ist ein einheitliches Managementkonzept erforderlich. Dieses Managementkonzept regelt diese Ressourcenverteilung bzw. die Zugangsmechanismen zu den Ressourcen einheitlich.

OSA definiert eine Anzahl von Schnittstellen für Zugangsmanagement, welche das Verwalten der jeweiligen Regeln erleichtern. Zusätzliche Schnittstellen sind vorgesehen, um die Kommunikation zwischen Endanwendungen und spezifischen Diensten zu unterstützen, sodass der Benutzer diese Ereignisse überwachen kann. Alle Schnittstellen basieren auf einem einheitlichen, darunter liegenden so genannten Policy-Informationsmodell. Die policy-basierten Dienste beinhalten:

- Load balancing, um mittels dieser Policies Netzlasten zu verwalten.
- Charging-Dienste, die mittels Policies Charging-spezifische Kriterien bestimmen.
- Call Management-Dienste, die Policies einsetzen, um den Verbindungsaufbau zu steuern.

4.8 Die Moby Dick-Architektur

Nachdem nun in diesem Kapitel mehrere Mechanismen und Konzepte zur Kommerzialisierung des Mobilens Internets vorgestellt wurden, die alle mehr oder weniger stark in verschiedenen Architekturen verankert sind, soll in diesem Abschnitt eine komplette Kommuni-

kationsarchitektur vorgestellt werden, welche den Anspruch hat, die Kommerzialisierung des Mobilens Internets angemessen zu unterstützen.

Während zahlreiche aktuell eingesetzte Kommunikationsnetze noch mehr oder weniger stark auf Konzepten und Elementen der verbindungsorientierten Kommunikationsarchitektur basieren (dies bezieht sich sowohl auf die ISDN-basierten Netze, als auch auf 2G/3G-Netze), wird aktuell davon ausgegangen, dass sich traditionelle paketorientierte Netze mittelfristig durchsetzen und eine Plattform sowohl für die bisherigen paketvermittelnden Datendienste, als auch für Sprachdienste bilden. Dieses Evolutionsstadium wird aktuell in der Literatur mit B3G (Beyond 3G) bezeichnet.

Um mobile Netze der nächsten Generation so zu erweitern, dass die grundlegenden Herausforderungen wie Mobilitätsverwaltung, Ortserkennung, Benutzeridentifikation, Verkehrslenkung und Ressourcenkontrolle des drahtlosen Mediums geeignet unterstützt werden, sind noch einige Anstrengungen zu unternehmen, um basierend auf den existierenden Netzarchitekturen sowohl eine Rückwärtskompatibilität zum existierenden Internet, als diese neuen Dienste unter Berücksichtigung der Skalierbarkeit zu unterstützen.

Die stetig wachsende Bedeutung von IP als Konvergenzschicht deutet sich bereits in der aktuellen 3G-Architektur an, bei der IP in einer Tunnelform bereits unterstützt wird. Hier ist zwar eine erste Tendenz zu einem IP-dominierten Netz sichtbar, jedoch handelt es sich hierbei im Kern noch um ein verbindungsorientiertes Netz, auf dem leitungsvermittelnde Konzepte überlagert wurden.

Viele Netzbetreiber stehen nun vor der Herausforderung, paketorientierte Elemente weiter in eine konvergierende Netzplattform zu integrieren. Aktuell wird allgemein davon ausgegangen, dass diese Anstrengungen nur einen Zwischenschritt hin zu einem reinen IP-basierten Netz schreiben. Weiter besteht die weit verbreitete Meinung, dass alle hierfür notwendigen Mechanismen bereits zur Verfügung stehen und nur noch integriert bzw. teilweise angepasst werden müssen. Es gibt jedoch aktuell keine verlässlichen Informationen, die den hierfür erforderlichen Aufwand exakt beschreiben [81].

Im Rahmen von Moby Dick [39] wurde eine Architektur entwickelt, welche die bestehenden Mechanismen evaluierte und so integrierte, das sowohl qualitative als auch quantitative Aussagen bezüglich dieses Aufwandes getroffen werden können [81]. Hier ist die Fähigkeit, schnell und effektiv neue Dienste, die kurzfristig nachgefragt werden, in ein solches Netz integrieren zu können, eine Kernherausforderung für die Betreiber dieser zukünftigen Netzinfrastruktur. Diese Kernanforderung entsteht aus dem Druck, der auf dem Internet lastet, die konzeptionelle Ausrichtung mehr in Richtung eines etablierten Geschäftsmodells zu treiben. Dies wird ein existenzielles Kriterium für den Erfolg dieser zukünftigen Netzarchitektur sein. Basierend auf dieser Prognose wurde im Moby Dick-Projekt eine Architektur entworfen, die die Mechanismen von drei verschiedenen traditionellen Disziplinen der IETF geeignet auf einer heterogenen mobilen Netzplattform integriert. Es wurde erstmalig ausschließlich IP als Konvergenzschicht eingeführt, die das Internet-Protokoll mit explizitem Ausschluss jeglicher proprietären und anslusstechnologieabhängigen Elemente in einem heterogenen Zugangsnetz platziert. Dies beinhaltet jegliche Signalisierungs- bzw. Netzmanagementfunktionalität ebenso wie den Datenverkehr. Die drei Kerndisziplinen sind QoS zur Verwaltung der Ressourcen (als Ressource wurde im Rahmen dieses Projektes lediglich Bandbreite definiert), Mobilitätsmanagement, um dem Benutzer jeglichen Netzzellenwechsel vorzuenthalten und Authentisierung, Autorisierung, Accounting, Auditierung und Charging (A4C), welche die Grundlage für die Einführung kommerzieller Dienste schaffen.

Die im Projekt angegangenen Herausforderungen können folgenden Kategorien zugeordnet werden:

- **Mobilität:**
Hier wurde das Problem der Aufrechterhaltung eines existierenden Datenflusses angegangen, während ein Benutzer eine Netzzelle wechselt. Die gesamte Mobilitätsverwaltung wurde mit Konzepten und Mechanismen der Schicht 3 durchgeführt. Es wurde besonderes Augenmerk darauf gelegt, dass dieser Netzübergang technologieunabhängig und für den Benutzer transparent ist. Während Mobile IPv6 als allgemeines Protokoll zur Mobilitätsverwaltung über administrative Netzgrenzen nicht berücksichtigt werden kann, wurde in Moby Dick zwischen Mobilitätsszenarien zwischen administrativen Netzen und Mobilitätsszenarien innerhalb eines administrativen Netzes (aber ggf. zwischen verschiedenen Netztechnologien) unterschieden. Beide Szenarien wurden optimiert. Es wurde angenommen, dass Mobilität innerhalb eines Netzes sehr viel öfter auftritt, als der Netzwechsel zwischen verschiedenen Betreibern. Mobile IPv6 (vergl. Kapitel 2.1.4), bietet leider keine ausreichende Lösung für Mobilitätsverwaltung zwischen administrativen Bereichen und wurde somit nur für die Mobilitätsverwaltung innerhalb eines administrativen Netzes eingesetzt.
- **Benutzerverwaltung:**
Benutzerverwaltung ist immer unter zwei unterschiedlichen Blickwinkeln zu betrachten. Dies sind die betriebswirtschaftliche Sichtweise und die technologische Sichtweise. In der Regel wird in einem ganzheitlichen Benutzerverwaltungskonzept die betriebswirtschaftliche Sicht auf die technische Sicht abgebildet. Das Kernziel der rein technischen Benutzerverwaltung ist eine allgemeine Infrastruktur zu erstellen, die *A4C* geeignet realisiert. Diese Strategie muss vollständig kompatibel zum Mobilitätsmanagement sein und muss gleichzeitig Geschäftsmodelle und Preisschemata berücksichtigen. Eine enge Verzahnung mit den Mechanismen zur Ressourcenverteilung und Betrachtungen der Leistungsfähigkeit/Skalierbarkeit erhöhen hier die Gesamtkomplexität.
- **Ressourcenverteilung:**
Die Probleme, die es im Rahmen dieser Disziplin zu lösen gab, lassen sich in zwei Teile einteilen. Dies sind eine effektive Ressourcenverteilung auf dem Zugangsnetz und eine effektive Ressourcenverteilung auf dem Kernnetz. Dies erfordert ein geeignetes QoS-Konzept für den Ende-zu-Ende Datenfluss zwischen zwei Benutzern, von denen mindestens einer mobil ist. Hier ist eine Herausforderung Dienste zu definieren, die dann auf geeignete Dienstgüteklassen abgebildet werden. Eine weitere Herausforderung ist die entsprechende Schnittstelle zu einem Abrechnungsmodell (Charging).
- **Herausforderungen der Netztechnologie:**
Jegliche Netztechnologie verfügt hinsichtlich des Bandbreitenaspektes über charakteristische Übertragungskapazitäten. Hieraus folgt, dass die Kontrollfunktionen, welche von physikalischen Eigenschaften abhängen, von den Kontrollfunktionen getrennt werden, die von der IP-Schicht übernommen werden können.

Wenn man die oben aufgeführten Punkte isoliert voneinander betrachtet, repräsentiert jeder dieser Punkte technologische Herausforderungen. Versucht man diese miteinander mit Konzepten und Mechanismen der Schicht 3 zu integrieren, erhält man den Rahmen für eine Architektur für eine zukünftige Netzplattform für Mobilkommunikation.

Im Moby Dick-Projekt wurde, basierend auf den beschriebenen Vorgaben, eine Architektur entwickelt, prototypisch in ihren Kernfunktionen implementiert und demonstriert, welche in Abbildung 39 aus konzeptioneller Sicht dargestellt ist.

Generell besteht ein Netz aus verschiedenen administrativen Domänen. Eine administrative Domäne ist der Teil des Netzes, der zu einem Netzbetreiber gehört, der prinzipiell gleichzeitig Netzzugang über mehrere Netztechnologien anbietet. Die verschiedenen Netzbetreiber unterhalten weiter Vereinbarungen untereinander, um Kunden des jeweiligen anderen Betreibers Ressourcen des eigenen Netzes für Mobilitätsszenarien anzubieten (Roaming Agreements). Das Moby Dick-Kernnetz unterstützt verschiedene Dienstgütern. Konkret unterstützen die Zugangsrouten eine Differenzierung der Nachrichtenströme nach den in Kapitel 2.1.2.3 vorgestellten Konzepten der IETF (Differentiated Services). In Abbildung 39 sind drei verschiedene Domänen abgebildet:

- Domäne A: stellt nur WCDMA-Zugang über die jeweiligen Zugangsnetze bereit
- Domäne B: stellt sowohl WCDMA als auch WLAN (802.11) über die jeweiligen Zugangsnetze bereit
- Domäne C: stellt WLAN und 802.3 Ethernet bereit

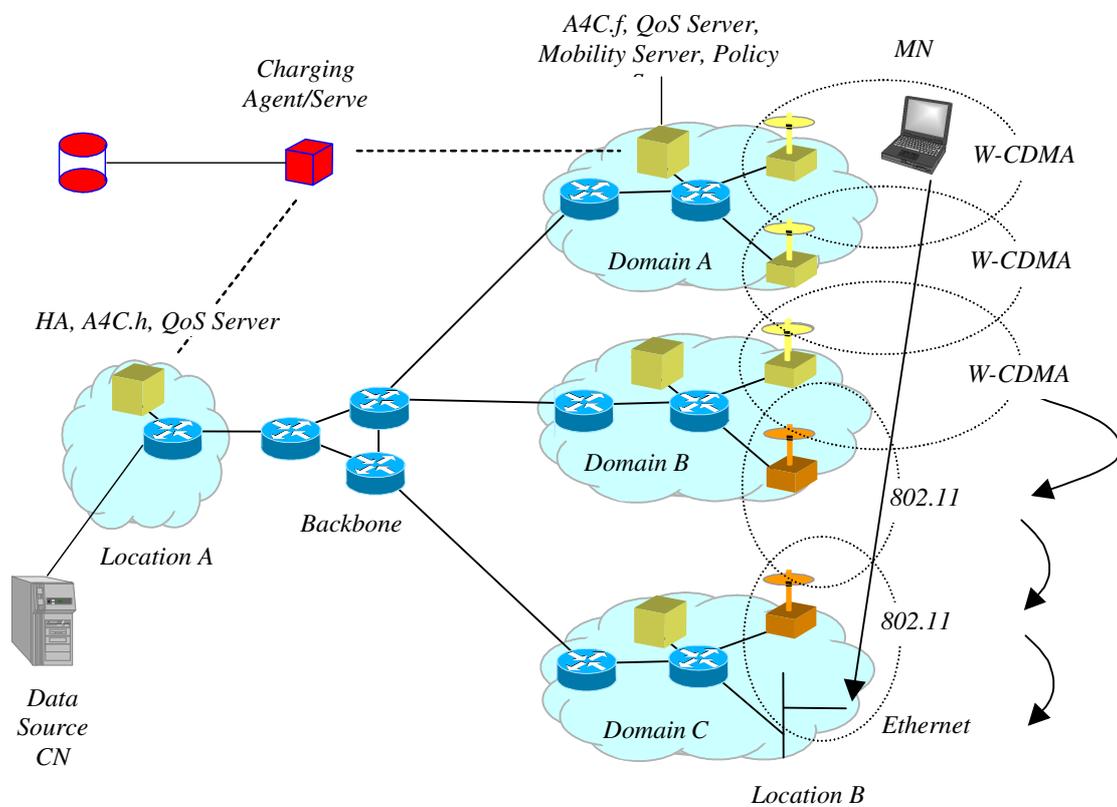


Abbildung 39: Moby Dick-Architektur

Der Correspondent Node (CN) ist über eine zunächst unbekanntere Dienstgüteunterstützende Zugangstechnologie an das Kernnetz angebunden. Der Heimagent (HA), der Heimat-QoS-Broker und der A4C.h-Server sind innerhalb einer administrativen Domäne platziert. Für die Verwaltung und den Betrieb eines Zugangsnetzes in Moby Dick sind folgende Komponenten vorgesehen:

- Ein QoS-Broker, der für die Verwaltung und das Ausführen von QoS-spezifischen Anforderungen erforderlich ist. Diese Anforderungen werden über so genannte QoS-Profilen, die benutzerspezifisch sind, signalisiert.
- Zugangsrouten: Hierbei handelt es sich um Netzknoten, die verschiedene Zugangstechnologien auf Kernnetzseite abschließen. Die Technologien sind in Moby Dick entweder die beiden drahtlosen Varianten WCDMA und WLAN (802.11) oder Ethernet

(802.3). Konzeptionell könnte jedoch jede weitere Netztechnologie mit bidirektionaler Kommunikation eingebracht werden. Jeder Zugangsrouten deckt ein IP-Netz ab, welches er unterhält und stellt die Verbindung aus dem drahtlosen Bereich in das Kernnetz her.

- A4C-Server für die Abwicklung von Authentisierung, Autorisierung, Accounting, und Charging.

Optional für den Zugangsbereich sind Paging-Agenten (PA). Ein PA ist außer für die eigentliche Funktion (Paging) außerdem für eine effizientere Verwaltung des Leistungsverbrauchs eines mobilen Endgeräts verantwortlich. Das Mobile Endsystem wird in eine Art Schlafzustand versetzt und vom PA bei Bedarf geweckt.

4.8.1.1 QoS in Moby Dick

Durch die konsequente Nutzung von IPv6 auf der Vermittlungsschicht konnte auf 3G-Netzkomponenten *SGSN*, *GGSN*, usw. (siehe Kapitel 2.2.2.4) verzichtet werden. Der *PDP-Kontext* Mechanismus wurde somit signifikant vereinfacht. Hier konnte vor allem auf die Signalisierung zwischen Mobilem Endsystem und dem *SGSN* verzichtet werden. Das Transformieren zwischen UMTS QoS-Attributen auf DiffServ-Attribute erfolgte nach einem spezifizierten Schema.

Tabelle 4: Abbilden von DiffServ QoS auf UMTS QoS in Moby Dick

IP-Attribute	PHB	EF	AF1	AF2	BE
	Expected forwarding service	Niedere Latenz, Jitter und Paketverlust	Latenz L Jitter J Paketverlust P	Mittlere Latenz, Jitter und Paketfilter P	N/A.
Verkehrsprofil	Spitzenrate R MTU M	Spitzenrate R Mittlere Rate r maximale Burstlänge B	Spitzenrate R Mittlere Rate r maximale Burstlänge B	Spitzenrate R	
UMTS-Attribute	Verkehrsklasse	Conversational	Streaming	Streaming	Background
	Maximale Bitrate (kbit/s)	R (limitiert auf 2048)	R (limitiert auf 2048)	R (limitiert auf 2048)	R (limitiert auf 2048)
	Reihenfolge-sicherung	nein	nein	nein	Ja (TCP) nein (UDP)
	Maximale SDU Größe (octet)	M (begrenzt auf 1500)	s() (1)	1500	1500
	SDU Fehlerrate	10^5	nahe bei P	nahe bei P	10^3
	SDU Format-information	IP	IP	IP	N/A.
	Zustellung korrupter SDU	ja (Sprache) nein(Daten)	ja	Ja	nein
	Ende-zu-Ende Verzögerung (ms)	100	nahe bei L	250	N/A.
	Garantierte Bitrate (kbit/s)	R (limitiert auf 2048)	$g()$ (limitiert auf 2048) (3)	$g()$ (limitiert auf 2048) (3)	N/A.
	Traffic handling priority	N/A.	N/A.	N/A.	N/A.
	Systempriorität	1	2	2	3

In Moby Dick wurden die IETF DiffServ-Klassen, wie in Kapitel 2.1.2.3 vorgestellt, auf die QoS-Klassen der UMTS-Architektur abgebildet [2]. Tabelle 4 fasst die im Moby Dick-Projekt spezifizierten Abbildungen zusammen. Die generische QoS-Architektur ist in Abbildung 40 dargestellt.

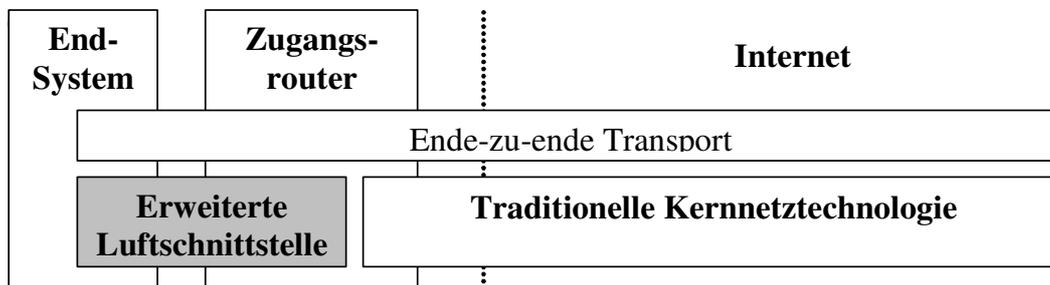


Abbildung 40: Generische Moby Dick QoS-Architektur

4.9 Existierende Vorschläge für Internet-Preismodelle

Nachdem nun ein Überblick über die verfügbaren technischen Mechanismen und Konzepte zur Kommerzialisierung des Mobilens Internets gegeben wurde, soll nun ein Überblick über existierende Preismodelle gegeben werden, welche im weiteren Verlauf dieser Arbeit relevant sind.

Bei der Betrachtung der verschiedenen Preismodelle soll zunächst festgehalten werden, dass die aktuell implementierte Preispolitik, welche die einst staatlichen Telekommunikationsbetreiber aus einer Monopolstellung heraus am Markt durchsetzen konnten, im Kontext dieser Arbeit überdacht werden muss. Dies wurde ausführlich in [69] diskutiert. Zusammengefasst kann festgehalten werden, dass der Parameter *Verbindungsdauer* in einem Preismodell für die jüngst von den Benutzern bevorzugten Kommunikationsszenarien in einem Mobilens Internet ungeeignet ist. Der Grund hierfür ist, dass sich asynchrone Arbeitsszenarien wie WWW, E-Mail usw. weiterer Akzeptanz erfreuen und synchrone Szenarien relativ an Bedeutung verlieren. Durch den Charakter des Internets kann der Parameter Verbindungsdauer keinerlei Rückschlüsse über Ressourcenbelegung liefern und ist somit als Mittel der Nachfragesteuerung ungeeignet. Weiter ist dieser Parameter nicht mit dem „Always-On“ Prinzip des Internets vereinbar und es darf bezweifelt werden, dass die Benutzer mittelfristig einen solchen Parameter akzeptieren bzw. dieser sich am Markt durchsetzen wird. Parameter, die sich nur auf Datenmenge bzw. Bandbreite beschränken, sind nicht geeignet, die Nachfrage derart zu steuern, dass die Ressourcen effektiv ausgelastet sind. Zusätzlich müssen der Aufwand für das implementierte Preismodell und die damit verbundenen Kosten vertretbar sein. Diese Aspekte wurden auch schon in [17] aufgenommen und teilweise implementiert.

Es soll an dieser Stelle jedoch festgehalten werden, dass sich die „historisch verbindungsorientierten“ Parameter vorzüglich eignen, um das Kommunikationsnetz kommerziell erfolgreich zu betreiben. Folglich werden die darüber liegenden Geschäftsmodelle, welche sich strukturell der leitungsvermittelnden Parameter bedienen, von den Netzbetreibern nur ungern aufgegeben.

Die Einführung einer Flat-Rate ist aus Sicht des Aufwandes, wie im weiteren Verlauf dieser Arbeit noch diskutiert wird, sehr attraktiv, weil hier die Verwaltungskosten und somit die Kosten für die Netzinfrastruktur minimiert werden. Es wird keine aufwändige Messtechnik benötigt, die den aktuellen Grad der Netzstauung ermittelt. Ferner wird keinerlei Management benötigt, welches aus dem Grad der Netzstauung Preise ermittelt, die dann dem Benutzer angeboten werden. Zusätzlich müsste in einem solchen Fall die Rückmeldung des Benutzers

wieder ins Netz integriert werden, was mit zusätzlichem Aufwand verbunden ist. Die Flat-Rate ist jedoch aus Sicht der Nachfragesteuerung völlig ungeeignet. Es hat sich gezeigt, dass mit dem Gut Bandbreite im Falle eines Flat-Rate Gebührenmodells wenig effektiv umgegangen wird und die Möglichkeit über den Preis Nachfrage quasi zur Unterstützung der Netzplanung zu steuern, verloren geht.

Diese Problematik wurde als Forschungsthema aufgegriffen und es wurden verschiedene Alternativen erarbeitet, von denen nun eine exemplarische Auswahl vorgestellt werden soll. Prinzipiell lassen sich diese Vorschläge in zwei Gruppen einteilen. Einmal sind dies Vorschläge, welche unter Inkaufnahme eines hohen Aufwandes eine möglichst effektive Steuerung der Ressourcenvergabe erreichen, bei der die Kosten für das Netz minimiert werden; weiter gibt es Vorschläge, bei denen diese hohen Aufwände gescheut werden mit dem Nachteil, dass keine umfassende Nachfragesteuerung ermöglicht wird. Das ideale Preismodell hat bei minimalem Aufwand die maximale Einflussnahme auf die Nachfragesteuerung zu leisten. Folgende Preismodelle wurden vor diesem Hintergrund diskutiert:

- Flat-Rate
- Paris Metro Pricing
- Smart Market
- Verbindungsorientiertes Preismodell
- Expected Capacity Allocation
- Priority Pricing
- Responsive Pricing

In [69] wurden die vorgeschlagenen Preisschemata im Rahmen einer Effektivitäts/Aufwandsmatrix relativ zueinander bewertet. Das ideale Preisschema hat eine hohe Effektivität bezüglich des Ressourcenverbrauches bei gleichzeitig sehr geringen Kosten für die Implementierung bzw. des Gesamtaufwandes, welcher die durch das Management belegten Ressourcen, die zusätzlich benötigte Messtechnik und jeglichen Aufwand bezüglich zusätzlicher Datenhaltung umfasst. Ein schlankes und somit kostengünstiges Preisschema hat den Nachteil, dass das Benutzerverhalten und somit die Nachfrage nicht effizient gesteuert werden kann. Auch dies hat zur Folge, dass die Kosten für Netzaufbau steigen. Abbildung 41 zeigt das Ergebnis dieser Bewertung.

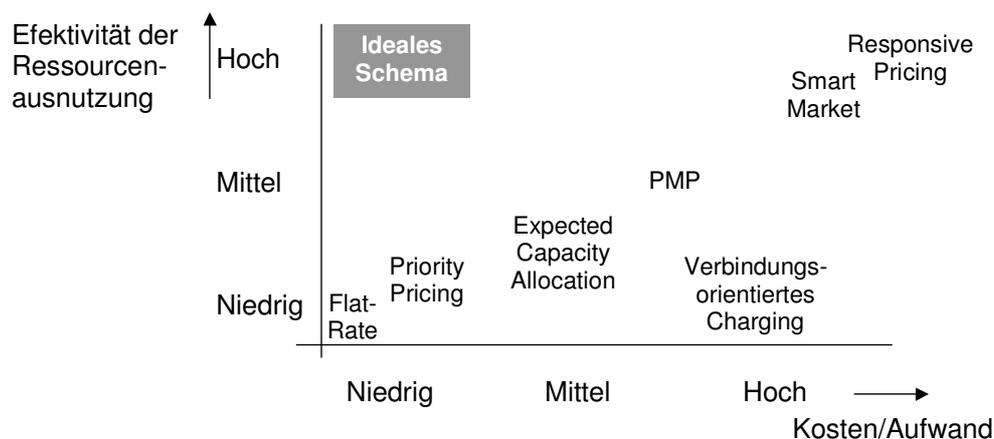


Abbildung 41: Bewertung der Preismodelle aus [69]

4.9.1 Flat-Rate

Die Flat-Rate bringt sehr geringe Verwaltungskosten mit sich. Es muss hier keine gebrauchsspezifische Datenverwaltung installiert und unterhalten werden. Weiter ist keine aufwändige Messtechnik erforderlich und es muss lediglich ein statisch fixierter Zahlungseingang über-

wacht werden. Da hier beim Kostenaufwand nur qualitative Aussagen getroffen werden können, werden die Kosten im Folgenden fiktiv auf die Kosten der Flat-Rate bezogen.

Der Nachteil der Flat-Rate ist, dass sie zu übermäßigem Ressourcenverbrauch führt, und keinerlei Möglichkeit besteht, die Netzkosten zu minimieren bzw. die Nachfrage entsprechend netzoptimierend zu steuern.

4.9.2 Prioritätsbasierte Preisbildung

Dieses Schema wurde in [45] vorgestellt und ist aus Sicht des Aufwandes für die Abrechnungs-Infrastruktur dem Schema der Flat-Rate ähnlich. Kernidee ist das Etablieren von verschiedenen aggregierten Prioritätsklassen, für deren Benutzung dann jeweils eine gestaffelte "Flat-Rate" erhoben wird. Hierfür müssen die Datenpakete mit Prioritäten versehen werden. Dies wurden in den letzten Jahren innerhalb der IETF Differentiated Services-Arbeitsgruppe standardisiert (siehe Kapitel 2.1.2.1). Somit hat dieses Preisschema einen leicht erhöhten Aufwand im Netzmanagement, da hier lediglich zusätzliche Konfigurations- und Netzplanungsarbeiten vorgenommen werden müssen. Bezüglich der Nachfragesteuerung hat dieses Schema ebenfalls nur leichte Vorteile gegenüber der Flat-Rate. Es kann zwar durch geeignete Prioritätenvergabe ein Teil der Netzstauungen vermieden werden, die Gleichstellung der Benutzer verhindert jedoch aus preispolitischer Sicht nicht den Effekt der wenig effektiven Ressourcenallokation. Ferner ist hier kein Eingriff in die Nachfragesteuerung mit Hilfe von preispolitischen Mitteln erforderlich. Elemente dieses Preisschemas wurden in [18] erfolgreich eingeführt.

4.9.3 Kapazitätsallokation

Clarks Kapazitätsallokation ermöglicht einem Benutzer erstmals aktiv auf die Dienstgüte Einfluss zu nehmen [29], welche sich dann im zu entrichtenden Preis widerspiegelt. Da diese Preise jedoch relativ statisch sind, ist die Netzdimensionierung ein wesentlicher Erfolgsfaktor. Jedoch wird erwartet, dass die Netzplanung in einer mobilen Umgebung wesentlich komplexer ist. Zusätzlich würde eine Realisierung dieses Ansatzes zu einem erhöhten Aufwand an Netzinfrastruktur zur Erfassung der Verkehrsdaten führen. Das Verkaufen von Kapazitätsbudgets erfordert, dass eine Architektur entwickelt und implementiert wird, die diese Budgets ermittelt und zu den Stellen im Netz übermittelt, an denen sie konsumiert werden. Bezüglich der Nachfragesteuerung könnte die Nachfrage durch geeignetes Festsetzen der Preise beeinflusst werden. Jedoch gibt es keinerlei Elemente im Netz, welche die garantierten Kapazitäten entsprechend der aktuellen Nachfrage mit geeigneten Preisen versehen. Somit ist der Mechanismus zur Steuerung der Nachfrage nur langfristig unter Berücksichtigung des Gesamtverkehrsaufkommens durchführbar. Die Randbedingung, dass permanent der Empfänger der Pakete für die Gesamtkosten aufkommt, ist im Falle eines TCP-Verkehrs zwischen Benutzer und Web Servers eine durchaus geeignete Möglichkeit. Da aber z.B. ein Web Server in diesem Falle über die vertraglichen Details zwischen Benutzer und Netzbetreiber in Kenntnis gesetzt, und diese Kommunikation über Providergrenzen hinweg speziell signalisiert werden müsste, ist der Aufwand hierfür doch erheblich. Weiter wäre dies ein Verstoß gegen das Paradigma des „Anrufer-Bezahlt-Prinzips“, das sich in Europa etabliert hat. Ob dieses Paradigma aufgegeben werden kann und welche Folgen dies gegebenenfalls auf das Benutzerverhalten hätte, ist eine interessante Frage, würde aber den Rahmen dieser Arbeit sprengen und wird somit nicht weiter verfolgt. Im Falle von UDP-Verkehr, also jeglichen zukünftigen Telefonieverkehr, ist hier erheblicher Signalisierungsbedarf gegeben. Dieser zusätzliche Aufwand in Form von Verkehr über das Netz ist allerdings beträchtlich.

4.9.4 Paris Metro Pricing

Das „Paris Metro Pricing“ (*PMP*) geht von einer festen Teilung des Marktes innerhalb eines administrativen Bereiches aus. Odlyzko schlägt vor, den Markt in zwei gleiche Teile zu teilen und entsprechend unterschiedliche Preise für die Benutzung der einzelnen Ressourcen zu erheben ([98], [99] und [100]). Die Kernidee dieses Vorschlages ist das Etablieren eines Hochpreissegmentes, ohne das dieses Segment mit zusätzlichen qualitativen Leistungen versehen ist. Hier wird also eine Übervorteilsstrategie nach Kapitel 3.2.2 vorgeschlagen. Nur durch die Tatsache, dass sich hier einige Benutzer von diesem Hochpreissegment abwenden, können die Benutzer, die dieses Segment in Anspruch nehmen, mit einer erhöhten Wahrscheinlichkeit befriedigt werden. Die Motivation eines Benutzers, den höheren Preis zu bezahlen, ist nur dann gegeben, wenn er davon in Kenntnis gesetzt wurde, dass das Kontingent, das zum niedrigeren Preis verkauft wurde, bereits ausgeschöpft ist. Ein solches System etabliert quasi zwei Märkte, die beide von einem Unternehmen kontrolliert werden und ist nur dann in der Lage auf die Nachfrage steuernd Einfluss zu nehmen, wenn sich die Benutzer permanent zwischen den beiden Märkten entscheiden können. Odlyzko schlägt vor, maximal vier verschiedene Märkte zu etablieren.

Da es sich bei dem drahtlosen Zugangsmedium um ein grundsätzlich sehr knappes Gut handelt, soll nun kurz eine überschlägige Betrachtung bzgl. des Ressourcenverbrauches der erforderlichen Signalisierung durchgeführt werden:

Es wird von einer Nachrichtenlänge von 100 byte und einer Nachricht pro Sekunde ausgegangen. Dies entspricht einer Bandbreite von 800 bit/s und erscheint zunächst als vertretbarer Aufwand (0,4 % der Gesamtkapazität von 2 Mbit/s). Zusätzlich kommen nun noch die Informationen der Benutzer hinzu, sich für den einen oder anderen Markt zu entscheiden. Dies hängt von der absoluten Anzahl der Benutzer und ihres Verkehrsprofils ab. Befinden sich wenige Benutzer im Netz, welche dies stark nutzen, ist der Aufwand für zusätzlichen Kommunikationsbedarf deutlich geringer als im Falle sehr vieler Benutzer mit sporadischem Netzzugriff, da diese mit einer höheren Wahrscheinlichkeit die Dienstklasse wechseln. Bei der Annahme, dass ein Benutzer pro Sekunde einen solchen Dienstklassenwechsel anzeigt, sind weitere 800 bit/s Bandbreite für Netzmanagement belegt. Somit ergibt sich ein zusätzlicher Bedarf von knapp einem Prozent der maximal zur Verfügung stehenden Ressourcen. Dieser Aufwand ist relativ gering. Bezüglich der Nachfragesteuerung ist anzunehmen, dass die Benutzer auf den erhöhten Preis reagieren und es würden sich langfristig zwei Märkte etablieren, für die es unterschiedliche Nachfragefunktionen gibt. Da jeder einzelne Benutzer eine individuelle maximale Zahlungsbereitschaft hat, kann somit durch geeignete Preisfestsetzung die Nachfrage gesteuert werden.

4.9.5 Smart Market

Der „Smart Market“ wurde in [86] vorgestellt und ermöglicht einem Benutzer eine permanente Partizipation an einem Auktionsverfahren, bei dem (im Kontext dieser Arbeit) Netzressourcen versteigert werden sollen. Hiermit soll die Nachfrage geeignet gesteuert werden. Das Ermitteln eines Auktionspreises beim Smart Market wird gemäß den Regeln der Vickery-Auktion [130] festgelegt (vergleiche Kapitel 3.2.3.1). Die Kosten, welche eine Einführung einer Vickery-Auktion ins Kommunikationsnetz verursachen würde, wären erheblich. Darüber hinaus muss an dieser Stelle festgehalten werden, dass die Kommunikation der Gebote/Preise die zum Verkauf anstehende Ressourcen reduziert, da diese Kommunikation ebenfalls über das Netz transportiert werden müsste und linear mit der Zahl der Benutzer in einer Zelle steigt. Ein weiterer erheblicher Aufwand ist die zeitliche Synchronisation der Gebote. Wie bei jeder Auktion gibt es ein Datum, zu dem die Auktion stattfindet und zu der sich die Bieter entsprechend versammeln. Beim einem Smart Market, der Internet Bandbreite versteigert,

müsste permanent das Auktionsdatum kommuniziert werden. Da sich die Bieter nicht an einem Punkt versammeln können, müssen die Gebote in Form einer Signalisierung übermittelt werden. Dies wurde bei der Delta-Auktion ([41] und [110]), einem weiteren Auktionsverfahren, gelöst, jedoch ist die Umsetzung dieses Auktionsverfahrens mit nennenswertem Aufwand in einem Mobil Internet verbunden.

Generell ist die Ressourcenbelegung nicht deterministisch und hängt vom Verhalten aller sich im drahtlosen Netz befindlicher Benutzer ab. Dies hat zur Folge, dass eine solche Versteigerung permanent durchgeführt werden muss, da die Daten, auf deren Basis die Preise gebildet wurden, keinerlei Schlüsse über die Nachfrage im nächsten Zeitintervall zulassen, sondern lediglich eine Feststellung der Ressourcenbelegung im vergangenen Zeitintervall sind.

Abschließend lässt sich festhalten, dass der Smart Market zwar auf den ersten Blick einen sehr geeigneten und plausiblen Mechanismus zur Nachfragesteuerung darstellt und sich generell Auktionsverfahren in vielen vergleichbaren Märkten bewährt haben. Eine Implementierung im mobilen Internet ist wegen der hohen zeitkritischen Anforderungen jedoch mit erheblichem Aufwand verbunden. Noch offen ist das Problem der Behandlung von TCP-Verkehr (vergl. [27]). Hier müssten die Auktionsinstanzen den Datenverkehr in beide Richtungen in der Auktion berücksichtigen. Dies wäre mit weiterem Aufwand verbunden.

4.9.6 Responsive Pricing

Das Responsive Pricing [84] ist ebenfalls ein Preisschema, das dem Benutzer einen hohen Grad an Interaktion mit dem Netz ermöglicht. Hier ist vorgesehen, dass der Benutzer aktiv auf Basis der Entscheidungen des Auktionators agieren kann. Dies ist sowohl sitzungsbasiert als auch in einem langfristigen Rahmen vorgesehen. Hierzu ist ein im Vergleich zum Smart Market erhöhter Aufwand erforderlich, da den Benutzern permanent die Entscheidungen des Auktionators übermittelt werden müssen. Dieses Verfahren ähnelt sehr der Englischen Auktion (vergleiche Kapitel 3.2.3.1 und [69]) und hat den Nachteil, dass die Anzahl der Iterationsschritte, welche erforderlich sind, um zu einer Preisfindung zu kommen, unbestimmt sind.

Die Kosten für die Implementierung des Priority Pricing wären sehr hoch und es darf bezweifelt werden, ob die technische Realisierung im Netzbereich unter Berücksichtigung der Skaleneffekte umsetzbar ist. Da Rückmeldungen, zum Endkunden übertragen werden müssten, würde sich die zum Verkauf stehende Menge an Ressourcen erheblich reduzieren. In einem System mit 100 mobilen Teilnehmern in einer Zelle welches jedem Teilnehmer sekundlich eine Information über den aktuellen Stauungsgrad in Form eines aktualisierten Preises sendet, hätte ein Verkehrsaufkommen bei einer angenommenen Nachrichtenlänge von 100 byte nach Tabelle 5.

Tabelle 5: Ressourcenverteilung bei 50 byte Rückmeldung (1 s Rückmeldungsfrequenz)

PAKETLÄNGE	TEILNEHMER	SENDE-FREQUENZ	RESSOURCENVERBRAUCH
50 byte	100	1 s	400 kbit/s

Das Beispiel basiert auf angenommenen Werten. Die Nachrichtenlänge von 50 byte orientiert sich an vergleichbaren Diensten. So ist z.B. das DNS-Paket [122] vergleichbar groß. Der angenommene Wert von 100 Teilnehmern erscheint auf den ersten Blick als zu hoch. Unter Berücksichtigung der Tatsache, dass durch das „Always-On“ Prinzip des Internet diese Information nicht nur zu den aktiven Teilnehmern im Bereich der Sendezelle gesendet werden müsste, sondern zu jedem Teilnehmer, der sich im Einzugsbereich der Sendezelle befindet und ein potentieller Kandidat für eine Ressourcenallokation im nächsten Zeitintervall ist, kann

dieser angenommene Wert eher als zu konservativ betrachtet werden. Im obigen Beispiel wären 20 % der verfügbaren Ressourcen für die Realisierung des Auktionsverfahrens belegt und könnten nicht den Mobilendkunden zum Verkauf angeboten werden. Es darf bezweifelt werden, ob Netzbetreiber ein solches Preissystem implementieren würde.

4.10 Zusammenfassung und Beschreibung des weiteren Vorgehens

In diesem Kapitel wurden existierende Konzepte und Mechanismen vorgestellt, welche aktuell zur Verfügung stehen, um das Mobile Internet in ein kommerzialisierendes Mobiles Internet zu transferieren. Es wurde weiter die Moby Dick-Architektur vorgestellt, welche ein Konzept für eine Architektur für ein Mobiles Internet liefert, welches jegliche Mobilitätsverwaltung und das kommerzielle Dienstmanagement auf Schicht 3 löst. Diese Architektur verinnerlicht den aktuellen Trend zum Internet als Konvergenzschicht weitaus stärker als alle anderen Ansätze der relevanten Standardisierungsgruppen. Die Moby Dick-Architektur soll für den weiteren Verlauf dieser Arbeit als Referenzarchitektur betrachtet werden. Die Einführung von A4C in Moby Dick ist ein wesentlicher und notwendiger Schritt in Richtung kommerzialisierendes Mobiles Internet. Jedoch sind diese Mechanismen nicht ausreichend, wie im nächsten Kapitel noch detaillierter erarbeitet wird. Generell kann festgehalten werden, dass der Kommerzialisierungsgedanke im Internet erst in jüngster Zeit, also über 30 Jahre nach seiner ersten Konzeptionsphase, an Bedeutung gewinnt.

Weiter sind die vorgestellten verfügbaren Mechanismen und Konzepte der existierenden mobilen und festen Sprachnetze nicht geeignet, um ein Mobiles Internet in ein kommerzielles Netz zu transformieren bzw. das existierende 3G-Netz in Richtung eines ALL-IP-Netz im Sinne von Moby Dick weiterzuentwickeln. Die im letzten Teil dieses Kapitels vorgestellten theoretischen Konzepte zur Preisbildung verfolgen zwar das gemeinsame Ziel, das Internet zu kommerzialisieren, jedoch darf bei einigen bezweifelt werden, ob sie sich als technisch realisierbar erweisen.

Im folgenden Kapitel sollen nun Annahmen getroffen werden, welchen den Rahmen dieser Arbeit einschränken. Weiter wird detailliert aufgezeigt, wo die existierenden Mechanismen im Kontext dieser Arbeit als nicht ausreichend im Sinne der in Kapitel 3 aufgezeigten kommerziellen Aspekte betrachtet werden können und sich daraus weitere Anforderungen an die Infrastruktur ergeben.

5 Annahmen und Problemdiskussion zum Mobilem Internet

5.1 Einleitung

In diesem Kapitel werden Annahmen getroffen, welche dem Rest dieser Arbeit den geeigneten Rahmen bietet, um das Mobile Internet zu einer kommerziellen Netzplattform ausbauen zu können, welche das existierende 3G-Netz in allen Belangen ersetzt. Im Rahmen des Moby Dick-Projektes wurde bereits ein erster Schritt in diese Richtung unternommen und eine solche Netzarchitektur erfolgreich entwickelt und demonstriert. Diese Architektur wird für den weiteren Verlauf dieser Arbeit als Referenzarchitektur für ein zukünftiges Mobiles Internet vorausgesetzt.

Im Moby Dick-Projekt wurden keinerlei Überlegungen angestellt, wie ein Betreiber des Mobilem Internets Dienste kommerziell anbieten kann. Es wurden zwar technische Mechanismen auf Basis von IETF AAA entworfen, die einen solchen Kommerzialisierungsprozess technisch unterstützen [128], jedoch ohne umfangreiche betriebliche Überlegungen anzustellen, wie sie u.a. auch in Kapitel 3 gemacht wurden.

In diesem Kapitel wird ein exemplarisches Preismodell entworfen, welches aus [69] abgeleitet ist und Bezug auf die in Kapitel 4.9 vorgestellten Vorschläge von Internet-Preismodellen bzgl. ihrer Implementierbarkeit nimmt.

Weiter soll in diesem Kapitel aufgezeigt werden, wo existierende Konzepte, Methoden, Mechanismen nicht ausreichen, um ein Mobiles Internet, unter kommerziellen Aspekten betreiben zu können. Diese weitere Diskussion fokussiert sich dann auf den Bereich der Abrechenbarkeit kommerzieller Dienste aus Sicht eines Mobilem Internet-Betreibers.

5.2 Evolution und Konvergenz

In der Wirtschaft gibt es prinzipiell zwei Möglichkeiten Unternehmen zu vereinigen. Entweder ein Unternehmen wird von einem anderen Unternehmen übernommen, welches dann die eigene Philosophie in einer recht „undemokratischen“ Form dem anderen Unternehmen aufzwingen kann, oder aber es gibt den so genannten „*Merger of Equals*“, also einen Zusammenschluss zweier gleichberechtigter Partner, bei dem versucht wird, aus beiden etablierten Kulturen und Unternehmensphilosophien Synergien zu erzeugen. In der Moby Dick-Architektur, welche aktuell im Nachfolgeprojekt Daidalos [33] verfeinert und erweitert wird und die Funktionalität Benutzerverwaltung, Mobilitätsverwaltung und das Dienstgütemanagement auf die Vermittlungsschicht (Mobile IPv6) hebt, werden jegliche Zwischenstände des Migrationsprozesses nicht berücksichtigt. Das Internet nimmt technisch eine absolut dominierende Form ein und verdrängt jegliche Elemente der traditionellen Mobilkommunikation vollständig. Diese Annahme geht weit über die in Kapitel 2.2.2.4 und Kapitel 2.2.2.5 vorgestellten ALL-IP-Ansätze hinaus. Da jedoch das Internet unter dem Aspekt der Kommerzialisierung im Vergleich zu dem nun verschwindenden POTS-basierten Telefonnetz als rückständig bezeichnet werden kann, muss zunächst im Detail evaluiert werden, wo bzw. wie Mechanismen aus Kapitel 3 angemessen umgesetzt werden können. In jedem Fall wird angenommen, dass Geschäftsmodelle, wie aktuell im Telekommunikationsmarkt etabliert, aus kommerzieller Sicht das Mobile Internet verändern werden.

Generell kann davon ausgegangen werden, dass eine Prognose bezüglich des Endstadiums eines Migrationsprozesses mit sehr viel Ungenauigkeit behaftet ist, jedoch soll mittels Ver-

gleichen aus anderen Bereichen des täglichen Lebens die Plausibilität der hier getroffenen Annahmen untermauert werden.

5.3 Aspekte einer kommerziellen B3G-Infrastruktur

Nun sollen unterschiedliche Aspekte der Kommerzialisierung des Mobilens Internets eingeführt werden, um einen verbreiterten Blickwinkel zu gewinnen, unter dem die später vorgestellten Elemente dieses Prozesses eingeordnet und bewertet werden können. Die gewählten Sichtweisen orientieren sich an den Bedürfnissen aller vom Prozess der Kommerzialisierung des Mobilens Internets betroffenen Akteure. In diesem Zusammenhang soll erwähnt werden, dass z.B. in [74] Schwachpunkte bei der Kommerzialisierung des Internets aufgezeigt wurden, diese jedoch nicht das Mobile Internet bei seiner Migration mit dem 3G-Netz entsprechend berücksichtigt haben.

5.3.1 Benutzerorientierte Sichtweise

Aus Sicht des Benutzers ist der Aspekt „Benutzerakzeptanz“ ein bedeutendes Kriterium. Dieser Aspekt kann in verschiedene Dimensionen eingeteilt werden. Zunächst ist die Akzeptanz der Schnittstelle zwischen dem Mensch und dem IT-System wesentlich. Diese Schnittstelle soll einfach und verständlich gehalten werden. Es hat sich in verschiedenen Marktforschungen gezeigt, dass ein zu komplexes System, welches vom Benutzer/Verbraucher nicht verstanden wird, nicht angenommen wird. Die Stellen, an denen Transparenz aus Benutzersicht gefordert wird, ist die Mensch-Maschine Schnittstelle, in der praktisch alle technischen Parameter abstrahiert werden sollten. Dies ist vor allem vor dem Hintergrund einer überalternden Gesellschaft ein Kernkriterium, denn nur wenn dieser neuen Technologie der gesamten Gesellschaft zugeführt werden kann, hat das kommerzielle Internet eine tragfähige Zukunft.

Ein weiteres Kriterium neben der Schnittstelle „Mensch-Maschine“ ist die Transparenz des Preissystems. Ein zukünftiges Mobiles Internet soll sowohl die traditionellen Kunden der mobilen sprachbasierten Netze ansprechen, als auch die Kunden des Internets, dessen Datenverkehrsaufkommen heute das der Sprachkommunikation übersteigt. Hier ist der Konflikt zu lösen, dass sprachbasierte Netze auf einem vom Benutzer akzeptierten Geschäftsmodell basieren, diese Parameter jedoch nicht einfach auf das Internet übertragbar sind. Die technischen Parameter, die zur Abrechnung in sprachbasierten Netzen herangezogen werden (Gesprächsdauer, Tageszeit, Entfernung), wurden in Kapitel 3.5 ausführlich erarbeitet und konnten sehr einfach aus einer monopolistischen Situation heraus am Markt etabliert werden. Diese Parameter eignen sich jedoch nicht, den Internet-Benutzer geeignet anzusprechen. Weiter kann basierend auf diesen Parametern das Benutzerverhalten nicht ressourcenoptimiert gesteuert werden, da Entfernung und Verbindungsdauer keinerlei Externe Kosten gemäß [127] generieren (vergl. Kapitel 3.6.3.3).

Wesentlicher Erfolg der Internets ist das „Always-On“ Paradigma, welches im Internet implizit realisiert ist und als solches keine Ressourcen verbraucht (in IPv6 wird das Belegen einer IP-Adresse nicht mehr als Ressourcenengpass betrachtet). Ebenso ist „Entfernung“ ein Parameter, der im Internet eine untergeordnete Rolle spielt, bzw. der sich auch während einer Sitzung ändern kann.

Die Herausforderung liegt also hier in der Implementierung eines Preismodells, das von beiden Benutzergruppen akzeptiert wird und zum gesamtwirtschaftlichen Erfolg des Netzbetreibers beiträgt.

5.3.2 Netzbetreiberorientierte Sichtweise

Ein Netzbetreiber wird in der Marktwirtschaft am gesamtwirtschaftlichen Erfolg gemessen. Alle weiteren Ziele sind Unterziele, welche je nach aktueller Situation langfristig zu diesem Erfolg beitragen sollten. In der jüngeren Vergangenheit waren diese Kernziele (nicht zuletzt hervorgerufen durch den Druck der Kapitalmärkte) durch die Ziele „Marktanteil“ und „Umsatzsteigerung“ verdrängt worden. Aktuell ist jedoch wieder eine Rückbesinnung an die marktwirtschaftlichen Kernziele der langfristigen Gewinnentwicklung zu beobachten.

Kommunikationsnetze sind ein Gut mit einer sehr charakteristischen Kostenstruktur, die sich aus einem relativ hohen Festkostenanteil und entsprechend niedrigem variablem Kostenanteil zusammensetzt [117]. Dieses Verhältnis wurde durch die Lizenzvergabe der UMTS-Lizenzen (in Deutschland und UK) verstärkt. In einer solchen Umgebung werden langfristig ausschließlich die variablen Kosten (gemäß Kapitel 3.6), den Unternehmenserfolg nachhaltig beeinflussen.

Eine zentrale Anforderung an eine Mobile Internet-Infrastruktur ist das technische Umsetzen der in Kapitel 3 vorgestellten Konzepte der Preispolitik in einer Art, welche die verschiedenen Modelle und Konzepte benutzerorientiert unterstützt und die variablen Kosten minimiert. Zusätzlich ist eine gewisse Flexibilität gefragt, um die Mechanismen des Marketings (vor allem der Preispolitik) entsprechend flexibel und dynamisch in das Netz einzuführen bzw. auf entsprechende Marketingaktivitäten der Konkurrenz reagieren zu können.

5.3.3 Regulatorische Sichtweise

Der Blick der Regulierungsbehörden ist immer auf die Gesamtentwicklung eines Wirtschaftsraumes ausgelegt. Hier wurde in [132] gezeigt, dass es Situationen gibt, in denen ein Markt, überlässt man ihn ohne regulatorischen Eingriff sich selbst, nicht die volkswirtschaftlich optimalen Zustände erreicht. Ein regulatorischer Eingriff hilft hier mittelfristig nicht nur den Benutzern, sondern auch den Unternehmen. Diese Sichtweise ist im weiteren Verlauf dieser Arbeit nur insofern relevant, dass hier ggf. weitere dynamische Anforderungen an einen Netzbetreiber gestellt werden. Diese Sichtweise soll jedoch nicht weiter berücksichtigt werden.

5.4 Randbedingungen eines B3G-Preismodells

Ein Preismodell für ein B3G-Netz hat aus Sicht der Benutzerakzeptanz die Aufgabe, sowohl von traditionellen 2G/3G-Benutzern, als auch von den Internet-Nutzern akzeptiert zu werden. Parallel hierzu verfolgt ein Preismodell das Ziel, den Unternehmenserfolg eines Unternehmens zu maximieren. Betrachtet man die aktuellen Umsatzzahlen der Internet und Mobiltelekommunikationsbetreiber, so kann zunächst festgehalten werden, dass die Mobiltelekommunikationsbetreiber mehr Gewinn erwirtschaften als die Internetbetreiber. Es hat sich hier jüngst die Erkenntnis bei den Telekommunikationskonzernen durchgesetzt, dass die Trennung zwischen Internet und Festnetz nicht vorteilhaft ist und diese beiden operativen Einheiten kurz bis mittelfristig wieder verschmolzen werden. Mittelfristig macht in logischer Konsequenz davon die operative Trennung zwischen Festnetz und Mobiltelekommunikationsnetz ebenso wenig Sinn.

5.4.1 Roaming zwischen Betreibern

Die Frage ob zwischen verschiedenen Netzbetreibern ein Roaming-Abkommen getroffen wird, welches es einem Benutzer erlaubt, zwischen Netzbetreibern zu wechseln, erscheint aus Sicht der Benutzer attraktiv, ist jedoch aus Sicht der Netzbetreiber eher unerwünscht. Der Grund hierfür ist, dass ein Netzbetreiber versucht, den Kunden an sein Netz zu binden und in jedem Fall keine Kunden „freiwillig“ an einen konkurrierenden Netzbetreiber abgibt. Grund

für dieses Verhalten ist das Streben eines jeden Anbieters eines Produktes, sich von dem Konkurrenten positiv abzuheben. Hat man einen Kunden an ein Produkt gebunden, ist eine künstliche Austrittsbarriere zu schaffen. Diese Austrittsbarriere soll den Kunden langfristig an das Produkt binden. Somit wird für den weiteren Verlauf dieser Arbeit angenommen, dass für Endkunden, wie auch aktuell im 2G-Netz üblich, überall dort ein Netzwechsel zum Konkurrenten nicht aktiv unterstützt wird, wo der Primäranbieter einen Dienst anbieten kann. Das bedeutet, dass ein Kunde nur eingeschränkt aktiv in die Netzauswahl eingreifen wird und so lange wie möglich bei einem Netzbetreiber eingebucht ist. Roaming-Vereinbarungen werden nur dann relevant, wenn ein Netzbetreiber in einem bestimmten geographischen Bereich keine Infrastruktur zur Verfügung hat und quasi für dieses Gebiet eine Vereinbarung mit einem anderen Betreiber trifft.

5.4.2 Kundenbeziehungen

Aktuell kann man in fast allen Bereichen der Wirtschaft in spezifischen Marktsegmenten einen Trend weg vom spezialisierten Nischenanbieter und hin zum Komplettanbieter beobachten. Dies tritt vor allem bei Massenmärkten auf, in denen Käufer mit großer Marktmacht diese Forderungen erfolgreich am Markt durchsetzen konnten. Dies hat das Käuferverhalten generell beeinflusst und heute sehen wir uns mit einer Situation konfrontiert, in der Komplettanbieter anstreben, immer umfangreichere Dienstbündel anzubieten. Diese Entwicklung wird auch im Telekommunikationssektor dafür sorgen, dass z.B. ein Mobilfunkkunde nicht gleichzeitig mehrere Vertragsbeziehungen mit unterschiedlichen Betreibern akzeptiert und sich am Ende für denjenigen Betreiber entscheidet, der die gewünschten Dienste (auch Dienste von Drittanbietern) aus einer Hand kostengünstig anbietet. Somit verlieren die Drittanbieter den direkten Kontakt zum Kunden, der dann dem so genannten Primäranbieter überlassen bleibt.

Die Abrechnung zwischen Primäranbieter und Drittanbieter kann über verschiedene Modelle erfolgen, die im weiteren Verlauf dieser Arbeit nicht weiter erörtert werden.

Diese angenommene Tendenz bringt die Primäranbieter mittelfristig in eine sehr komfortable Position, da sie über den direkten Kundenkontakt neue Geschäftsmodelle entwickeln können, bei denen ggf. das historische Kerngeschäft der Übertragung von Sprache eine weniger zentrale Rolle spielt. Dies kann z.B. das Anbieten von kostenlosen Telekommunikationsdiensten sein, wenn geeignete Werbemaßnahmen den Bereitstellungsdienst unterbrechen. Durch den direkten Kontakt zum Kunden in Kombination mit Maßnahmen zur Pflege der Kundenbeziehungen (*CRM*) kann das Benutzerverhalten individuell erforscht werden. Diese Informationen werden dann geeignet für eine Profilbildung der Benutzer verwendet. Konkret können damit z.B. bedarfsgerechte Werbeinformationen dediziert an Kunden adressiert werden, die dann zu einer Erhöhung der Rücklaufquote führen, für deren Erlangen Drittanbieter gerne bereit sind, Geld auszugeben.

Es soll an dieser Stelle ausdrücklich festgehalten werden, dass Telekommunikations/Internetanbieter nicht ausschließlich Primäranbieter sein werden, sondern durchaus auch in die Rolle eines Sekundäranbieters gedrängt werden können. Wegen Ihres großen Kundenstamms haben die Telekommunikationsanbieter im Rennen um die Rolle des Primäranbieters eine vorteilhafte Ausgangsposition; jedoch haben z.B. Banken/Kreditkartenunternehmen eine ebenfalls komfortable Perspektive, um diese Rolle einnehmen zu können. Ähnliche Modelle sind prinzipiell auch von aktuellen Drittanbietern erstrebenswert und man kann in jedem Fall von einem Kampf zwischen diesen verschiedenen Parteien um den Kunden ausgehen. So könnte zum Beispiel ein Spieleanbieter einen Exklusivvertrag mit einem Netzanbieter abschließen und der Endkunde würde eine Rechnung über die Nutzung von Spielen einschließlich Netzzugang als Dienstbündel beziehen.

Im weiteren Verlauf dieser Arbeit wird also davon ausgegangen, dass einzig der Primäranbieter den direkten Kontakt zum Kunden hat, ein Kunde jedoch prinzipiell Geschäftsbeziehungen zu mehreren Primäranbietern gleichzeitig unterhalten kann. In jedem Fall wird in dieser Arbeit davon ausgegangen, dass die Anzahl der parallel unterhaltenden Geschäftsbeziehungen aus Sicht des Endkunden begrenzt bleibt. Diese Annahme basiert auf dem in Kapitel 4.4 vorgestellten Ansatz, welcher jedoch die Rollen der verschiedenen Akteure konzeptionell weiterentwickelt.

Die oben dargestellten Beziehungsmodelle erfordern ein System zur Dienstbündelung bzw. Produktbündelung wie in Kapitel 3.4 beschrieben. Solche Dienstbündel werden dem Kunden dann als Gesamtpaket angeboten. Die Komposition dieser Bündel ist vielfältig und komplex. Sie erstreckt sich im Allgemeinen über eine Vielzahl von verschiedenen Diensten mit verschiedenen Schwerpunkten, bei denen in der Tendenz der ursprüngliche Telekommunikationsdienst eine Nebenrolle spielt. Weiter wird davon ausgegangen, dass sich diese Bündel benutzerspezifisch ändern und z.B. in Form von Sonderaktionen zeitlich befristet angeboten werden, jedoch ggf. unbegrenzt gültig sind.

Eine Infrastruktur zur Nutzungsdatenerfassung inklusive dem nachgelagerten Aufbereiten der Nutzungsdaten muss zunächst grundsätzlich in der Lage sein, Nutzdaten, d.h. IP-Pakete erfassen zu können, um daraus Nutzungsdaten zu generieren. Es kann dabei jedoch nicht von einheitlichen Parametern für die Nutzungsdatenerfassung über alle Benutzer hinweg ausgegangen werden. Eine benutzerspezifische und benutzerindividuelle Parametrisierung der gesamten Abrechnungsinfrastruktur ist in diesem Zusammenhang deshalb notwendig, weil dies einen signifikanten Kostenblock eines Betreibers darstellt, wie in Abbildung 27 dargestellt. Es wird daher angenommen, dass sich zwischen Flexibilität der IT-Infrastruktur und Aufwand für Nutzungsdatenerfassung, nachgelagerter Aufbereitung und Rechnungserstellung ein Gleichgewicht einpendelt. Dies wird dadurch erreicht, dass den Benutzern vordefinierte Dienstbündel angeboten werden, die der Primäranbieter komponiert und dem Benutzer innerhalb dieser Bündel keine weitere Wahlmöglichkeit lässt. Das geeignete Komponieren dieser Bündel ist dabei Aufgabe der Marktforschung und soll hier nicht weiter verfolgt werden.

Die Entscheidungsbasis, ein Bündel einzuführen, ist allein über die zu erwartende Nachfrage gegeben. Somit ist zu erwarten, dass bestimmte Dienstbündel vom Markt verschwinden werden, und (betrachtet man ein solches Dienstbündel als Produkt) jedes dieser Produkte eine begrenzte Lebenszeit hat. Nimmt ein Betreiber ein solches Produkt aus seinem Angebotsportfolio, bedeutet dies, dass es keine neuen Benutzer gibt, welche dieses Produkt beziehen können. Dies bedeutet jedoch nicht zwingend, dass dieses Produkt von den bisherigen Benutzern nicht mehr weiter konsumiert wird. Es ist davon auszugehen, dass Benutzer sich nur dann für ein neues Produkt entscheiden, wenn Ihnen ein entsprechendes Angebot für ein neues Produktbündel gemacht wird.

5.4.3 Prepaid

Wie in Kapitel 3.6 aufgezeigt, hat ein Netzbetreiber verschiedene Kostenarten zu schultern. Ein Teil der Kapitalkosten nach Kapitel 3.6.3.3 entsteht dadurch, dass der Netzbetreiber in Vorleistung gehen muss, d.h. den Dienst bereitstellt und in der Regel erst am Monatsende eine Rechnung ausstellen kann. Dieser Liquiditätsverlust ist erheblich und verstärkt sich durch die Tatsache, dass sehr viele Rechnungen unbeglichen bleiben. Hier hat die Prepaid-Bezahlform einen signifikanten Vorteil [77]. Die aktuellen Statistiken über unbeglichene Telefonrechnungen werden von den Netzbetreibern nicht gerne kommuniziert und es sind somit keine verlässlichen Daten verfügbar. Jedoch kann von einem nennenswerten „fiktiven Kostenblock“ ausgegangen werden.

Dieser fiktive Kostenblock erweitert sich um einen realen Kostenblock, da zusätzliche Aufwendungen im Bereich Zahlungseingangsprüfung, Mahnung und ggf. das Einleiten von Rechtsmitteln entstehen. In diesem Zusammenhang ist das Anbieten von Prepaid-Verträgen in mehrerer Hinsicht vorteilhaft. Man könnte sich nun fragen, warum die Betreiber nicht komplett zum Prepaid-Modell übergehen. Es hat sich jedoch gezeigt, dass die Kunden mit diesem Modell weniger gut an das Unternehmen gebunden werden, und es unter Prepaid-Kunden eine höhere Wechselbereitschaft zu anderen Anbietern gibt. Es wird zwar mit Mechanismen wie Preisnachlässen bei Wiederaufladen einer Prepaid-Karte versucht, diese Wechselbereitschaft zu reduzieren, jedoch hat sich hier noch nicht der gewünschte Erfolg für den Netzbetreiber eingestellt. Zusätzlich bietet ein Vertragskunde für den Betreiber eine langfristige Perspektive, die zu erwartenden Umsätze abzuschätzen. Aus diesen Zusammenhängen lässt sich schließen, dass Prepaid-Modelle vor allem in jenen Marktsegmenten eine attraktive Alternative darstellen, in denen mit einer erhöhten Wahrscheinlichkeit Rechnungen nicht beglichen werden.

Es wird deshalb im weiteren Verlauf dieser Arbeit davon ausgegangen, dass es auch weiterhin einen signifikanten Anteil an Prepaid-Kunden geben wird, jedoch die Unternehmen Festverträge bevorzugen.

5.4.4 Preispolitik

Dieser Abschnitt beschreibt den Ausgangspunkt dieser Arbeit bezüglich der Preispolitik. Die bisher getroffene generelle Annahme für eine „Mobile Internet-Infrastruktur“ hat Auswirkungen auf die technischen Möglichkeiten zur Umsetzung einer Preispolitik. So ist z.B. ein reines „verbindungs-dauerabhängiges“ Preismodell in einer IP-Umgebung nicht einfach zu realisieren bzw. durchzusetzen. Andere Elemente der „verbindungsorientierten Preispolitik“ behalten jedoch ihre weitere Existenzberechtigung in einem konvergierenden Netz und werden im Folgenden dargestellt. Diese Elemente sind im Wesentlichen die Dienstbündelung (vergl. Kapitel 3.4), die Preisdifferenzierung (vergl. Kapitel 3.2.3.2) und, wie bereits erwähnt, die Prepaid-Bezahlform.

Dienste werden aus Sicht des Netzbetreibers in gebündelter Form und in strikt limitierter Konfiguration und Granularität angeboten. Ein Endbenutzer hat nur dann die Möglichkeit, nennenswert auf diese Elemente der Preispolitik Einfluss zu nehmen, wenn sie einen signifikanten Umsatz versprechen. Dies ist in der Regel nur bei Großkunden zu erwarten.

Üblicherweise kann man in der 2G-Welt bei jedem Anbieter zwischen ca. 10 verschiedenen Tarifen wählen, die auf verschiedene Benutzergruppen zugeschnitten sind. Diese Tarife entsprechen verschiedenen Produktbündeln. Hier werden die Benutzer entsprechend verschiedenen Klassen typisiert und bzgl. der Zahlungsbereitschaft nach Kapitel 3.2.3 klassifiziert. In der Regel hat dabei der Benutzer nicht die Möglichkeit, ein für sich individuelles Servicepaket zusammenzustellen. Benutzer haben dies akzeptiert, und es wird im weiteren Verlauf dieser Arbeit davon ausgegangen, dass dieses Konzept in ein neues, konvergierendes Netz übernommen wird.

Innerhalb der angebotenen Dienstklassen kommen Elemente wie Preisdifferenzierung (vergl. Kapitel 3.3.3) zum Einsatz. Es sei in diesem Zusammenhang erwähnt, dass die Zuordnung von Benutzern zu diesen Produktbündeln/Tarifen als relativ statisch betrachtet werden kann. D.h. in der Regel bindet sich ein Benutzer für zwei Jahre mittels eines Vertrages an eine Dienstklasse. In diesem Zeitraum ist ein Wechsel eher unüblich und in der Regel mit finanziellen Nachteilen für den Benutzer verbunden. Ein dynamisches Wechseln zwischen diesen Produktbündeln wird für den weiteren Verlauf dieser Arbeit als nicht marktrelevant betrachtet

und explizit ausgeschlossen. Die Gründe hierfür sind das Bestreben eines Netzbetreibers, Marktsegmente in einem Telekommunikationsnetz etablieren zu können. Ein solches Marktsegment zeichnet sich dadurch aus, dass Image, Flexibilität und andere Zusatznutzen (und natürlich das Preisniveau) in einem Produktbündel so am Markt positioniert werden, dass bestimmte Benutzer (z.B. als Statussymbol, aus Gründen der Flexibilität, wegen der höheren Zahlungsbereitschaft usw.) ein höherwertiges Bündel beziehen, ohne dies konkret nutzen zu können. Als vergleichbares Beispiel sei hier die Preispolitik der traditionellen Fluglinien erwähnt. Wenn ein Reisender nicht weiß, wann er einen Rückflug antreten kann, bucht er ein Business-Ticket. Die Fluggesellschaft erlaubt ihm in der Regel nicht, die Reise mit einem Economy-Ticket zu beginnen und im Bedarfsfall den Rückflug in ein Business-Ticket umzuwandeln. Würde man diese Möglichkeit einführen, gäbe es keinen Grund mehr von Beginn an Business-Ticket zu buchen. Man würde das Markesegment „Business Class“ gefährden, was wohl zu einer Verringerung von Umsatz und Gewinn führen würde.

Somit soll neben der Annahme, dass Prepaid-Mechanismen, die im 2G-Netz sehr erfolgreich eingeführt wurden, auch weiterhin ein wesentliches Element der zukünftigen Preispolitik sein werden, das Etablieren von Marktsegmenten ein weiterer wesentlicher Aspekt in einem kommerzialisierten Mobil Internet sein. Dieses Etablieren von Marktsegmenten und die Unterstützung von Prepaid-Mechanismen führen zu einem streng benutzerspezifischen Ansatz, in dem eine dienstorientierte Sichtweise hinderlich für die Kommerzialisierung des Mobil Internet ist.

5.4.5 Benutzerverwaltung

Neben einem Preismodell, das sich so weit wie möglich an dem traditionellen, aus der Telefonie abgeleiteten Preismodell orientiert, werden in diesem Abschnitt noch einige Elemente der Benutzerverwaltung dargestellt, die für den weiteren Verlauf dieser Arbeit relevant sind.

Zunächst soll in dieser Arbeit davon ausgegangen werden (vergl. Kapitel 5.4.1), dass Netzbetreiber über Roaming-Abkommen verbunden sind und ihren Benutzern dadurch eine erweiterte Erreichbarkeit anbieten. Der Benutzer ist in einem bestimmten Netz „zu Hause“ und tritt mit Netzbetreibern, in deren Netz er sich dann sozusagen als Gast aufhält, nicht kommerziell in Kontakt. Im Falle eines Roamings kann ein Benutzer jedoch die im Netz eines Roaming-Partners konsumierten Ressourcen über die eigene Rechnung abrechnen. Dieses Konzept ist im Internet zunächst nicht verbreitet, soll aber für den weiteren Verlauf dieser Arbeit angenommen werden. Natürlich ist es einem Benutzer möglich, gleichzeitig mehrere Verträge mit verschiedenen Betreibern zu unterhalten, jedoch wird dieser Fall im Rahmen dieser Arbeit nicht weiter verfolgt. Vorwegnehmend kann jedoch hier schon festgehalten werden, dass dieser Fall sehr wohl von den getroffenen architektonischen Vorschlägen mit abgedeckt wird.

Der Grund für diese zunächst getroffene Annahme ist die Erkenntnis aus der Marktforschung, dass ein Benutzer in der Regel nicht eine sehr große Anzahl von gleichzeitigen Geschäftsbeziehungen mit verschiedenen Anbietern akzeptiert. Dies hätte mehrere verschiedene Rechnungen zur Folge, welche ein Benutzer für das Konsumieren eines Dienstes erhalten würde. Erfahrungen in im 2G-Netz haben gezeigt, dass Kunden in einem solchen Fall ein Dienstbündel nach Kapitel 3 mit klarer und überschaubarer Rechnungsaufstellung bevorzugt konsumieren.

Weiter haben sich die Benutzer aus der 2G-Welt daran gewöhnt, Einzelgesprächsnachweise in der Abrechnung zu erhalten. Dies ist im Prinzip eine explizite Aufstellung der Nutzungsdaten, bei der zwar einige datenschutzrechtliche Rahmenbedingungen eingehalten werden müssen,

jedoch es einem Benutzer ermöglicht wird, die Plausibilität der Rechnung nachvollziehen zu können. Das Generieren dieser Daten erlaubt nicht bzw. nur eingeschränkt die Anwendung statistischer Methoden und basiert auf realen Messungen. Im Rahmen dieser Arbeit wird die Möglichkeit, statistische Methoden zu diesem Zweck einzuführen, nicht berücksichtigt. Es wird davon ausgegangen, dass die Abrechnungen auf real gemessenen Daten beruhen.

5.4.6 Zukünftige Dienststruktur

Das Telefonienetz hat historisch einen Einheitsdienst – die Vermittlung und Übertragung von Sprache – im Angebot. Dieses Angebot wurde dann im Laufe der Zeit in technische Parameter wie z.B. Bandbreite transferiert, um es einem Benutzer zu ermöglichen, anstatt Sprache Daten (z.B. Fax usw.) zu übertragen. Aus Sicht des Netzes jedoch unterscheidet sich die Übertragung eines gesprochenen Wortes nicht von der Übertragung eines geschriebenen Wortes. Mit der mobilen Telefonie wurde dann neben diesen Dienst der Short Message Service (SMS) erfolgreich am Markt platziert. Hier werden in einer SMS bis zu 180 Zeichen übertragen. In jedem Fall kann hierbei festgehalten werden, dass die absolute Anzahl der Dienste überschaubar ist, und sich somit grundlegend vom Internet unterscheidet.

Das Internet ist historisch ein reines Datennetz, das eine nahezu unbegrenzte Dienstvielfalt anbietet. Die Anforderungen solcher Dienste an das Netz sind sehr unterschiedlich und weder vom Benutzer, noch von Netzbetreiber bzgl. ihren spezifischen Anforderungen exakt beschreibbar.

Dadurch, dass in der Regel keine Kommunikation zwischen Netzschicht und Anwendung stattfindet, und die Signalisierung bei IP implizit in den Datenpaketen (Header) abgewickelt wird, ist das Internet sehr offen und nimmt sehr einfach beliebig viele Anwendungen auf, die dann historisch nach dem so genannten „Best Effort“ Prinzip bedient werden. Diese Tatsache verhindert, dass sich aus Sicht des Netzbetreibers verschiedene Marktsegmente ausgebildet haben, in der Kunden mit größerer Zahlungsbereitschaft eben die Anwendungen mit einer ihrer Zahlungsbereitschaft angepasste Dienstgüte konsumieren können.

Im weiteren Verlauf dieser Arbeit wird angenommen, dass im zukünftigen Mobilien Internet eine sehr große Vielfalt an Diensten angeboten wird, die dann vom Netzbetreiber gemäß Kapitel 3.4 geeignet gebündelt werden. Diese Bündel bestehen aus einer Kombination von Anwendungen, welche mit einer bestimmten Dienstgüte und ggf. anderen netzspezifischen Parametern angeboten werden. Ein solches Bündeln der Dienste ist aus Sicht des Netzbetreibers sehr wichtig, um ein überschaubares und verwaltbares Netzmanagement inklusive Abrechnungsinfrastruktur geeignet implementieren zu können. Diese Annahmen werden auch aus Sicht der Benutzerakzeptanz untermauert, da nach aktueller Kenntnis Benutzer sehr großen Wert auf die Übersichtlichkeit und Verständlichkeit einer Rechnung legen.

5.5 Preismodell für ein B3G-Netz

In diesem Abschnitt soll, basierend auf den Gedanken aus Kapitel 3, ein Preismodell entwickelt werden, welches ein zukünftiger Betreiber eines Mobilien Internets einsetzen könnte, um sich am Markt zu platzieren. Generell kann festgehalten werden, dass ein allgemeiner Trend hin zur Definition von höherwertigen Diensten und Inhalten erfolgt. Dies bedeutet, dass nicht ausgeschlossen werden kann, dass der Transport von Daten mittel bis langfristig nicht mehr als Kernkomponenten eines umfassenden Preismodells zu betrachten ist, sondern geeignet mit weiteren Parameters aus dem Anwendungs- bzw. Inhaltebereich verzahnt wird. Diese Arbeit wird sich jedoch auf die Entgelterfassung beim Datentransport beschränken.

Nachdem ausgewählte, in der öffentlichen Diskussion stehenden Vorschläge für ein Preismodell im Telekommunikationsumfeld in Kapitel 4.9 diskutiert wurden, soll nun in diesem Kapitel ein Vorschlag für ein alternatives Preismodell vorgestellt werden, das den Namen *Differenzierte Flat-Rate* erhält und erstmals so in [69] vorgestellt wurde. Dieses Preismodell soll ausgehend vom PMP-Modell nach Kapitel 4.9.4 erläutert werden. Eine Einteilung des Marktes, wie es beim PMP erfolgt, entspricht der Preisdifferenzierung nach [98], [99] und [100] (vergl. Kapitel 3.2.3.2).

Dieses Schema hat den Nachteil, dass es nicht in der Lage ist, zwischen den bestimmten Arten des Internetverkehrs zu differenzieren. So verzeihen z.B. TCP-Anwendungen kurzfristige Netzstauungen sehr großzügig und werden in der Regel vom Anwender nicht wahrgenommen. Bei UDP-Anwendungen ist die Verzögerung ein wesentlicher Parameter, der die vom Benutzer wahrgenommene Dienstgüte wesentlich beeinflusst. Paketverluste bei Multimediaanwendungen haben hier sehr unterschiedliche Auswirkungen. Bei einem Multimedia-Datenstrom werden z. B. in einigen Paketen Rahmeninformationen transportiert. Ein Verlust eines solchen Paketes würde einen Totalausfall von Audio bzw. Video zur Folge haben, da dies den Verlust der Synchronisation eines Medienstroms zur Folge hätte. In anderen Paketen von Multi-Media-Datenströmen würde ein Paketverlust lediglich z.B. eine kurzzeitige Farbveränderung eines Bereiches zur Folge haben und somit die Kommunikation weniger stark einschränken. Bei Verlust eines Paketes, das Audioinformationen transportiert, wäre ein lediglich leichtes „Knacken in der Leitung“ hörbar.

Dieses Beispiel verdeutlicht, dass es einem Kunden nicht zumutbar ist, technische Parameter selbst zu definieren und sich aktiv mit einer Klassifikation seiner Anwendungen nach Kapitel 4.8.1.1 auseinanderzusetzen. Da aber genau diese Parameter dazu verwendet werden sollen, die Ressourcen bzw. die Nachfrage der Ressourcen zu verwalten, ist es sinnvoll, eine Adaptionsschicht einzuführen. Diese Schicht hat die Aufgabe, dem Benutzer die wesentlichen technischen Parameter vorzuenthalten. An ihrer unteren Schnittstelle zum Netz wird, für den Benutzer nicht sichtbar, die entsprechende Information dem Netz übermittelt. Derartige Adaptionsschichten werden als Element der in Kapitel 3.4 vorgestellten Produktbündelung in vielen anderen Sektoren eingesetzt und wurden auch prinzipiell in Kapitel 4.5 vorgestellt. Analog hierzu sollen in dieser Arbeit exemplarisch drei Produktbündel spezifiziert werden. Dies sind:

- Professionelles Paket
- Komfort-Paket
- Standard-Paket

Prinzipiell ähnelt ein solches Etablieren von Produktbündeln dem PMP (Kapitel 4.9.4). Auch hier wird der Markt statisch in verschiedene Teilmärkte eingeteilt. Eine solche Einteilung ist im Mobil Internet jedoch wegen der bereits erwähnten Gründe nicht optimal. Aus diesem Grund wird in Anlehnung an den Smart Market (Kapitel 4.9.5) eine Flexibilisierungskomponente hinzugefügt. Beim Smart Market würden je nach Nachfrageverhältnissen bis zu 100 % der verfügbaren Ressourcen den Benutzern der höchsten Priorisierungsstufe zugeteilt werden, da diese bereit waren, den höchsten Preis zu bezahlen. D.h., die statische Aufteilung der Marktsegmente wird dynamisiert. Der *Differenzierte Flat-Rate* Vorschlag kann somit als eine Kombination aus Smart Market und Paris Metro Pricing betrachtet werden. Das Einteilen der Kunden in drei Gruppen hat zur Folge, dass der Markt in drei Teilbereiche eingeteilt wird, denen jedoch keine statischen Mengen (Ressourcen) zugewiesen werden. Die Gesamtmenge an verfügbaren Ressourcen bleibt konstant und die Mengenallokation der Teilbereiche wird hier dynamisch vorgenommen.

Die technischen Parameter, die für das vorgeschlagene Preisschema verwendet werden, sollten lediglich maximale Bandbreite und Datenmenge sein. Da jegliche Kommunikation mit dem Benutzer zu erhöhten Kosten in der Datenhaltung und bei der Messinfrastruktur beiträgt, sollte diese minimiert werden. Somit ist eine Signalisierung der Zahlungsbereitschaft eines Benutzers, wie beim Smart Market vorgeschlagen, nur auf einem externen Weg möglich. Dies hat zur Folge, dass diese Signalisierung für einen größeren Zeitraum ihre Gültigkeit besitzen muss, da eine externe Kommunikation nicht in Echtzeit möglich ist.

Die Vorteile aus einem solchen System sind die Steigerung der verkaufbaren Ressourcen und ein Wegfallen des Aufwands für das Verwalten der Zahlungsbereitschaft und der damit verbundenen Auktionatoreinrichtungen. Innerhalb einer jeden Klasse gibt es beim vorgeschlagenen Verfahren einen mengen/volumenbegrenzenden Mechanismus, der dem aus Kapitel 3.5.3.2 ähnelt. Der zusätzliche Aufwand bzgl. der Funktionalität im Netz besteht darin, die verbleibenden Budgets der Ressourcen aus dem Netz an die aktuellen Netzzugangsknoten zu signalisieren.

Tabelle 6 zeigt exemplarisch ein detailliertes *Differenziertes Flat-Rate* Preisschema. Aus der Tabelle wird ersichtlich, dass jedem Dienst, der vom Netzanbieter unterstützt wird, pro Klasse ein volumenbasierter und ein bandbreitenbasierter technischer Parameter zugewiesen ist. In Ausnahmefällen kann natürlich nur einer der beiden Parameter spezifiziert werden.

Tabelle 6: Exemplarisches Preisschema

	PREMIUM USER	ADVANCED USER	STANDARD USER
Email	100 kbit/s 2 GB/Monat	80 kbit/s 1GB/Monat	60 kbit/s 500 MB/Monat
Telnet	40 kbit/s	40 kbit/s	40 kbit/s
WWW	200 kbit/s 10 GB/Monat	150 kbit/s 7 GB/Monat	100 kbit/s 4 GB/Monat
FTP	100 kbit/s 2 GB/Monat	80 kbit/s 1GB/Monat	60 kbit/s 500 MB/Monat
DNS	40 kbit/s	40 kbit/s	40 kbit/s
UDP	400 kbit/s 2 GB/Monat	200 kbit/s 1 GB/Monat	150 kbit/s 500 MB/Monat
Spiele Server	200 kbit/s 10 GB/Monat.	200 kbit/s 4 GB/Monat	0

Die Parameter, die bei diesem exemplarischen Preisschema gewählt sind, sollen keine quantitativen Vorgaben darstellen, sondern nur als qualitative Anregungen verstanden werden. Die genaue Parametrisierung ist sicherlich nicht statisch und sollte mit der Zeit an das Benutzerverhalten angepasst werden. Aus Sicht der Betreiber ist dies jederzeit und ohne Vertragsänderungen mit den Kunden möglich, da die adaptive Zwischenschicht den Kunden diese Parameter vorenthält.

Der technische Parameter Bandbreite spiegelt nur eine maximal zulässige Bandbreite wieder, die vom System jedoch nicht garantiert wird. Die mengenmäßige Beschränkung der einzelnen Dienste pro Klasse soll dem Benutzer die Knappheit der verfügbaren Ressourcen nach Kapitel 3.2.4 kommunizieren. Diese technischen Parameter bleiben dem Kunden jedoch ebenfalls verborgen. Den Kunden werden ausschließlich die Klassen der Adaptionsschicht kommuniziert. Für jedes dieser logischen Pakete wird dann eine Flat-Rate angewendet. Durch die Mengenbegrenzung wird der Nachteil der Flat-Rate des unkontrollierten Ressourcenverbrauches

eingeschränkt. Durch Einsatz von QoS-Mechanismen nach in Kapitel 2.1.2.3, kann den Datenpaketen der einzelnen Benutzer eine jeweilige Priorität zugewiesen werden. Dieses Zuweisen würde automatisch im Endgerät geschehen und wäre fest vorinstalliert. In der Realisierung wäre hier denkbar, dass ein Kunde bei Vertragsabschluß eine bestimmte Software zur Verfügung gestellt bekommt, die den ausgehenden Verkehr entsprechend der gewählten Klasse markiert. Diese Software wäre frei verfügbar. Bevor der im Netz registrierte Benutzer eine Anwendungssitzung beginnen kann, müsste er sich Authentisieren und Autorisieren (siehe Kapitel 4.4). Dies würde z.B. beim Einschalten des Endgeräts und gleichzeitiger Identifikation des Benutzers geschehen. Nach erfolgreicher Autorisierung/Authentisierung werden die entsprechenden Parameter zum Netzknoten übertragen, welcher die Ressource freigibt. Dieser Prozess ist prinzipiell identisch zu dem in 3GPP spezifizierten Verfahren aus Kapitel 2.2.2. Im Falle eines Festvertrages würden die Regeln nach Tabelle 6 übertragen, welche beim Prepaid-Benutzer noch um das verbleibende Guthaben erweitert werden würde. Der Zugangsrouten hätte Kenntnis von der vertraglichen Leistung, die dem jeweiligen Benutzer zusteht und kontrolliert lediglich, ob das von diesem Benutzer ausgesendete Verkehrsprofil im Rahmen des Vertrages ist. Sollte ein Kunde die gegen seinen Vertrag verstoßen, so werden die über den Vertrag hinaus gehenden Datenpakete verworfen.

Durch die Priorisierung des Verkehrs und durch die Zuordnung der Benutzer zu den verschiedenen Klassen würde das System die Ressourcen automatisch im Falle einer Netzstauung entsprechend der Zahlungsbereitschaft der Kunden verteilen. Tabelle 7 zeigt einen exemplarischen Fall einer Netzstauung. Im Beispiel befinden sich 12 Benutzer in einer Zelle, von denen vier Benutzer den Status *Premium User*, vier weitere den Status *Advanced User* und weitere vier Benutzer den Status *Standard User* haben. Im Beispiel wird davon ausgegangen, dass jeder der Benutzer über ausreichend Budget der entsprechenden Dienste verfügt. Weiter steht gleichzeitig der Bedarf aller 12 Teilnehmer an, eine Sitzung in der maximal zulässigen Kapazität zu starten. Daraus ergäbe sich folgende Nachfrage (Tabelle 7):

Tabelle 7: Exemplarische Gesamtnachfrage der Ressourcen

Premium User	4* 400kbit/s	1600 kbit/s
Advanced User	4*200 kbit/s	800 kbit/s
Standard User	4*150 kbit/s	600 kbit/s
Gesamtnachfrage		3000 kbit/s

Im Beispiel übersteigt die Nachfrage das maximale Angebot eines angenommenen 3G-Zugangsnetzes von 2 Mbit/s. Das System ist völlig überbucht. In einem solchen Fall würden die Ressourcen bei Implementierung des Differenzierten Flat-Rate Schemas nach Tabelle 8 verteilt.

Tabelle 8: Exemplarische Verteilung der Ressourcen

PREMIUM USER	4* 400 KBIT/S	1600 KBIT/S
Advanced User	4*100 KBit/s	400 KBit/s
Standard User	4*0 KBit/s	0 KBit/s
Gesamtnachfrage		2000 KBit/s

Premium User würden zu 100 % bedient. Dies führt zu einer Netzbelegung von 1,6 Mbit/s. Um die restlichen 400 kbit/s konkurrieren vier *Advanced User* mit einer Nachfrage von 800 kbit/s, und vier *Standard User* mit einer Nachfrage von 600 kbit/s. Die Nachfrage der *Advanced User* kann nur zu 50 % befriedigt werden. Dies belegt weitere 400 kbit/s und somit

bleiben für die *Standard User* keine Ressourcen übrig. Diese Nachfrage würde nicht befriedigt werden.

Natürlich ist das Sendeverhalten einer IP-basierten Datenquelle nicht so statisch wie im Beispiel angenommen und die Verteilung der Senderate über die Zeit kann als nicht konstant angenommen werden, was jedoch nur dazu führen würde, dass von den *Premium Usern* nicht verwendete Bandbreite sofort der nächst niedrigen Klasse ohne Einschränkung zur Verfügung gestellt wird.

Eine Realisierung auf Basis einer fiktiven Smart Market-Implementierung hätte in diesem Fall das identische Resultat ergeben. Bei Betrachtung dieses Falls aus Sicht des Smart Marktes haben im Prinzip zwölf Benutzer Gebote abgegeben, von denen zufälligerweise je vier Gebote identisch waren. Die Gebote der *Premium User* waren hoch genug und wurden angenommen. Die Gebote der *Advanced User* waren nicht hoch genug um vollständig angenommen zu werden, können jedoch teilweise bedient werden, und die Gebote der *Standard User* waren zu niedrig und die Pakete werden verzögert. Generell darf davon ausgegangen werden, dass *Premium User* ihre Pakete mit einer höheren Zahlungsbereitschaft versehen als *Advanced User* und diese wiederum eine höhere Zahlungsbereitschaft wie die *Standard User* haben. Somit kann durch die Eingruppierung der Benutzer in drei Klassen dieses System als Smart Market mit verschiedenen, festen Gebotsstufen betrachtet werden, bei denen im Falle einer Situation, bei der mehrere Benutzer gleicher Prioritätsklasse um nicht ausreichend zur Verfügung stehende Ressourcen konkurrieren, die freien Ressourcen entsprechend gerecht aufgeteilt werden. Lediglich die verdeckten Gebote werden vom Benutzer bei Vertragswahl abgegeben.

Prinzipiell kann man sich vorstellen, dass im Falle einer sehr starken Überlastung des Systems, also für den Fall, dass die Nachfrage in einer Klasse größer ist als die Summe aller im Angebot befindlichen Ressourcen, nach zwei Prinzipien vorgegangen wird. Eine Ablehnungsstrategie würde diejenigen Benutzer, die ihren Bedarf zuerst angemeldet haben, zu 100 % bedienen und die restlichen Benutzer müssten abgelehnt werden. Eine weitere Möglichkeit ist die Verteilungsstrategie, bei der eine faire Verteilung der Ressourcen unter den Parteien mit identischem Gebot angestrebt wird. Diese Möglichkeit wird beim *Differenzierten Flat-Rate Verfahren* vorgeschlagen, da hierzu weniger Netzausrüstung erforderlich ist. Es muss nicht festgehalten werden, welcher Benutzer seine Nachfrage zu spät angemeldet hat, und es müssen keine Ressourcen aufgewendet werden, um die Ablehnung zu kommunizieren. Parteien mit identischem Gebot sind also immer die Benutzer der gleichen Klasse. Es wird im Rahmen dieser Arbeit explizit nicht vorgeschlagen einen solchen Fall mit technischen Methoden zu lösen, da davon auszugehen ist, dass ein Ablehnen von *Premium Usern* mangels Ressourcenverfügbarkeit entweder umgehend Konkurrenten anzieht, oder aber den Netzbetreiber umgehend veranlasst, sein Netz an dieser Stelle weiter auszubauen.

Sollte ein Benutzer beim Differenzieren Flat-Rate Verfahren sein mengenmäßig beschränktes Budget nicht aufbrauchen, so verfällt es zum Monatsende.

Dadurch, dass Benutzer einer niederen Klasse während Zeiten mit höheren Stauungswahrscheinlichkeit einen garantiert schlechteren Dienst als die *Premium User* erhalten, werden sehr schnell Erfahrungswerte gesammelt, die das Benutzerverhalten entsprechend beeinflussen. Folglich werden diese Benutzer vorwiegend zu jenen Zeiten auf das Netz zugreifen, an denen die Wahrscheinlichkeit einer Netzstauung geringer ist. Als Alternative hierzu könnten sie ihre Zahlungsbereitschaft durch Wahl in eine höhere Klasse entsprechend erhöhen. Ein

Benutzer hat im Falle einer Netzstauung also folgende zwei Möglichkeiten, diese mittelfristig aufzulösen:

- Wahl einer höheren Klasse,
- Verschiebung seiner Sendewünsche auf einen späteren Zeitpunkt.

Sollten zu einem Zeitpunkt nur *Standard User* das Netz belegen, ist darauf zu achten, dass den Benutzern nicht mehr Ressourcen als vertraglich vereinbart zur Verfügung gestellt werden. Dies ist nicht im Einklang mit der im aktuellen Internet angewendeten Strategie, aus Sicht des Marketings jedoch sehr wichtig, da nur so mittelfristig ein Marktsegment etabliert werden kann.

Das *Differenzierte Flat-Rate Verfahren* könnte man auch als statische Quantisierung eines Smart Market bezeichnen. Die Anzahl der Quantisierungsstufen ruft die Frage hervor, wie viel Stufen aus Sicht des Datenmanagements optimal sind. Hier sind auf der einen Seite der zusätzliche Aufwand in der Benutzerverwaltung zu erwähnen, und auf der anderen Seite die technischen Fähigkeiten des Systems, zwischen mehr als drei Klassen differenzieren zu können. Ein zusätzlicher Aspekt wäre hier die Kundenakzeptanz. Es darf vermutet werden, dass eine zu große Anzahl von Klassen bei den Benutzern für Verwirrung sorgt und dies zur Folge haben kann, dass sich die Benutzer für keines der Angebote entscheiden.

5.6 Problemdiskussion

In der Mobilkommunikation existiert ein etabliertes Abrechnungskonzept, das den Anforderungen dieses Netzes voll genügt. Es wird jedoch hier die Frage aufgeworfen, ob dieses Konzept den neuen Anforderungen des Mobilnetzes im Sinne dieser Arbeit wirklich genügt. Die Gründe hierfür sind zum einen die Tatsache, dass die Parameter, welche zur Abrechnung herangezogen werden, in der IP-Welt in dieser Form nicht auftreten (z.B. Entfernung), wie u.a. in [22] zusammenfassend dargestellt. Weiter ist im Internet keine Trennung zwischen Signalisierung und Daten vorhanden. Dies führt dazu, dass die Messpunkte, welche die Nutzungsdatensätze erfassen, dezentralisiert implementiert werden müssen. Berücksichtigt man in diesem Zusammenhang die Tendenz zu sehr kleinen Netzzellen, die dann gemäß der Annahme dieser Arbeit IP-basiert sind, so ist eine höchst dezentrale Architektur zur Nutzungsdatenerfassung erforderlich.

Im Mobilnetzes werden gemäß Mobile IP (Kapitel 2.1.4) einem Benutzer während einer Sitzung mehrere IP-Adressen zugewiesen, sollte er zwischen den Zellen wechseln, was als Folge der Moby Dick-Referenzarchitektur dem Regelfall entspricht. In diesem Fall verfügen die dezentralen Messpunkte über keinerlei sitzungsbasierte Informationen und sind nicht in der Lage, die Nutzungsdaten frühzeitig einem Benutzer zuzuordnen. Die Ansätze aus DOCSIS, wie in Kapitel 4.5.1.1 vorgestellt, haben hier keinerlei Mobilitätsunterstützung im Sinne von Mobile IP vorgesehen und sind hierfür nicht vollständig anwendbar. Jedoch ist die Definition der zu erfassenden Datenstrukturen prinzipiell verwendbar. Weiter zeigen die Grundlagen der Preispolitik, dass es aus ökonomischen Gesichtspunkten absolut notwendig ist, jegliche Bereitstellung von Dienstgüte und somit auch die daran anschließende Abrechnung nicht dienstspezifisch, sondern benutzerspezifisch abzuwickeln. Diese benutzerspezifische Sichtweise ist eine grundlegende Voraussetzung für einen Betreiber, um verschiedene Märkte zu erschließen und den Unternehmenserfolg zu maximieren ([83], [88]). Jedoch sei an dieser Stelle festzuhalten, dass auch DOCSIS keine benutzerspezifische Parametrisierung der Datenerfassung zulässt, sondern einen dienstspezifischen Ansatz verfolgt.

Geht man von einer ALL IP-Architektur (vergl. Kapitel 5.2) aus, welche mit jeglichen leitungsvermittelnden Konzepten bricht, jedoch bzgl. Preispolitik und Benutzerverwaltung auf Konzepte des 2G-Umfeldes zurückgreift, dann ist festzuhalten, dass die IETF AAA-Architektur, wie in Kapitel 4 bzw. in [105] dargestellt, nicht ausreicht, um vor allem mobilen Benutzern integrierte Dienste kommerziell, betriebskostenminimiert und skalierbar anbieten zu können. Dies ist vor allem wegen der im Internet drastisch steigenden Dienstvielfalt ein Problem, welchem über die Definition von Dienstbündeln, wie in Kapitel 3 vorgestellt, entgegnet werden kann. Jedoch ist es bei der existierenden oligopolistischen Marktform erforderlich, bei der Benutzerverwaltung nicht unnötig hohe Fixkosten zu erzeugen, bzw. durch ein weniger effektives Vorgehen sich im Angebotsrepertoire unnötig in der Dienstvielfalt beschränken zu müssen. Siehe hierzu auch [87] und [85].

Betrachtet man sich die weit verbreiteten Mechanismen und Konzepte der Preispolitik und deren Einbindung in das Marketing, so kann man feststellen, dass aus marketingpolitischer Sicht bedeutende Elemente wie Dienstbündelung und Preisdifferenzierung, aber auch die marketingspezifischen Produktlebenszyklen und deren preispolitische Parametrisierung nicht ausreichend in der von der IETF getriebenen Internet-Architektur berücksichtigt sind. Da genau diese Einbindung das Kernziel dieser Arbeit ist, sollen im weiteren Verlauf dieser Arbeit Anforderungen, die ausschließlich technischer Natur sind, untergewichtet behandelt werden.

Bei Betrachtung der Konzepte Dienstbündelung und Preisdifferenzierung lässt sich festhalten, dass im Falle der Preisdifferenzierung benutzerspezifische Dienstbündel zu einem Produkt zusammengefasst werden. Diese Bündel werden idealerweise als Ergebnis der Marktforschung komponiert und decken in der Regel einen sehr unterschiedlichen Zeithorizont ab. Sowohl aktuelle Ereignisse, als auch andere kurzfristige Trends können hier sehr kurzfristig einen Bedarf für die Definition eines neuen Produktes feststellen und es kann nicht davon ausgegangen werden, dass der Parameter Bandbreite Kernbestandteil eines solchen Bündels ist. Es ist hier durchaus möglich, dass Bandbreite als Parameter mittel bis langfristig seine Bedeutung verliert, jedoch vom Betreiber als Bestandteil eines höher liegenden Dienstes im Bündel angeboten wird. In einem solchen Fall würde dieser Parameter für die Abrechnung keine weitere Rolle spielen.

In jedem Falle kann aus Preisdifferenzierung und der Tatsache, dass sich verschiedene Benutzer in der Regel für verschiedene Produkte interessieren, geschlossen werden, dass jegliche Abrechnung benutzerspezifisch zu erfolgen hat. Als zweite Anforderung kann festgehalten werden, dass sich die Dienstbündel sehr dynamisch ändern können sollen und sich auch die Parameter, welche Bestandteil des Preismodells sind, sehr dynamisch ändern. So ist es für einen Anbieter eines Produktes im Markt sehr wichtig, schnell auf Aktionen eines Wettbewerbers reagieren zu können.

Eine weitere Anforderung, die mehr auf einen effektiven Netzbetrieb zielt, ist die Möglichkeit, eine maximale Dienstvielfalt bei minimalen Benutzerverwaltungskosten anbieten zu können. Dies ist für einen Netzbetreiber eine sehr elementare Anforderung, da sich ein Betreiber über die Vielfalt der angebotenen Dienste vom Konkurrenten geeignet abgrenzen kann.

Unter Dienstvielfalt versteht man in diesem Zusammenhang die Anzahl der Bündel, die am Markt platziert werden. Da generell jedes inkrementelle Dienstbündel inkrementelle Betreiberkosten zur Folge hat, handelt es sich hierbei um ein Optimierungsproblem. In jedem Fall ist jedoch eine Minimierung der von dieser Relation unberührten Kosten erforderlich, da auch über eine Optimierung dieses Parameters die Anzahl der Dienste pro Benutzerverwaltungskosten maximiert werden kann.

Bezugnehmend auf Abbildung 27 kann festgehalten werden, dass das Verarbeiten der erfassten Berichte sehr rechenintensiv ist und eine aufwändige technische Infrastruktur erfordert. Somit kann die Anforderung der Nutzdatenminimierung in Verbindung mit einer optimierten Nutzdatenerfassung als weitere elementare Anforderung an ein Mobiles Internet abgeleitet werden.

5.7 Einschränkungen und Schwächen der existierenden Konzepte

Die zu erfassende Datenmenge beim Prozess der Nutzungsdatenerfassung wird neben der absoluten Benutzeranzahl und deren Verhalten von drei weiteren prinzipiellen Faktoren beeinflusst. Dies sind die Vielfältigkeit des Preismodells bzw. der Dienstbündel, zu denen sich die Kunden einbuchen können, die Vielfalt/Granularität der Dienstgüte, die Benutzer ggf. während einer Sitzung ändern können und die Möglichkeit die entsprechenden Daten, die zur Rechnungserstellung relevant sind, so früh als möglich erfassen und weiterverarbeiten zu können.

Bevor nun die vorab zusammengefassten Einschränkungen näher erläutert werden, soll einleitend festgehalten werden, dass im Internet historisch der gesamte Dienstbereitstellungsprozess keiner benutzerspezifischen Sichtweise unterliegt, sondern eine sehr starke Dienstorientierung vorzufinden ist. Eine solche Sichtweise war in einem nichtkommerziellen Netz, welches in seiner Anfangsphase weitgehend von staatlichen Einrichtungen ausgebaut wurde und dabei die kommerziellen Aspekte unberücksichtigt ließ, auch weniger erforderlich. Zusätzlich war die IETF aus ideologischer Sicht nicht auf eine Kommerzialisierung des Internets vorbereitet und hat sich erst in jüngster Zeit einem Themenbereich, der die Schnittstelle zwischen Technologie und ökonomischen Aspekten beschreibt, wie z.B. AAA, geöffnet. In diesem Zusammenhang sei auf [68] verwiesen.

Ist das Preisschema vielfältig und von einer Struktur, dass eine sehr spezifische Nutzungsdatenerfassung ermöglicht wird, müssen jegliche Details der zur Abrechnung relevanten Daten erfasst werden. Durch die nicht vorhandene Vorbereitung des Netzes auf kommerzielle Anforderungen aus Sicht eines Netzbetreibers verbunden mit der eher dienstspezifischen Sichtweise in der IETF müsste bei der Nutzungsdatenerfassung jeweils die Vereinigungsmenge aller in allen Preismodellen spezifizierten Parameter erfasst werden. D.h. die Netzarchitektur ginge davon aus, dass sich hinter jedem Benutzer derjenige mit den feingranularsten Anforderungen bzgl. der Nutzungsdatenerfassung verbirgt. Die Anzahl der Daten, die in einem solchen Fall unnötigerweise erfasst werden, hängt nun natürlich von der Vielfältigkeit der preispolitischen Maßnahmen ab. Erlaubt man z.B. einem Benutzer dynamisch die Dienstgüte zu ändern, so muss der Betreiber über eine Infrastruktur zur Nutzungsdatenerfassung verfügen, welche ein Erfassen in diesen Quantisierungsstufen ermöglicht. Ist diese Infrastruktur nicht in der entsprechenden Granularität verfügbar, macht ein solches Element in der Preispolitik keinen Sinn, da kein Rückkopplungsmechanismus über den Preis zum Benutzer implementiert werden kann. Erfahrungen im Internet haben gezeigt, dass Benutzer ohne einen solchen Rückkopplungsmechanismus über den Preis immer die bestmögliche Qualität wählen würden.

Eine weitere Schwäche existierender Ansätze ist das Fehlen eines Konzeptes, das die unterschiedlichen Paradigmen zwischen dem Internet und der Telefonie bezüglich der Zahlungsströme zum einen zwischen Betreiber und Kunden und zum anderen zwischen Betreibern regelt. Im aktuellen Telefonnetz bezahlt eine Sitzung prinzipiell der Anrufende. Beim Internet bezahlt jeder Kommunikationsteilnehmer seinen eigenen Anteil direkt an den Netzbetreiber. Dieses Konzept hätte den Vorteil, dass Zahlungsströme zwischen Netzbetreibern und die dazugehörigen Abrechnungsinfrastrukturen in diesem statischen Fall ersatzlos wegfallen. Wenn

jedoch ein mobiler Benutzer die Ressourcen bei einem Roaming Partner konsumiert, ist ein Zahlungsfluss zwischen diesen beiden Betreibern erforderlich.

Weiter ist die Trennung zwischen Signalisierung und Daten im Telefonnetz vorteilhaft, bindet man die Abrechnungsinfrastruktur an die Signalisierung. Dies ist konzeptionell im Internet nicht möglich. Somit muss die Nutzungsdatenerfassung im Mobilien Internet verteilt erfolgen.

Die oben erwähnten, eher generischen, Schwächen der existierenden Lösungsansätze haben ihre Ursache u.a. darin, dass die existierende IETF-Architektur über kein geeignetes Sitzungskonzept verfügt, das eine Konsolidierung der aufbereiteten Nutzungsdaten zu einem sehr frühen Zeitpunkt, idealerweise bei ihrer Entstehung, ermöglicht. Die aktuelle IETF AAA-Architektur nach Kapitel 4.4 billigt z.B. dem AAA-Server, der die Rolle des AAA.f ausfüllt, keinerlei Intelligenz zu und reduziert seine Funktionalität dahingehend, dass jegliche Daten lediglich weitergeleitet werden. Das Nichtvorhandensein einer intelligenten Bearbeitung der Nutzungsdaten in einer auswärtigen Netzdomäne ist hierbei ein gravierender Nachteil. Eine proaktive Rolle der AAA-Infrastruktur würde sich in einem solchen Umfeld als sehr vorteilhaft erweisen, da hier vorzeitig eine Vorauswahl der für die weitere Verarbeitung der Nutzungsdaten zum Zweck der Rechnungserstellung erfolgen könnte. Dies beinhaltet auch eine Delegation wesentlicher Funktionen der Heimatdomäne an die auswärtige Netzdomäne, also vom AAA.h zum AAA.f. Weiter ist festzuhalten, dass jegliche Prepaid-Unterstützung in der existierenden AAA-Architektur nicht ausreichend berücksichtigt wurde. Es wurden zwar jüngst Vorschläge in die IETF eingebracht (siehe hierzu [82] und [46]), jedoch können diese im Sinne dieser Arbeit als nicht benutzerorientiert betrachtet werden.

Die IETF RTFM-Architektur (Kapitel 4.4.5) ist sehr komfortabel einsetzbar, jedoch ist hier ein benutzerspezifisches Erfassen der Daten ebenso wenig vorgesehen, wie deren detaillierte Anbindung an das AAA-System. Da die detaillierte Konfiguration eines Meters letztendlich von dem benutzerspezifischen Preismodell bestimmt wird, kann zusammengefasst eine Schwäche der aktuellen Internet-Architektur festgestellt werden, die darin liegt, dass keine Signalisierung der preispolitisch relevanten Parameter auf die Schicht 3-Netzarchitektur, bei Bedarf über Betreiber Grenzen hinweg, in einer standardisierten und im betriebswirtschaftlichen Sinne effektiven Art und Weise verfügbar ist.

Die Policy-basierte Accounting-Architektur hat keinen konsequent benutzerorientierten Ansatz, sieht jedoch eine effiziente dienstspezifische Nutzungsdatenerfassung mit nachgelagerter Aufbereitung der Nutzungsdaten vor. Weiter ist der Vorschlag nicht in die Mobilitätsverwaltung integriert, was eine effiziente Kommunikation im Mobilitätsfall erschwert. Darüber hinaus fehlt diesem Vorschlag ein Sitzungskonzept für die Nutzungsdatenerfassung über verschiedene Zugangsnetze innerhalb eines Betreibers und zwischen Betreibern.

Weiter ist festzuhalten, dass in einem Mobilien Internet die Zuordnung zwischen der Nutzungsdatenerfassung und des Sitzungsmanagements, wie z.B. in [41] dargestellt und im Prinzip von der IETF AAA-Architektur nach Kapitel 4.4.1 und Kapitel 4.4.2 übernommen wurde, zu erheblichem Mehraufwand beim Billing führt.

5.8 Zusammenfassung

In diesem Kapitel wurden die für den weiteren Verlauf dieser Arbeit relevanten Annahmen getroffen, die eine entsprechende Fokussierung für den Rest dieser Arbeit zulassen. Die Annahmen beschreiben technisch das Mobile Internet und gehen auf ein exemplarisches Preismodell näher ein. Dieses Preismodell ist dazu geeignet, sowohl zwischen Kunde und Betreiber, als auch zwischen zwei Betreibern implementiert zu werden. Die beiden Betreiber kön-

nen hier horizontal oder vertikal kooperieren. In Tabelle 6 wurde hierfür exemplarisch ein Spiele-Anbieter erwähnt. Bei Betrachtung des Preismodells wurde im weiteren Verlauf dieses Kapitels herausgearbeitet, dass eine benutzerorientierte Sichtweise elementar ist und in existierenden Ansätzen bislang nicht ausreichend implementiert wurde.

Um zu einer benutzerspezifischen verteilten Nutzungsdatenerfassung zu gelangen, sind basierend auf den aktuellen Ansätzen innerhalb der IETF AAA-, IETF RTFM- und IRTF AAAArch-Arbeitsgruppen einige Modifikationen erforderlich, die sich prinzipiell in drei Bereiche einteilen lassen. Dies ist die Entwicklung einer Signalisierungsarchitektur, welche die dezentrale Nutzungsdatenerfassung benutzerspezifisch steuert, als auch die Nutzungsdaten benutzerspezifisch konsolidiert. Weiter ist ein Sitzungskonzept erforderlich, das die benutzerspezifische Nutzungsdatenkonsolidierung unterstützt. Hierbei ist es wichtig, dass dieses Konzept über Betreibergrenzen hinweg arbeitet. Um sowohl Sitzungsarchitektur und Signalisierungsarchitektur entsprechend zu unterstützen, muss in den Zugangsroutern eine entsprechende Architektur implementiert werden. Im folgenden Kapitel sollen nun diese drei Forderungen näher vorgestellt werden.

6 Vorschläge zur Weiterentwicklung von Mechanismen im kommerzialisierten Mobil Internet

6.1 Einleitung

In Kapitel 5 wurde die Notwendigkeit von drei wesentlichen Verbesserungsvorschlägen zur Kommerzialisierung des Mobil Internet herausgearbeitet. Dies sind eine Signalisierungsarchitektur, ein Sitzungskonzept und eine Netzknotenarchitektur für Zugangsrouten, welche in [67] erstmals vorgestellt wurde und nun detaillierter spezifiziert werden.

Die Signalisierungsarchitektur basiert auf dem im Projekt Moby Dick entwickelten Ansatz mit den für diese Arbeit spezifischen Erweiterungen. Das Sitzungskonzept ist neu und lehnt sich an die Vorschläge aus [68] und [72] an. Die Architektur der Netzknoten erfolgt auf Basis der IETF RTFM-Architektur nach Kapitel 4.4.5 mit geeigneter Anbindung an die IETF/IRTF AAA-Architektur nach Kapitel 4.4 mit spezifischen Erweiterungen. Diese Erweiterungen erfolgten auf Basis der im Projekt Moby Dick entworfenen Knotenarchitektur. Hierzu wurde die IETF RTFM-Architektur um eine benutzerspezifische Nutzungsdatenerfassung erweitert. Weiter wurde ein Sitzungskonzept für mobile Benutzer definiert, welches den Prozess der Nutzungsdatenerfassung in einer echten heterogenen ALL-IP-Netzinfrastruktur integriert.

Die drei Kernvorschläge adressieren den wesentlichen Mangel einer fehlenden Integration kommerzieller Prozesse in das Mobile Internet. Somit ist der Kernbeitrag dieser Arbeit darin zu sehen, dass fehlende Schlüsselmechanismen und -funktionen, welche hier vorgestellt werden, den Migrationsprozess in Richtung Mobiles Internet nicht bei den aktuellen Vorstellungen von 3GPP, 3GPP2 und Parlay enden lassen (siehe hierzu auch die Kapitel 2.2.2.4, 2.2.2.5 und 4.7.3).

6.2 Entwurf einer Signalisierungsarchitektur

Um zunächst eine detaillierte Signalisierungsarchitektur vorstellen zu können, sollen zunächst die beiden relevanten Schlüsselszenarien aus Benutzersicht detailliert dargestellt werden. Diese Szenarien basieren auf der im Projekt Moby Dick erarbeiteten Lösung und wurden im Rahmen dieser Arbeit mit einigen Erweiterungen versehen. Bevor nun die detaillierten Mechanismen vorgestellt werden, soll zunächst kurz eine Übersicht über Schlüsselszenarien des Mobil Internets gegeben werden.

6.2.1 Schlüsselszenarien für das zukünftige Mobile Internet

Die nun vorgestellten Schlüsselszenarien Registrierung und Netzwechsel sind im Vergleich zum Moby Dick-Ansatz um eine Übertragung der Kontextinformation für die Nutzungsdatenerfassung bei Zellwechsel erweitert worden. Ebenso ist im Vergleich zu [39] die Rolle des QoS-Brokers modifiziert worden. Bevor nun diese Szenarien vorgestellt werden, soll zunächst detailliert auf die verschiedenen Beziehungen zwischen den Anbietern von Netz bzw. IT-Diensten untereinander und jeweils mit dem Endkunden eingegangen werden.

6.2.1.1 Arten der Föderierung

Unter Föderierung bezeichnet man im Kontext dieser Arbeit das organisierte Zusammenarbeiten verschiedener Netz-/Dienst-/Inhalte-Anbieter, um für den Benutzer einen Dienst erbringen zu können, den ein einzelner Anbieter nicht in der Lage ist zu leisten. Es ist davon auszugehen, dass die Föderierung von administrativen Netzen zwischen Betreibern verschiedene Ausprägungen annehmen kann. So ist zwischen folgenden drei Arten der Föderierung zu unterscheiden:

Bei der Föderierung erster Art ist ein Netzbetreiber nicht in der Lage, für seine Kunden einen Dienst flächendeckend anzubieten und vereinbart mit einem ggf. konkurrierenden Betreiber, dass in einem begrenzten und genau definierten geographischen Bereich der Dienst eines anderen Betreibers verwendet wird. Diese Art der Föderierung bleibt dem Kunden in der Regel verborgen. Weiter ist es für den Kunden z.B. aus der Rechnung nicht ersichtlich, dass der Dienst eines anderen Betreibers benutzt wurde. Es ist sogar im Sinne des Betreibers, dem Benutzer vorzuenthalten, dass die flächendeckende Dienstleistung nicht aus eigener Kraft geleistet werden kann. Diese Föderierung kommt in Deutschland sehr häufig bei den kleineren Mobilfunk Anbietern vor, die in bestimmten Bereichen, in denen sich der Netzausbau wirtschaftlich nicht lohnt, das Netz eines Konkurrenten mitbenutzen.

Als Föderierung zweiter Art soll hier die Art der Föderierung eingeführt werden, bei dem zwei Dienstanbieter ein offenes Abkommen eingehen, welches die Kunden aus dem jeweils anderen Netz entsprechend dem Abkommen behandelt. Dieses Abkommen (wenigstens Teile daraus) werden offensiv zu Vermarktungszwecken verwendet. Oftmals findet man diese Art der Föderierung zwischen Landesgrenzen vor. Auch in diesem Fall kann davon ausgegangen werden, dass die Netzknoten untereinander Daten austauschen. Dies entspricht dem Roaming-Fall bei Mobilkommunikationsbetreibern.

Die Föderierung dritter Art soll hier ein Abkommen zwischen einem Netzbetreiber und einem Anbieter höherwertiger Dienste wie z.B. Inhalte bezeichnet werden. Hier konkurrieren die Betreiber nicht, sondern der Netzbetreiber bietet Dienste des föderierten Drittanbieters unter eigener Rechnung dem Kunden an. Dies kann ggf. exklusiv sein, d.h. das Angebot des Drittanbieters ist auf das Netz eines Netzanbieters limitiert. Aus Sicht eines Netzbetreibers ist angestrebt, den Dienst des jeweiligen Inhalte-Anbieters exklusiv anbieten zu können, der Inhalte-Anbieter strebt dagegen eher an, sein Angebot nicht auf das Netz eines Betreibers beschränken zu müssen.

6.2.1.2 Registrierung

Als Registrierung bezeichnet man den Vorgang, bei dem sich ein mobiles Endsystem erstmalig an einem Netz anmeldet. Dies geschieht mittels Übertragung von Vertraulichkeitsinformationen vom Endsystem zur Authentisierungsinstanz, welche innerhalb der IETF vom AAA-Server eingenommen wird. Der Registrierungsprozess wird gestartet, nachdem ein Endsystem erfolgreich eine Care-off-Adresse (*CoA*) nach Kapitel 2.1.4.4 bezogen hat. Optional kann sichergestellt werden, dass diese erhaltene/konfigurierte Adresse nicht schon vergeben ist. Jedoch kann auch hier aus Zeitgründen vereinfacht eine Adresskonfiguration auf Zufallsbasis erfolgen [8]. Generell ist festzuhalten, dass die Registrierung immer an ein Zeitintervall gebunden ist, nach dem sich ein Endsystem unabhängig von der Kommunikation neu registrieren muss.

Der prinzipielle Signalisierungsablauf der Registrierung ist in Abbildung 42 detailliert dargestellt. Damit ein vagabundierender Teilnehmer voll in eine Mobilitätsarchitektur integriert ist, müssen zwei Voraussetzungen erfüllt sein. Zum einen muss seine neu bezogene CoA weltweit sichtbar sein, d.h. dritte Kommunikationsteilnehmer müssen ihn erreichen können und er muss selbst in der Lage sein, Daten entsprechend versenden zu können. Beide Voraussetzungen sind nach Bezug einer CoA noch nicht gegeben, da der Zugangsrouten für diese CoA noch keine Ressourcenfreigabe vollzogen hat. Die prinzipielle Erreichbarkeit eines mobilen Endsystems wird, wie in Kapitel 2.1.4 beschrieben, durch das Registrieren beim Heimagenten mittels eines Binding Updates (BU) und die Möglichkeit, eigene Pakete zu versenden, wird mittels einer Freischaltung der Ressourcen im Zugangsrouten erreicht (was das Versenden eines BUs mit einschließt).

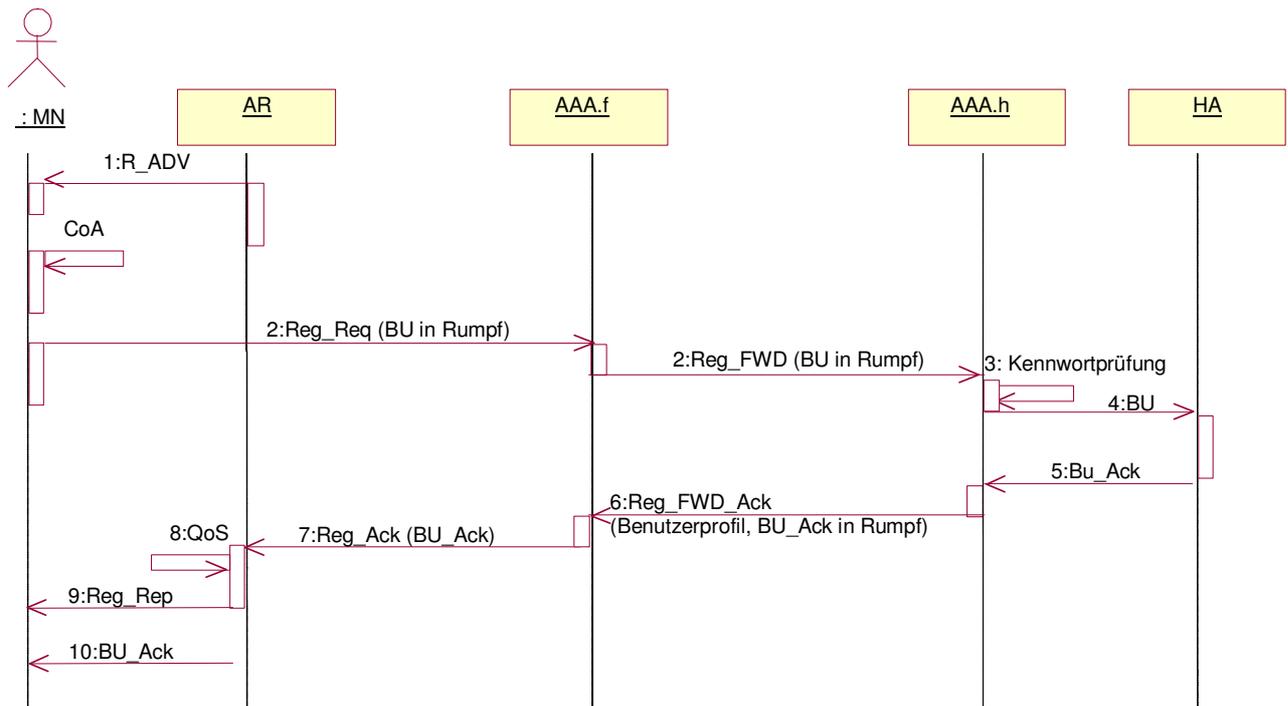


Abbildung 42: Registrierung

Nach Erhalt des *Router Advertisement* (1)/Bezug der *CoA* ist ein *MN* lediglich in der Lage, eine Registrierungsnachricht auszusenden. Eine Registrierungsnachricht ist ein Datenpaket, welches an den Zugangsrouten (AR) adressiert ist. Der Zugangsrouten schickt dieses Paket weiter zu einem fest konfigurierten AAA-Server (2), der nun die Rolle eines AAA.f-Servers nach Kapitel 4.4 einnimmt. Hier können einige wenige Ausnahmen gestattet werden, wie z.B. das Aussenden von Notrufen. In dieser Nachricht ist per Piggybacking (Huckepack-Mechanismus) der BU angehängt (4), (5).

Der AAA.f-Server entnimmt die Informationen bezüglich der Herkunft des Teilnehmers und prüft zunächst, ob mit der Domäne, die der Anfordernde als Heimatdomäne angegeben hat, entsprechend gültige Abkommen bestehen. Ist dies der Fall, wird die Registrierungsanforderung zu dem für das Heimatnetz verantwortlichen AAA.h weitergeleitet (2). Die genaue Adresse geht aus den statischen Vereinbarungen zwischen den involvierten Netzbetreibern hervor. Am dortigen AAA-Server, der die Rolle eines AAA.h einnimmt, wird der Benutzer authentisiert (3). Nach erfolgreicher Authentisierung erfolgt die Autorisierung in Form eines Übertrags des Benutzerprofils (6). Hier wird eindeutig spezifiziert, welche Dienste und Ressourcen dem Benutzer zur Verfügung gestellt werden dürfen.

Diese Informationen werden verwendet, um den Zugangsrouten entsprechend des Vertrags zwischen Benutzer und Primäranbieter zu konfigurieren (8). D.h. ab diesem Zeitpunkt ist der Teilnehmer, der das Endsystem mit der entsprechenden *CoA* verwendet, in der Lage, sich am Heimatnetz zu registrieren (*BU* darf passieren) und Ressourcen entsprechend den Vertragsparametern zu konsumieren. In diesem Zusammenhang soll erwähnt werden, dass ein mobiles Endsystem nach Abschluss von Nachricht (9) weltweit sichtbar ist, d.h. jeglichen Verkehr empfangen kann und frühestens nach Abschluss der Nachricht (8) Daten ins Internet senden kann. Im Folgenden sollen die Parameter von ausgewählten Nachrichten explizit spezifiziert und erläutert werden.

Router Advertisement und CoA-Konfiguration

Die Router Advertisement-Nachricht (*R_ADV*) sollte zwei verschiedene Parameter beinhalten. Dies ist zum einen die eigene, weltweit gültige IP-Adresse. Durch den Empfang dieser Adresse ist ein mobiles Endsystem gemäß der Mobile IP-Spezifikation nach Kapitel 2.1.4 in der Lage, eine gültige *CoA* zu konfigurieren. Nach [104] erscheint es am effektivsten, dass diese Konfiguration auf Basis einer Zufallszahl erfolgt.

Da nicht davon auszugehen ist, dass ein Endsystem basierend auf dieser Nachricht auf einen speziellen Netzbetreiber schließen kann, wird als weiterer Parameter zur Integration in diese Nachricht eine Netzbetreiber-ID vorgeschlagen. Diese ID ist geeigneterweise weltweit gültig. Eine 48 Bit lange ID bietet hier weltweite Eindeutigkeit. Die genaue Spezifikation wird im Rahmen dieser Arbeit nicht weiter verfolgt, jedoch würde das Hinzufügen dieses Parameters den Netzauswahlprozess beschleunigen und aus kommerzieller Sicht vereinfachen.

Registrationsanforderung

Die Registrierungsanforderung (*Reg_Req*) enthält neben einem Benutzernamen einen Vertraulichkeitsnachweis, welcher benutzerspezifisch ist und entweder manuell eingegeben oder aber z.B. automatisch aus einer vom Benutzer eingesetzten SIM-Karte ausgelesen und übermittelt wird.

Der Benutzername besteht aus einer eindeutigen Benutzer-ID, einer Betreiber-ID des Primärbetreibers und der aktuellen *CoA*. Somit ähnelt diese ID strukturell einer E-Mail-Adresse, welche im Prinzip aus den beiden erstgenannten Bestandteilen besteht. Das Erkennen dieser Nachricht beim Empfänger könnte dadurch erleichtert werden, dass die Pakete mit einer standardisierten Transportadresse (Kapitel 2.1.3.1) versehen werden.

Registrationsweiterleitung zur auswärtigen Domäne

Kommt die Registrierungsanforderung am Zugangsrouten an, so wird diese eindeutig als solche identifiziert und direkt an den *AAA.f*-Server weitergeleitet. Diese Weiterleitungsfunktion erfolgt bedingungslos. Zur Eingrenzung von so genannten DoS-Angriffen (Denial of Service) kann hier die Ankunftsrate von Registrierungsanforderungen mit der identischen *CoA* begrenzt werden. Darüber hinausgehende Maßnahmen werden im Rahmen dieser Arbeit nicht weiter verfolgt. In jedem Fall erfolgt diese Weiterleitung nur zu *AAA*-Servern von Betreibern, mit denen eine entsprechende Förderierung nach Kapitel 6.2.1.1 besteht.

Registrationsweiterleitung zur Heimatdomäne

Der *AAA.f*-Server wird im Rahmen dieser Arbeit mit mehr Funktionalität versehen als in [78] vorgesehen. Zunächst wird vorgeschlagen, dass der *AAA.f*-Server bei allen eingehenden Nachrichten die Betreiber-ID überprüft. Hier sind nun drei verschiedene Fälle möglich:

- Betreiber-ID entspricht eigener Betreiber-ID, welche nicht als Roaming-Partner hinterlegt ist. In diesem Falle wird die Anfrage verworfen. Ggf. wird eine negative Rückantwort zurückgegeben
- Betreiber-ID wird als Betreiber-ID eines Roaming-Partners identifiziert. In diesem Fall wird die Nachricht an einen festgekoppelten *AAA.h*-Server weitergeleitet. Hier ist jeder Betreiber-ID ein eindeutiger *AAA.h*-Server zugeordnet.
- Betreiber-ID entspricht der eigenen Netz-ID. In diesem Fall agiert der *AAA.f*-Server als *AAA.h*-Server und überprüft, ob die angegebene Benutzer-ID gültig ist. Je nach Ergebnis dieser Überprüfung erfolgt eine positive oder negative Antwort. Im Falle einer positiven Antwort wird zusätzlich ein Benutzerprofil der Antwort beigefügt.

Überprüfung der Benutzerangaben

Hier wird die Benutzer-ID auf ihre Gültigkeit überprüft. Bei positivem Ergebnis wird an die Antwort ein Benutzerprofil angefügt. Eine negative Antwort wird durch ein entsprechendes AVP (vergl. Kapitel 4.4.1) indiziert.

Struktureller Aufbau des Benutzerprofils

Prinzipiell besteht ein Benutzerprofil aus drei verschiedenen Teilen, die jeweils inkrementell zueinander aufgebaut sind. In der untersten Hierarchie ist ein netzorientiertes „Profil 2“ definiert, das alle Daten beinhaltet, welche für den Zugangsrouten (AR), der die Ressourcen für ein mobiles Endsystem bereitstellt, erforderlich sind. Weiter wird ein „Profil 1“ definiert, welches alle Informationen enthält, die zwar wie bei „Profil 2“ vom AAA.h-Server zum AAA.f-Server (vergl. Kapitel 4.4) transferiert werden, dort aber verbleiben.

Im Benutzerprofil, das in der Heimdomäne verbleibt, werden zusätzlich zu „Profil 2“ und „Profil 1“ weitere Informationen gespeichert, welche nicht an den AAA-Server eines Roaming-Partners weitergegeben werden. Ein Beispiel hierfür sind Namen, Anschrift und ggf. Bankverbindung eines Benutzers, aber auch andere Informationen, auf die noch in Kapitel 6.2.2 näher eingegangen wird. Der Strukturelle Aufbau des Benutzerprofils ist in Abbildung 43 dargestellt.



Abbildung 43: Struktur des Benutzerprofils

Antwort/Konfiguration des Zugangsrouters

Die Antwort des AAA.h-Servers beinhaltet im positiven Fall Informationen über die vertraglichen Details des Benutzers. Diese werden ggf. zwischen den beiden AAA-Servern gemäß dem Roaming-Abkommen der beiden involvierten Betreiber abgeglichen. Am Zugangsrouten erfolgt neben der Konfiguration der Dienstgüte des Benutzers eine Konfiguration der Nutzungsdatenerfassung. Konkret sind hier Dienstklassen, wie in Kapitel 5.5 vorgeschlagen, festgelegt. Anhand dieser Dienstklassen wird die Nutzungsdatenerfassung entsprechend des Vertrages auf die aktuelle CoA des Benutzers konfiguriert.

6.2.1.3 Netzübergang innerhalb einer administrativen Domäne

Ein Netzübergang innerhalb einer administrativen Domäne (HO) beschreibt den Übergang zwischen zwei Zugangsrouten (gleiche oder verschiedene Netztechnologie), welche nach erster Art föderiert sind (vergl. Kapitel 6.2.1.1). Es kann davon ausgegangen werden, dass sich die Zugangsrouten eines Netzes vertrauen und statische Sicherheitsbeziehungen zueinander unterhalten. Somit können Informationen, die für das Mobilitätsmanagement erforderlich sind, direkt zwischen diesen Zugangsrouten ausgetauscht werden.

Aus Gründen der Effektivität wird im Rahmen dieser Arbeit vorgeschlagen, dass ein AAA.h-Server jegliche Authentisierungs- und Autorisierungsvorgänge an den AAA.f-Server delegiere-

ren kann. In einem solchen Fall wird lediglich der Binding Update (vergleiche Kapitel 2.1.4) an das Heimatnetz übertragen. Abbildung 44 zeigt die vorgeschlagene Signalisierung. Diese basiert auf Ergebnissen des Moby Dick-Projektes, jedoch mit veränderter Parametrisierung der Nachrichten. Im Folgenden werden ausgewählte Nachrichten inklusive ihrer Parameter näher erläutert.

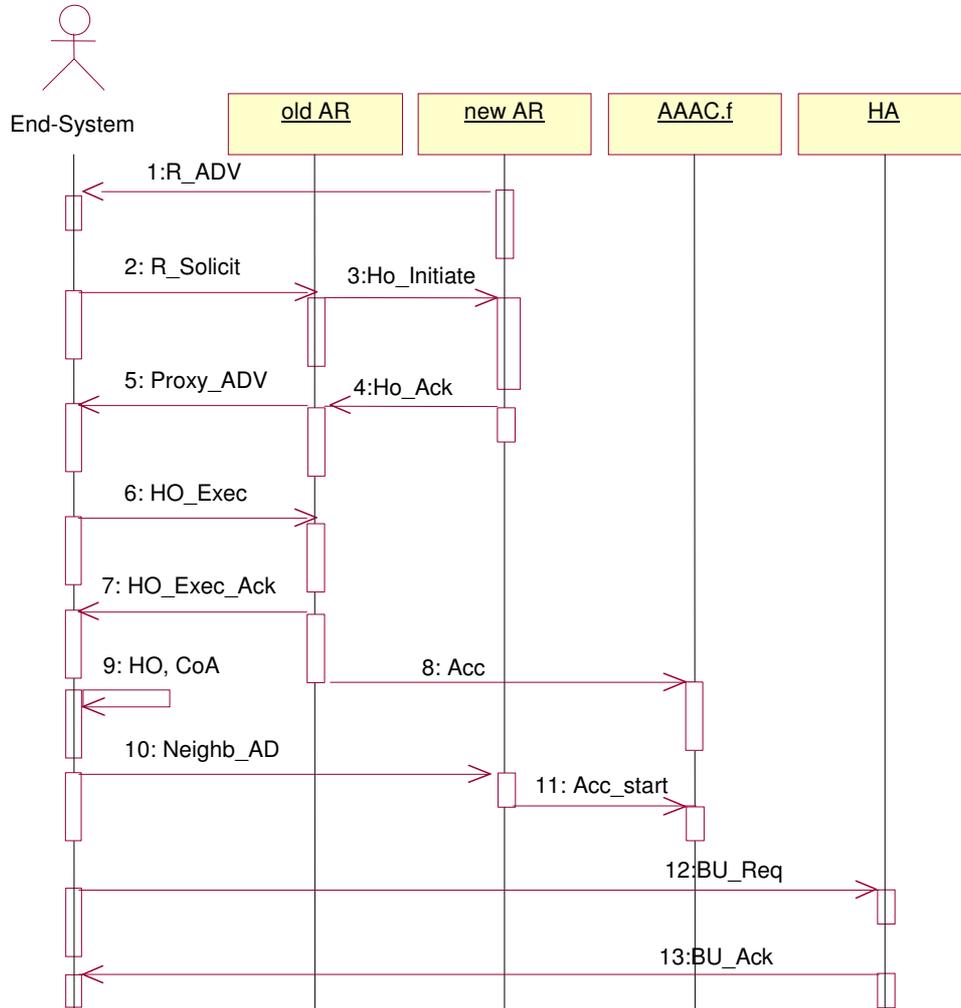


Abbildung 44: Netzübergang innerhalb eines Betreibers

Router Solicitation

Die Router Solicitation-Nachricht (2) wird von einem mobilen Endsystem ausgesendet, wenn ein interner Entscheidungsprozess zum Ergebnis gekommen ist, von einem Netz in ein anderes zu wechseln. Der detaillierte Mechanismus dieser Entscheidungsfindung soll im Rahmen dieser Arbeit nicht näher verfolgt werden.

HO-Initialisierung

Die HO-Initialisierung (3) ist eine Nachricht, die vom alten Zugangsrouten zum neuen Zugangsrouten gesendet wird. Die Parameter, die dieser Nachricht beigefügt werden, sind Meter-Konfiguration, QoS-Parameter, aktuelle CoA, neue CoA und „Profil 2“ gemäß Abbildung 43.

Aufbereitung der Nutzungsdaten

Die aufbereiteten Nutzungsdaten werden mittels des Diameter-Protokolls übertragen (8). Hier wird vorgeschlagen, neue AVPs basierend auf der in Kapitel 4.4.1 vorgestellten Architektur zu definieren, welche diese aufbereiteten Nutzungsdaten sitzungsbasiert weiterleiten.

Initialisierung einer neuen Sitzung zur Nutzungsdatenerfassung

Eine neue Sitzung zur Nutzungsdatenerfassung in einem Zugangsrouter wird mittels (11) signalisiert. Hier informiert ein Zugangsrouter den entsprechenden AAA-Server, dass er einen neuen Benutzer in seiner Netzzelle hat, für den nun ein neuer Prozess der Nutzungsdatenerfassung beginnt.

6.2.1.4 Netzübergang zwischen administrativen Domänen

Beim Netzübergang zwischen administrativen Domänen unterhalten die involvierten Netzbetreiber eine Geschäftsbeziehung zueinander. Dieser Fall entspricht einer Föderierung zweiter Art nach Kapitel 6.2.1.1. Es wird in einem solchen Fall nicht davon ausgegangen, dass vertrauliche, benutzerspezifische Daten zwischen den Netzknoten beider Betreiber ausgetauscht werden, da prinzipiell Dreiecksbeziehungen zwischen Betreibern, wie in Abbildung 45 dargestellt, nicht ausgeschlossen werden können. Es ist also erforderlich, eine neue Beziehung zwischen den involvierten Parteien gemäß einem Roaming-Abkommen zu etablieren, welche ausschließlich über die AAA-Infrastruktur vorgenommen wird. Der Betreiber, der in einem solchen Fall einen neuen Teilnehmer aufnimmt, muss direkt bei dessen Heimatdomäne die entsprechenden Details des Teilnehmers beziehen. Somit ist in diesem Fall eine neue Registrierung erforderlich. Dieser Fall soll in folgendem Beispiel (Abbildung 45) noch einmal näher verdeutlicht werden.

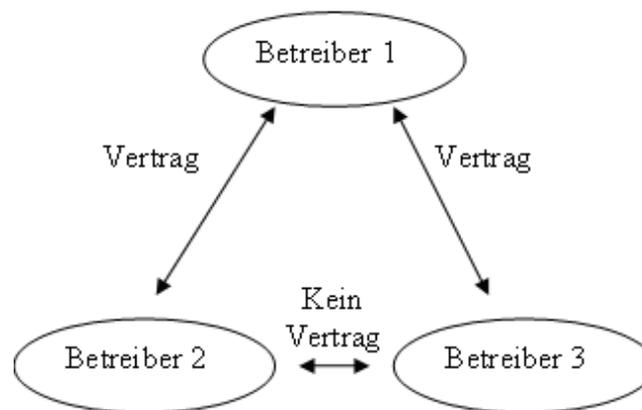


Abbildung 45: Dreiecksbeziehungen der Netzbetreiber

Möchte ein Benutzer, welcher bei Betreiber 1 zu Prepaid-Konditionen unter Vertrag steht, vom Netz des Betreibers 2 zum Netz des Betreibers 3 wechseln, kann nicht angenommen werden, dass sich Netzknoten von Betreiber 2 und Betreiber 3 vertrauliche Daten wie z.B. das Restbudget austauschen. Technisch wäre dies zwar denkbar, jedoch sind ohne eine vertragliche Beziehung dieser beiden Betreiber eventuelle Haftungsfragen nicht geregelt, welche dann aus organisatorischen Gründen eine solche Kommunikation verhindern.

Im Rahmen dieser Arbeit wird der Optimierung dieses Falls keine besondere Bedeutung beigemessen. Es wird in einem solchen Fall von einer neuen Registrierung ausgegangen, da Betreiber 3 Benutzer von Betreiber 1 akzeptiert.

Durch die gegebene Komplexität eines solchen Vorgangs ist anzunehmen, dass der Netzübergang zwischen zwei administrativen Domänen weniger häufig vorkommt. Aktuell ist dies z.B. bei einem Grenzübergang der Fall. Diese Annahme soll durch einen weiteren Aspekt weiter untermauert werden:

Es ist davon auszugehen, dass ein Netzbetreiber keinerlei Interesse hat, einen Kunden an einen Konkurrenten zu verlieren, sondern wird alles daran setzen, den Kunden an sich zu bin-

den. Somit werden die Netze dort zusammengeschaltet, wo ein Netzbetreiber keine Flächendeckung erreicht. In diesem Fall (Föderierung erster Art aus Kapitel 6.2.1.1) ist kein Betreiberwechsel für den Benutzer erkennbar und es gilt das Szenario nach Kapitel 6.2.1.3.

6.2.2 Genereller Ansatz

Abbildung 46 zeigt die prinzipielle Verbesserung im Gesamtprozess. Anstatt die Nutzungsdaten dem Benutzer erst im Charging Prozess zuzuführen, ist es im Mobilien Internet vorteilhaft, den Gesamtprozess benutzerorientiert zu gestalten.

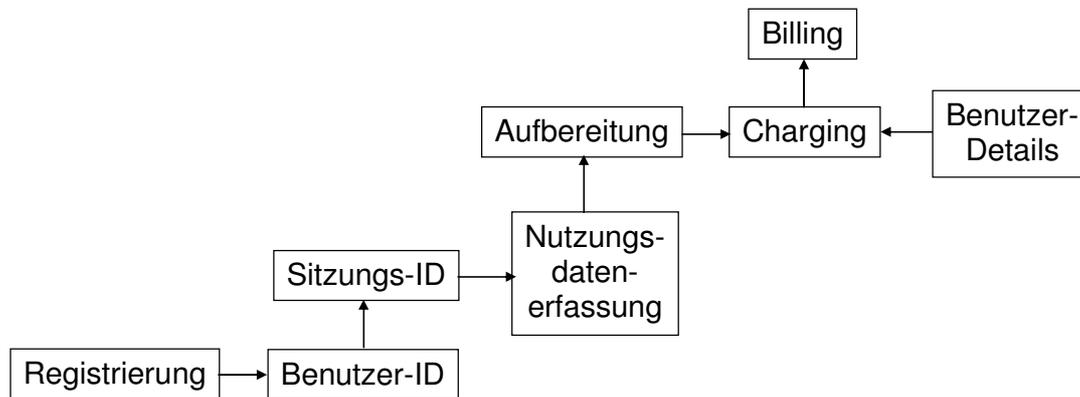


Abbildung 46: Konzeptioneller Vorschlag

6.2.3 Prinzipielle Konfigurationssignalisierung

Um nun eine Signalisierung, welche den Prozess zur Erfassung der Nutzungsdaten inklusive der nachgelagerten Aufbereitung dieser Daten steuert, vorzuschlagen, wird von folgendem Modell ausgegangen (Abbildung 47). Hier sind drei Domänen involviert, bei denen es sich um die Domäne des Heimatnetzes handelt und zwei weitere Domänen, welche als Roaming-Partner der Heimatdomäne agieren. Der mobile Teilnehmer befindet sich zunächst in einer Zelle der Domäne 1 und erreicht nach zwei Netzübergängen innerhalb dieser Domäne die Domäne 2. Es muss zwischen folgenden zwei Übertragungswegen unterschieden werden:

- Weg vom AAA.f-Server zum AAA.h-Server, der als Inter-Domänen Signalisierung realisiert ist,
- Weg zwischen Zugangsroutern (AR) und AAA.f-Server, der als Intra-Domänen Signalisierung realisiert ist

Weiter wird zwischen den beiden Szenarien „Registrierung“ und „Netzübergang innerhalb einer administrativen Domäne“ unterschieden. Der Fall Netzübergang zwischen zwei administrativen Domänen hat eine neue Registrierung zur Folge.

Die Konfigurationssignalisierung schließt sich an die Registrierung eines mobilen Teilnehmers an und wird im Rahmen dieser Arbeit spezifiziert und integriert. Sollte ein AAA.h-Server eine Registrierungsanforderung erhalten haben, so werden nach erfolgreicher Authentifizierung die zur Konfiguration der Netzknoten bzgl. einer optimierten Nutzungsdatenerfassung erforderlichen Informationen übertragen. Diese Übertragung erfolgt innerhalb der Registrierungsnachrichten aus Kapitel 6.2.1.2 (Nachricht 6 und Nachricht 7 aus Abbildung 42), also unter Verwendung des Diameter-Protokolls mit spezifischen AVPs. Wie in Abbildung 47 gezeigt, wird der Status der benutzerspezifischen Sitzungen entsprechend dem Bewegungsverhalten des Benutzers innerhalb einer administrativen Domäne direkt von Zugangsroutern zu Zugangsroutern in Echtzeit weitergereicht, also in den Netzübergang nach Abbildung 44 integriert.

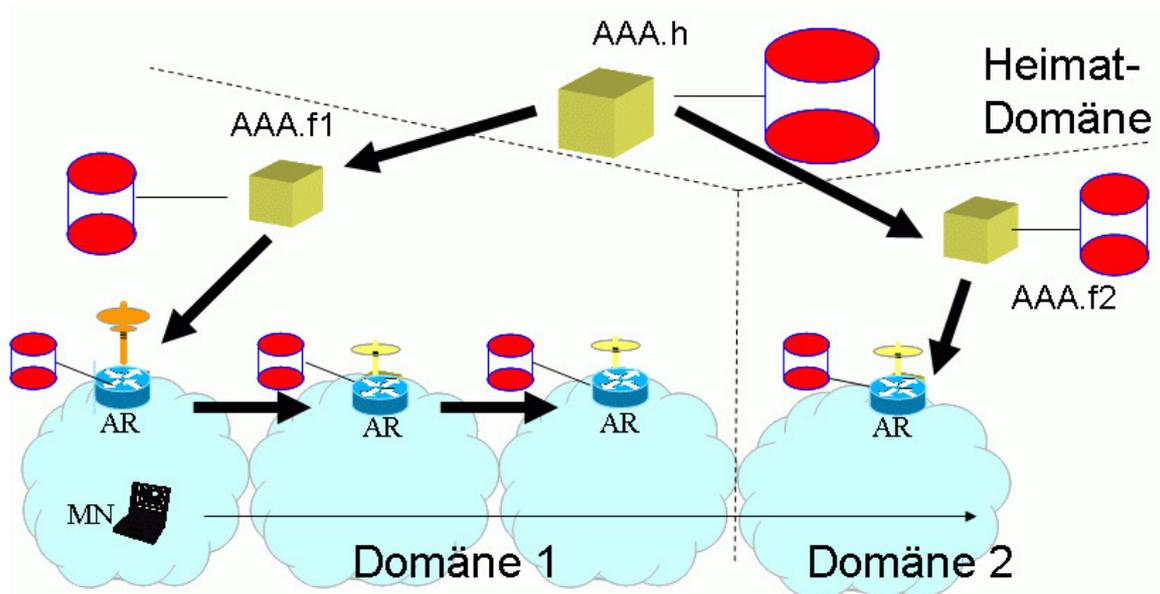


Abbildung 47: Prinzipielle Konfigurationssignalisierung

6.2.4 Prinzipielle Signalisierung der Nutzungsdaten

Die Nutzungsdaten werden dann je nach Anforderung in einem dreistufigen System zurück zur Heimatdomäne übertragen. Auf jeder dieser drei Stufen ist prinzipiell eine Datenkonsolidierung möglich. Diese Konsolidierung erfolgt benutzerspezifisch und orientiert sich am Vertrag des Benutzers mit dem Netzanbieter. Die Kommunikation erfolgt ebenfalls unter Verwendung des Diameter Protokolls (Kapitel 4.4.1). Abbildung 48 zeigt den prinzipiellen Datenfluss.

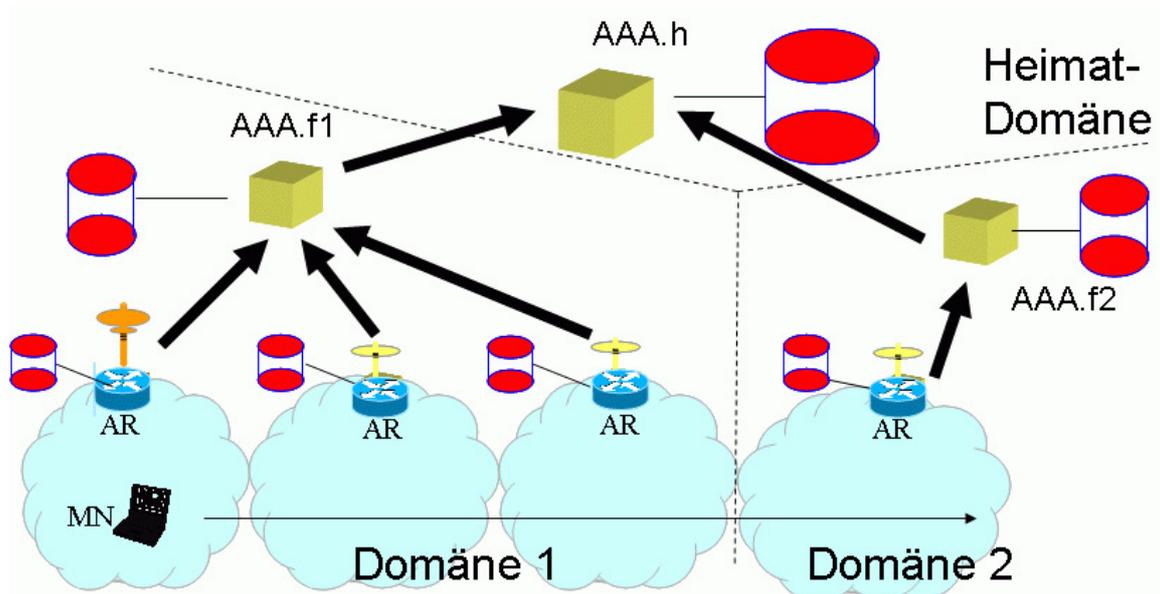


Abbildung 48: Datenfluss der Nutzungsdatenerfassung und Aufbereitung

Die Zeitpunkte der Datenübertragung von den Zugangsroutern zum AAA.f-Server und dann ggf. weiter zum AAA.h-Server werden benutzerindividuell festgelegt und orientieren sich am Vertragsmodell des Benutzers. Im Rahmen der existierenden Spezifikation nach Kapitel 4.4 ist keine Zwischenspeicherung dieser Daten sowohl in den Zugangsroutern, als auch im AAA.f-Server vorgesehen (vergl. Kapitel 5.7). Dies wird jedoch in dieser Arbeit vorgeschlagen. Prinzipiell ist die Kommunikation zwischen einem Zugangsroutern und dem dazugehö-

renden AAA.f-Server vom vertraglichen Zahlungsmodus des Benutzers abhängig. Bei einem Prepaid-Kunden sollte dieses Intervall möglichst klein gehalten werden; bei einem Kunden mit Festvertrag kann dieses Intervall entsprechend erweitert werden. Ein weiterer Faktor, welcher dieses Intervall beeinflusst, ist der resultierende Aufwand für den Datenspeicher. Im Falle eines Prepaid-Kunden sind verschiedene Strategien bezüglich der Kreditverwaltung denkbar. So könnten die Kredite im AAA.h-Server (Vorschlag nach [25]), im AAA.f-Server (dezentrale Haltung) und in AR (voll verteilte Haltung) gehalten werden. Generell kann das Konzept der Kreditverteilung jeweils Mischformen annehmen, welche dann ggf. zu weiteren Optimierungen führen, welche im weiteren Verlauf dieser Arbeit noch genauer betrachtet werden.

6.2.5 Sicherheitsaspekte

Die Kommunikation zwischen Zugangsroutern und den zugehörigen AAA-Servern und zwischen AAA-Servern, die sowohl verschiedenen als auch gleichen administrativen Netzen zugehören, kann als sicher angenommen werden, da die Beziehung zwischen diesen Komponenten als rein statisch bezeichnet werden kann. Hier wird vorgeschlagen, dass die jeweiligen Kommunikationspartner auf einen symmetrischen Schlüssel zurückgreifen. Das Verteilen und Verwalten dieser symmetrischen Schlüssel ist nicht Bestandteil dieser Arbeit.

Die Behandlung der Daten zwischen involvierten Betreibern im Roaming-Fall wird im Rahmen eines Roaming-Abkommens festgelegt. Auch hier wird davon ausgegangen, dass die vereinbarten Maßnahmen sowohl dem Sicherheitsbedürfnis des Benutzers, als auch dem Sicherheitsbedürfnis des Betreibers, Rechnung tragen. Auch dieser Aspekt wird im weiteren Verlauf dieser Arbeit nicht weiter verfolgt.

6.3 Vorschlag für ein Sitzungskonzept

Im Rahmen der AAA-Standardisierungsaktivitäten gibt es kein integriertes Sitzungskonzept, das den Anforderungen eines kommerziell betriebenen Mobiles Internets genügt. Im Rahmen dieser Arbeit soll nun ein solches Konzept vorgeschlagen werden. Ziel des Sitzungskonzeptes ist, die Voraussetzungen zu schaffen, um die verteilt entstehenden Daten effektiv zusammenführen zu können. Dies ist vor allem für Auditierung (falls integriert realisiert) und für die Aufbereitung der Nutzungsdaten in heterogenen verteilten Systemen, in dem sich die mobilen Teilnehmer uneingeschränkt bewegen und hierbei die existierenden Sitzungen aufrechterhalten können, erforderlich.

Das Konzept der Sitzung wird dazu verwendet, abrechnungsrelevante Vorgänge geeignet Benutzern zuordnen zu können. Diese Vorgänge können aus einzelnen Nachrichten bestehen, aus Transaktionen oder auch aus Sitzungen. Somit ist das Sitzungskonzept eine logische Konsequenz aus der in Kapitel 5.6 geforderten Benutzerorientierung, welche als technische Folgerung eine Neustrukturierung des Flusses der Nutzungsdatenberichte ermöglicht.

6.3.1 Beschreibung des Sitzungskonzeptes

Im Rahmen dieser Arbeit wird von einer Mobilität der Benutzer nach Kapitel 4.2.1 und 4.2.2 ausgegangen. Einem Benutzer ist also zeitweise ein bestimmtes Endsystem zugeordnet, welches im Internet durch eine IP-Adresse zuordenbar ist. Es soll an dieser Stelle darauf hingewiesen werden, dass im Rahmen dieser Arbeit nicht davon ausgegangen wird, dass ein Benutzer permanent ein Endgerät besitzt. Somit kann die Identifikation eines Benutzers nicht indirekt über das Endgerät (wie z.B. in 2G) erfolgen, sondern muss vom Netz auf Basis von AAA-Mechanismen, wie in Kapitel 4.3 beschrieben, erbracht werden. Ein Sitzungskonzept wird für folgende Zuordnungen vorgeschlagen:

- Authentisierung, Autorisierung und Aufbereitung der Nutzungsdaten mit der Dienstbereitstellung (Dienst Sitzung),
- Nutzungsdatenberichte, welche es ermöglichen, Sitzungen, die über mehrere IP- Zugangsnetze hinweg aktiv waren, geeignet zusammenzuführen. Dies kann im Extremfall auch zu zusammenhängenden Sitzungen führen, an denen mehrere Dienstanbieter bei der Ressourcenbereitstellung beteiligt waren,
- Verschiedene Dienst-Sitzungen, die logisch zusammen gehören.

Authentisierung, Autorisierung und das Konsumieren von Diensten sind zunächst drei getrennte Vorgänge. Um jedoch ein sinnvolles Konzept zur Nutzungsdatenerfassung und der nachgelagerten Weiterverarbeitung dieser Daten zu ermöglichen, welches als Basis für ein effektives Charging verwendet werden kann, ist ein Verbinden dieser Vorgänge erforderlich. Dies ist umso wichtiger, wenn die Funktionen von verschiedenen Instanzen in verschiedenen administrativen Domänen ausgeführt werden.

Der Konsum von Diensten muss geeignet auf Authentisierung und Autorisierung abgebildet werden, um eine Beziehung zwischen den Abrechnungsberichten (vergl. Kapitel 4.2.4.16 bis 4.2.4.18) und einem Benutzer, der für den Konsum der Daten letztendlich bezahlen muss, herzustellen. Natürlich ist hier eine gewisse Zuverlässigkeit sehr wichtig, da ein Missbrauch bzw. ein Fehler generell als sehr schwerwiegend zu kategorisieren ist. Weiter ist es zwingend erforderlich, dass jederzeit eine Rückverfolgung der Vorgänge im Netz möglich ist. Somit sollte eine Sitzung durch eine Sitzungs-ID eindeutig identifiziert werden können, und eine Auditierung, welche zu einem späteren Zeitpunkt erfolgt, unterstützen. Die Sitzungs-ID sollte einem Betreiber ermöglichen, die erfassten Nutzungsdaten geeignet und unter Berücksichtigung von Transaktionen nach Kapitel 4.2.4.6 zusammenzuführen, um letztendlich effektiv zu einer Rechnungserstellung zu gelangen.

In [112] wurde gezeigt, dass selbst in einem sehr einfachen Szenario, bei dem sich ein Benutzer über ein Modem in ein Netz einwählt, wenigstens zwei Berichte (Start/Stop) entstehen. Es ist somit zu erwarten, dass deutlich mehr als nur zwei Berichte pro Sitzung erzeugt werden, speziell wenn während einer Sitzung Zwischenberichte entstehen, oder wenn bestimmte Dienste innerhalb einer Sitzung nur zeitweilig konsumiert werden. Weiter ist davon auszugehen, dass die Sitzungs- bzw. Abrechnungsberichte in einem Mobilnet verteilt erzeugt werden.

Generell kann ein Dienst aus verschiedenen Einzeldiensten und somit aus Untersitzungen bestehen, welche zu einer Obersitzung gebündelt werden (vergl. Kapitel 4.2.4.4) und von verschiedenen Diensteanbietern erbracht werden. Dies können Diensteanbieter auf gleicher Netzebene (Föderierung erster Art und zweiter Art nach Kapitel 6.2.1.1) oder aber auf verschiedenen Netzebenen sein (Föderierung dritter Art nach Kapitel 6.2.1.1). Aus Sicht des Endverbrauchers handelt es sich jeweils um einen Dienst, welcher je nach Vertragsgestaltung bzw. Preismodell benutzerspezifisch ausgestaltet ist. In jedem Fall unterhält der Endverbraucher hier nur eine Beziehung mit einem einzelnen Anbieter (vergl. Kapitel 3.4). Ein geeignetes Sitzungskonzept kann hier eine Sitzungs-ID gemeinsam über alle involvierten Diensteanbieter hinweg vorsehen, jedoch ist auch eine individuelle Sitzungs-ID denkbar, die jeder einzelnen Untersitzung zugeordnet wird, denkbar. Jedoch ist hier dann ein entsprechender Mechanismus erforderlich, der in der Lage ist, die entsprechenden Zuordnungen durchzuführen.

Als Kernanforderung an ein Sitzungskonzept kann hier die weltweite Eindeutigkeit erwähnt werden. Diese Eindeutigkeit muss für jede beliebige Kombination zwischen Diensteanbieter und Benutzer gewährleistet sein.

6.3.1.1 Sitzungs-ID-Management

Die Sitzungs-ID wird von dem ersten AAA-Server, die eine Registrierungsauforderung bearbeitet, vergeben. Im Rahmen dieser Arbeit wird ein hierarchisches Zuordnungskonzept vorgeschlagen, das eine spätere Auditierung und ein Verknüpfen der Nutzungsinformationen zum Zweck der effektiven Datenverdichtung unterstützt.

Die Sitzungs-ID wird dann in allen einer Sitzung zuordenbaren erfassten Nutzungsdaten verwendet. Es wird vorgeschlagen, die Sitzungs-ID in einen Benutzerteil (Benutzer-ID) und einen Betreiberteil (Betreiber-ID) aufzuteilen.

6.3.1.2 Sitzungszuordnung

Eine Sitzung beginnt mit einer Dienstanforderung und endet mit einer expliziten Terminierung des Dienstes. Diese Terminierung kann prinzipiell sowohl auf Netzseite, als auch auf dem Endgerät angestoßen werden. Generell ist diese Prozedur vom speziellen Dienst und dem darüber liegenden Geschäftsmodell abhängig. Es ist davon auszugehen, dass zur Dienstleistung immer der Betreiber des Heimatnetzes und ggf. der Netzbetreiber des auswärtigen Netzes eingebunden sind. (Das Einbinden eines dritten Netzbetreibers, der die Verbindung zwischen den beiden Netzen darstellt, wird im Rahmen dieser Arbeit nicht weiter betrachtet). Erstreckt sich nun ein Mobilitätsszenario eines auswärtigen Benutzers über zwei verschiedene, vom Heimatnetzbetreiber unterschiedliche administrative Bereiche, so wird bei diesem Netzübergang eine Sitzung (im „alten“ Netz) terminiert und eine neue Sitzung (im „neuen“ Netz) gestartet.

Das Verknüpfen einer Sitzung mit den dazugehörigen Nachrichten sollte geeigneterweise zeitnah erfolgen, um vor allem zeitkritische Geschäftsmodelle wie z.B. Prepaid unterstützen zu können.

6.3.1.3 Flache vs. Hierarchische Adressierungskonzepte

Generell kann ein eindeutiges Sitzungsschema flach oder hierarchisch organisiert werden. Beide Konzepte haben ihre Vor- und Nachteile und sollen nachfolgend kurz diskutiert werden.

Als Beispiel für ein hierarchisches Schema ist die historische Adressierung im Telefonnetz zu erwähnen. Hier gibt es in oberster Ebene eine Identifizierung der Landeskennung. Die nächste Hierarchieebene ist dann nach den verschiedenen nachgelagerten Vermittlungsinstanzen aufgeteilt, welche historisch Regionen, Städte, bis hin zu ggf. Wohnvierteln abbilden.

Dieses System migriert nun langsam in ein eher flaches Adressierungsschema, welches nun dienstorientiert den Wert der Inhalte implizit signalisiert. Ein Beispiel hierfür sind die 0180/0190-Dienstnummern im öffentlichen Telefonnetz in Deutschland, welche nicht mehr hierarchisch aufgebaut sind

Der Vorteil eines hierarchischen Konzeptes liegt in der effizienten Unterstützung von Verkehrslenkung bzw. Weiterverarbeitung der Information. Mit Zunahme der Leistungsfähigkeit der IT-Infrastruktur reduziert sich die Bedeutung dieses Vorteils und man kann sozusagen fließend in ein flaches Adressierungskonzept übergehen, welches mehr Flexibilität bietet.

6.3.1.4 Sitzungsgraph

In Abbildung 49 ist die Struktur des vorgeschlagenen Sitzungsgraphen dargestellt. Die Wurzel eines Sitzungsgraphen ist eine Sitzung, welche einen Prozess beschreibt, der zu einer finalen Preisberechnung führt. Es handelt sich hierbei um einen unidirektionalen Graphen der,

ausgehend von einer Obersitzung, in viele Untersitzungen eingeteilt werden kann. Ausgehend von einer Untersitzung kann nicht auf eine Obersitzung geschlossen werden, was somit nur einer begrenzten Anzahl von Instanzen einen Gesamtblick auf eine solche Sitzung ermöglicht und somit Vertraulichkeitsaspekte berücksichtigt.

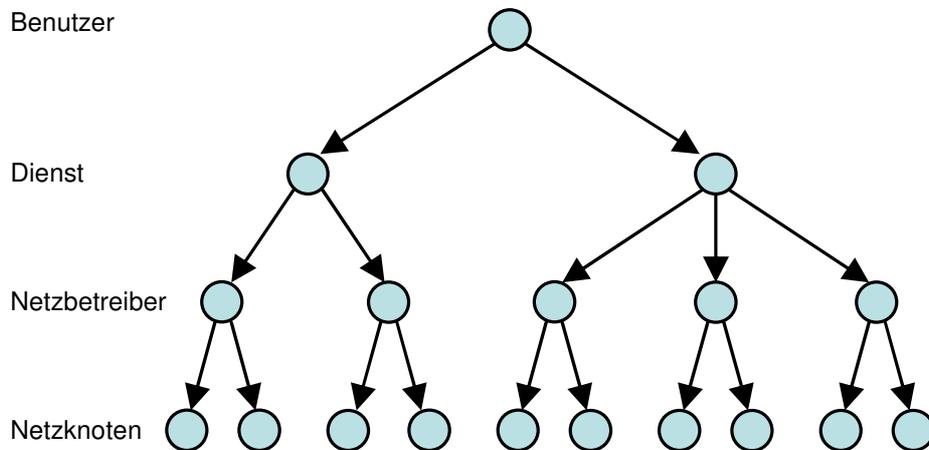


Abbildung 49: Sitzungsgraph

6.3.1.5 Sicherheitsaspekte des Sitzungskonzeptes

Generell kann davon ausgegangen werden, dass zahlreiche Daten, die ein Netzbetreiber über seine Vertragsnehmer erfasst und zu verschiedenen Zwecken weiterverarbeitet, für einen Betreiber sehr wertvoll sind, da sie für weitere benutzerbezogene Charakterisierungen und zu Marktforschungszwecken für den eigenen kommerziellen Nutzen verwendet werden könnten. Diese Informationen sollen jedoch einem potentiellen Wettbewerber vorenthalten werden. Dies sind jegliche Daten die u.a. Rückschlüsse auf die Identität, die Bankverbindung des Kunden, das Bewegungsverhalten und das Kaufverhalten eines Kunden zulassen.

Somit soll ein Betreiber Nutzungsdaten so erfassen, dass die Rechnungserstellung gemäß Vertrag vereinfacht wird. Personenbezogene Daten wie Name des Benutzers usw., aber auch Daten aus denen Informationen bzgl. der Netzauslastung eines Betreibers usw. gewonnen werden können, sollten so weit als möglich anonymisiert werden.

In diesem Zusammenhang ist davon auszugehen, dass die Parameter, die Bestandteil des Vertrages zwischen Kunde und Netzbetreiber sind, nicht zwingend zwischen verschiedenen Netzbetreibern in gleicher Form ausgetauscht werden. Hier können sich z.B. zwei Betreiber zu einem kostenlosen Zusammenschalten ihrer Netze verständigen, ein Betreiber jedoch die Belegung von Ressourcen im Netz des Roaming-Partners generell in Rechnung stellt. Somit hat die Nutzungsdatenerfassung und nachgelagerte Aufbereitung zwei verschiedene Zwecke:

- Es müssen die Daten gemäß des Vertrages zwischen Endkunde und Netzbetreiber erfasst werden.
- Die generierten Datensätze müssen bei Bedarf zur Abrechnung zwischen zwei Betreibern beitragen.

6.3.2 Definition der Sitzungsberichte

Das vorgeschlagene Datenmodell ist in Tabelle 9 dargestellt und basiert auf dem in Kapitel 4.5 vorgestellten IPDR-Modell, welches entsprechend den Mobilitätsanforderungen erweitert wurde.

Tabelle 9: Datenmodell der Sitzungsberichte

Allgemeine Attribute	
Sitzungs-ID	Weltweit eindeutig identifizierbare Kennzeichnung
Zielparameter	IP-Adresse / Transportadresse
Sendeparameter	IP-Adresse / Transportadresse
Nutzungsattribute	
Start Zeitpunkt	Zeitpunkt, zu dem die Sitzung gestartet wurde
Stop Zeitpunkt	Zeitpunkt, zu dem die Sitzung beendet wurde bzw. letztmalig als aktiv berichtet wurde
Dienstklasse	Bezeichnung des Dienstes
Class-ID	Bezeichnung welcher Benutzergruppe der Benutzer zuzuordnen ist
Menge	Nutzungsmenge
Authentisierungs- Transaktions-Referenz	Eine Referenz über den vollzogenen Autorisierungsprozess vor Dienstfreigabe

Die Sitzungs-ID besteht aus einem Benutzerteil und einem Betreiberanteil (vergleiche Kapitel 6.3.1.1 und Abbildung 50).

Startzeit und Stopzeit werden aus der IETF RTFM-Referenzimplementierung übernommen ([16], [21]) und haben eine Länge von je 8 byte.

Dienstklasse ist ein Parameter, welcher die Dienstbündel adressiert, die von einem Betreiber angeboten werden. In Kapitel 6.3.5 wird dieser Parameter detaillierter spezifiziert.

Zielparameter und Sendeparameter beschreiben das Tupel IPv6-Adresse/Transportadresse und haben somit je eine Länge von 18 byte. Beide Parameter werden aus der IETF RTFM-Referenzimplementierung übernommen ([16], [21])

Class-ID transferiert die Dienstklassen in konkrete Meterkonfigurationen. Die genaue Funktionsweise wird in Kapitel 6.4 detailliert vorgestellt.

Die Nutzungsdatenmenge repräsentiert die erfassten Nutzungsdaten pro Dienstklasse. Die genaue Spezifikation erfolgt in Kapitel 6.3.6.

6.3.3 Sitzungsverwaltung

6.3.3.1 Anforderungen

Die Anforderungen an eine Sitzungsverwaltung können in zwei verschiedene Bereiche eingeteilt werden. Zunächst ist die einfache und zweckgerichtete Weiterverarbeitbarkeit der Nutzungsberichte ein wesentliches Kriterium. Weiter sollen die gesetzlichen Vorgaben aus dem Bereich Datenschutz nicht verletzt werden. In diesem Zusammenhang soll ein Betreiber, welcher temporär einem auswärtigen Benutzer Ressourcen freigibt, nicht in der Lage sein, personenbezogene Daten (z.B. das detaillierte Bewegungs- oder Konsumverhalten eines personifizierten Teilnehmers) über verschiedene Sitzungen hinweg nachvollziehen können.

6.3.3.2 Annahmen und Ausgangspunkt

Bei der Einführung einer Sitzungsverwaltung wird davon ausgegangen, dass jeder Benutzer über Dienste eines Netzbetreibers Zugang zum Internet erhält. In der Regel unterhält dieser Netzbetreiber Roaming-Abkommen mit anderen Netzbetreibern. Es soll folgende Terminolo-

gie eingeführt werden:

Der Dienst, den ein Benutzer erhält, ist im Rahmen eines Vertrages zwischen Netzbetreiber und Benutzer definiert. Hier wird analog zum aktuellen 2G-Netz angenommen, dass die Dienstgüte, die im Heimatnetz gewährt wird, ebenso spezifiziert ist, wie die Dienstgüte, die in den jeweiligen Roaming-Bereichen gewährt wird. Diese Dienstgüte basiert ausschließlich auf den individuellen Abkommen der involvierten Betreiber.

Es soll an dieser Stelle nochmals ausdrücklich erwähnt werden, dass sich die Rolle eines auswärtigen Servers nicht nur auf einem Partner auf gleicher Netzebene beschränkt, sondern auch auf z.B. einen Inhalte-Anbieter ausgeweitet werden kann (vergl. Förderierung dritter Art aus Kapitel 6.2.1.1). Somit können mehr als zwei auswärtige Netzbetreiber gleichzeitig Ressourcen im Rahmen einer (Ober)Sitzung freigeben.

6.3.3.3 Spezifikation

Für die Einführung eines hierarchischen Sitzungskonzeptes sollen zunächst einige globale Überlegungen angestellt werden.

Bei IPv6 können (vergl. Kapitel 2.1.1.1, Abbildung 2) maximal 80 bit für die eindeutige Identifizierung eines Endsystems verwendet werden. Da jedes Endsystem eine eigene administrative Domäne bilden kann, sind folglich 80 bit zur Identifizierung der administrativen Domänen vorzusehen.

Zur Nachverfolgung des detaillierten Benutzerverhaltens (Auditierung) kann weiterhin vorgesehen werden, dass grundsätzlich die Möglichkeit besteht, den genauen Aufenthaltsort eines Benutzers nachvollziehen zu können. Bei Annahme einer Architektur nach Kapitel 4.8 und einem durchschnittlichen Zellradius von 100 Metern und weiterer Annahme, dass ein Netzbetreiber maximal das gesamte Festland der Erde abdeckt, ergibt sich folgende Kalkulation (Tabelle 10):

Tabelle 10: Kernparameter Beispielrechnung

Mittlerer Zellradius einer B3G-Zelle in km	0,1
Mittlere Zellgröße in km ²	0,03
Erdoberfläche (ohne Wasser) in km ²	150 000 000

Somit hat ein Netzbetreiber maximal 5 Milliarden Zugangsroutern in Betrieb. Diese Grenz-betrachtung ist rein theoretischer Art, führt jedoch zu einem Bedarf von max. 32 Bit, welche nach Abbildung 50 vorgeschlagen werden.

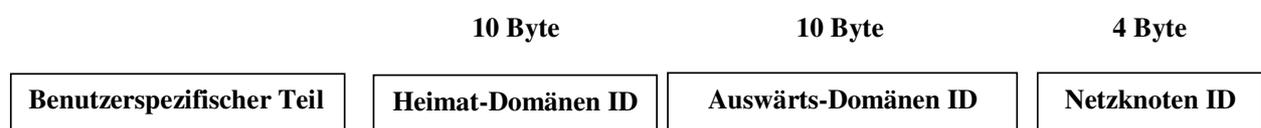


Abbildung 50: Aufbau der hierarchischen Sitzungs-ID

Die *Heimat-Domänen-ID*, die *Auswärts-Domänen-ID* und die *Netzknoten-ID* könnten jeweils in sich hierarchisch aufgebaut werden. Es ist davon auszugehen, dass durch Konsolidierungen im Telekommunikationsmarkt solch eine Adressierung mangels Flexibilität schnell in ein flaches Adressierungskonzept übergehen wird.

Eine Sitzungsverwaltung im Kontext dieser Arbeit erfordert die Einbindung von Zugangsroutern, und AAA-Servern in den beiden involvierten administrativen Bereichen. Die hierar-

chische Sitzungs-ID wird in Anlehnung an Abbildung 50 von den in Abbildung 51 dargestellten Komponenten vergeben.

Die detaillierte Struktur der *Netzknoten-ID* kann von jedem Betreiber selbst definiert werden und ist für die weiteren Betrachtungen nicht relevant. Hier sind vor Weitergabe an einen anderen Betreiber ggf. Anonymisierungsstrategien denkbar, welche zwar eine nachträgliche Auditierung ermöglichen, jedoch dem Betreiber des Heimatnetzes bestimmte Informationen vorenthalten und die Vertraulichkeit von personenbezogenen Daten sicherstellen. Im nächsten Abschnitt werden weitere Aspekte bezüglich der Vertraulichkeit vorgestellt.

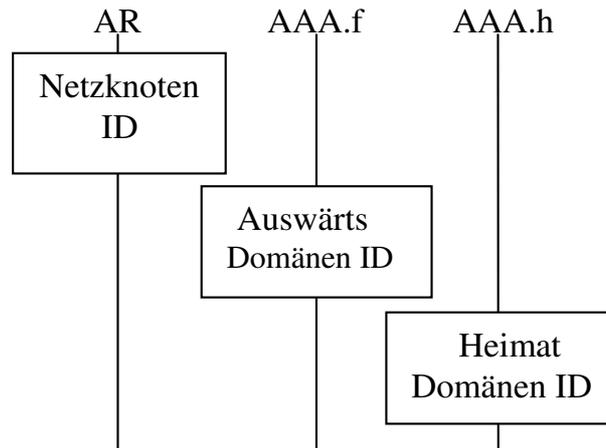


Abbildung 51: Zuweisung der Identifizier zur hierarchischen Sitzungs-ID

Um die Sitzungs-ID mit dem benutzerspezifischen Teil zu vervollständigen, wird ein weiteres Feld vorgeschlagen, welches die eindeutige Benutzerkennung festlegt. Die detaillierte Spezifikation dieses Feldes wird in dieser Arbeit nicht weiter verfolgt. Sollte ein Betreiber seine Benutzer mit einer laufenden Nummer identifizieren, sind hierfür weitere 10 Byte erforderlich.

6.3.4 Vertraulichkeit

Beim Abrechnen von Diensten ist die Vertraulichkeit der erfassten personenbezogenen Daten sehr wichtig. Konkret unterliegt das Erfassen und Weiterverarbeiten von personenbezogenen Informationen strengen gesetzlichen Bestimmungen.

Generell bietet sich aus Gründen der Vertraulichkeit an, jegliche personenbezogenen Daten zu anonymisieren. Grundsätzlich ist bei der Erfassung und Weiterverarbeitung dieser Daten aus Sicht der Vertraulichkeit zwischen folgenden zwei Fällen zu unterscheiden:

- Das Erfassen und Weiterverarbeiten von Daten, ohne dass der Benutzer davon in Kenntnis gesetzt wird
- Das Erfassen und Weiterverarbeiten von Daten unter voller Transparenz und Einwilligung des Benutzers.

Im Rahmen dieser Arbeit wird vorgeschlagen, von einem dienstspezifischen Erfassen von Nutzungsdaten zu einem benutzerspezifischen Erfassen dieser Daten zu gelangen, da dies vor allem ökonomische Vorteile bietet. Bei einem benutzerspezifischen Erfassen der Nutzungsdaten kann ausgeschlossen werden, dass nicht abrechnungsrelevante Daten eines Benutzers ohne dessen Kenntnis und Einwilligung erfasst werden. Dies schließt aus, dass Daten ohne Kenntnis des Benutzers weiterverarbeitet werden, was an folgendem Beispiel verdeutlicht werden soll:

Ein Benutzer mit einem Flat-Rate-Vertrag konsumiert verschiedene Dienste im Netz. Bei vollständiger Umsetzung der in dieser Arbeit gemachten Vorschläge ist ausgeschlossen, dass für diesen Benutzer überhaupt Nutzungsdaten erzeugt werden, da bei einer Flat-Rate keinerlei Nutzungsdaten abrechnungsrelevant sind. Somit kann für diesen Fall das Thema Vertraulichkeit als implizit gelöst betrachtet werden. Da ein Benutzer generell ein vertragliches Verhältnis mit seinem Heimbetreiber eingeht, welcher die entsprechenden Gebührensätze für die jeweiligen Leistungen beinhaltet, kann bei Vertragsabschluss die Einwilligung des Benutzers zur Erfassung dieser Daten mit eingefordert werden.

6.3.5 Benutzerklassen-Management

6.3.5.1 Anforderungen

Für das Management der Benutzerklassen ergeben sich folgende übergeordnete Überlegungen:

Ein Betreiber eines Dienstes bietet seinen Kunden die Auswahl aus einer festen Anzahl vordefinierter Dienstbündel in Anlehnung an Kapitel 5.5 an. Die Anzahl dieser verschiedenen Verträge, in denen dem Endbenutzer jeweils ein Produktbündel angeboten wird, ist generell überschaubar. Aktuell sind pro Betreiber ca. 20 solcher Bündel im Angebot, welche teilweise als Prepaid und teilweise als Festvertrag implementiert sind. Die Festlegung der theoretischen Anzahl der verschiedenen Vertragsmodelle, zwischen denen ein Kunde wählen kann, ist Aufgabe des Marketings und nicht Bestandteil dieser Arbeit. Es ist jedoch zu beobachten, dass hier zeitlich befristete Verträge mit Sonderkonditionen bzgl. Preisbildung und Dienstbündel verstärkt auf dem Markt erscheinen, d.h. mit einer zunehmenden Dynamik am Markt zu rechnen ist.

6.3.5.2 Beschreibung der Dienstklassenadressierung

Die Zahl der Preismodelle ist aus Sicht der Netzbetreiber übersichtlich, also so niedrig wie möglich zu gestalten, jedoch groß genug, um sich von einem Wettbewerber zu differenzieren und um ggf. Großkunden gesondert behandeln zu können. Im weiteren Verlauf dieser Arbeit wird hier von einem fiktiven Wert von 20 ausgegangen, was für diese Arbeit zu einem Nachrichtenformat von 1 byte führt (Abbildung 52). Wichtig in diesem Zusammenhang ist eine Erweiterbarkeit.

Bit 1-5 ermöglichen einem Netzbetreiber Preismodelle zu adressieren. Bit 6 und Bit 7 (als R-Bit gekennzeichnet) sind für zukünftige Verwendung bzw. Erweiterungen reserviert. Bit 8, das PP-Bit wird dazu verwendet, um schnell zwischen einem Prepaid-Kunden und einem Kunden mit Festvertrag unterscheiden zu können.

PP	R	R	5	4	3	2	1
----	---	---	---	---	---	---	---

Abbildung 52: Dienstadressierung

6.3.6 Datensatz der Nutzungsdaten

Die Parameter, welche zur Erfassung der Nutzungsdaten festgelegt werden, folgen in Anlehnung an den IETF RTFM-Vorschlag aus Kapitel 4.4.5, welcher folgendermaßen erweitert wurde:

- Erfassung von IPv6-Paketen
- Einführung eines Feldes für den DSCP nach Kapitel 2.1.2.3

- Einführung eines Feldes für Mobile IP-Erweiterungen. Konkret bedeutet dies die Einführung eines weiteren Feldes für die Heimatadresse (HA)

In Anlehnung an Abbildung 35 ergibt sich folgender Datensatz (Abbildung 53). Dies führt zu einer Datenmenge von 91 byte. Da in der Regel weitere Bytes für Kontrollfunktionen verwendet werden, wird für die Skalierungsbetrachtungen in Kapitel 7 überschlägig von einem Nutzungsdatenbericht von 100 byte pro Dienstklasse ausgegangen.

HA	IP SRC/ CoA	IP DST/ CoA	SRC Port	DST Port	DSCP
128 bit	128 bit	128 bit	8 bit	8 bit	8 bit
To Octets	From Octets	To PDU	From PDU	F_Time	L_Time
64 bit	64 bit	64 bit	64 bit	32 bit	32 bit

Abbildung 53: Datensatz eines Nutzungsdatenberichts

6.4 Zugangsrouterarchitektur

Die vorgeschlagene Architektur des Zugangsrouters ist in Abbildung 54 dargestellt und kann als Erweiterung der Moby Dick-Zugangsrouterarchitektur aus Kapitel 4.8 betrachtet werden. Kernbestandteil der Erweiterungen ist die Integration eines verteilten Konzeptes zur Nutzungsdatenerfassung.

Zentrales Element der Architektur ist der AAA-Manager, welcher intern Informationen mit den Komponenten Meter-Manager, Mobility-Manager und QoS-Manager austauscht. Extern kommuniziert der AAA-Manager mit dem AAA-Server und dem Endsystem. Die eingesetzten Protokolle mit den externen Komponenten sind PANA mit dem Endsystem (vergl. Kapitel 4.4.3) und Diameter mit dem AAA-Server (vergl. Kapitel 4.4.1). Für die Kommunikation mit den internen Komponenten wird kein explizites Protokoll vorgeschlagen. Weiter kommuniziert der AAA-Manager mit zwei Datenbanken, dem Benutzer Repository und dem Accounting Repository. Diese Kommunikation erfolgt mittels einer JDBC-Schnittstelle. In den folgenden Abschnitten soll auf die einzelnen Elemente näher eingegangen werden.

6.4.1 Meter und Meter-Manager

Meter und Meter-Manager basieren auf der IETF RTFM-Architektur, wie in Kapitel 4.4.5 vorgestellt, mit spezifischen Erweiterungen und einer Integration mit der in Kapitel 4.4 vorgestellten AAA-Architektur und berücksichtigen die Vorschläge aus [47]. Die Erweiterungen sind in Abbildung 55 dargestellt.

Kernkonzept ist eine kontrollierte Ansteuerung des Meters, welche für eine bestimmte IP-Adresse den Prozess der Nutzdatenerfassung startet bzw. beendet. Dieses Starten und Beenden wurde so gestaltet, dass die IP-Quelladresse mit weiterem Parameter, der Nutzerklasse, versehen wurde, welche in der Class-ID aus Tabelle 9 signalisiert werden. Konkret sind in der NeTraMet-Implementierung vordefinierte Erfassungsregeln konfiguriert, welche mittels der Class-ID adressiert werden.

Das Starten des Meters fügt lediglich die CoA des neuen Benutzers dynamisch einer dieser Erfassungsregeln hinzu. Als Protokoll für diese Kommunikation wird SNMP vorgeschlagen.

Der Kerngedanke hierbei ist, dass bei der Registrierung der darin beinhaltete Parameter Nutzerklasse angibt, welche Daten pro IP-Quelladresse überhaupt erfasst werden müssen. Dies erfolgt über eine dynamische Zuordnung der Erfassungsregeln der RTFM-Architektur zu IP-Adressen. Als weitere Modifizierung wurde der Meter-Manager aus Kapitel 4.4.5 mit einer neuen Funktionalität versehen. Dies ist ein Übertragen der aus dem Meter gewonnenen Rohdaten in eine Datenbank, welche eine erste Weiterverarbeitung im Zugangsroutern ermöglicht.

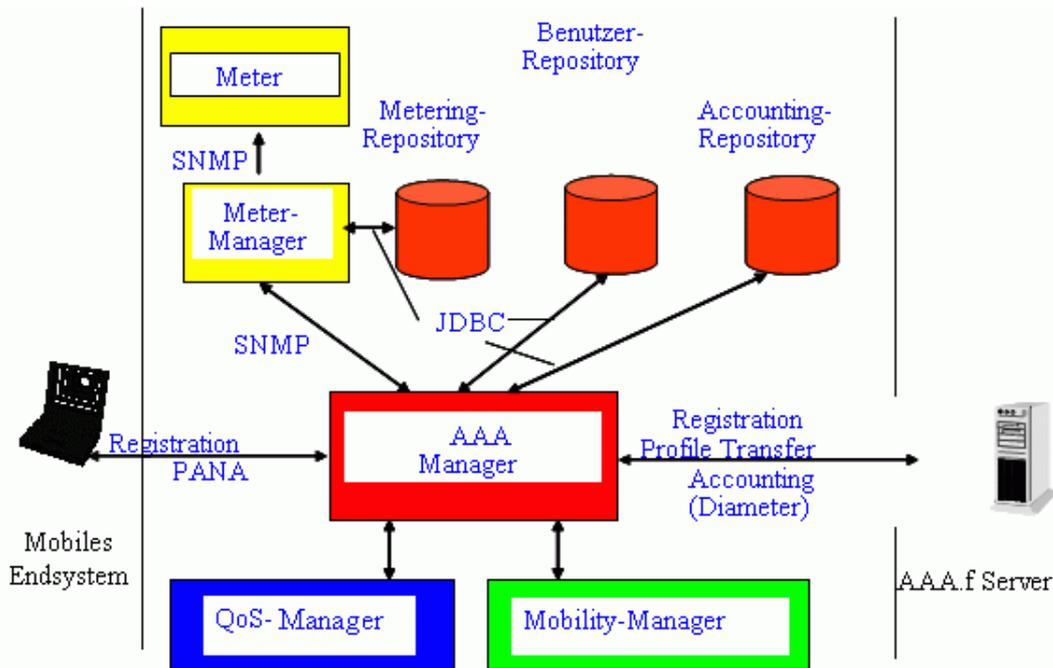


Abbildung 54: Architektur des Zugangsrouters (ohne Datenpfad)

6.4.2 Skalierbarkeit von Meter und Meter-Manager

Aktuelle drahtlose Netzchnittstellen haben eine maximale Kapazität von 50 Mbit/s (802.11 10 Mbit/s, UMTS 2 Mbit/s und 802.16 50 Mbit/s). Die IETF RTFM-Referenzimplementierung wurde als OC12-Monitor problemlos eingesetzt. Somit ist kein Engpass zu erwarten. Für nähere Informationen bzgl. der Leistungsfähigkeit wird auf [14], [16] und [15] verwiesen.

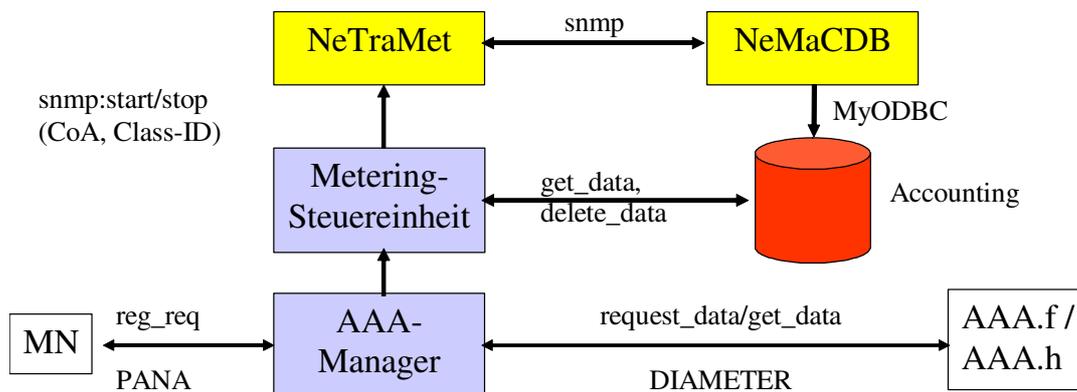


Abbildung 55: Erweiterungen der IETF RTFM-Architektur (Accounting)

6.4.3 AAA-Manager

Der AAA-Manager ist ein IETF AAA Client nach Kapitel 4.4.1, der zusätzlich die Funktionalität eines PANA Agent (PAA) nach Kapitel 4.4.3 erfüllt und mit der Metering-Steuereinheit kommuniziert.

Neben dem Anfordern der Registrierungsinformationen beim Betreiberwechsel bzw. erstmaligen Registrieren bei einem Betreiber hat der AAA-Manager noch die Aufgabe, die konsolidierten Nutzungsdaten benutzerspezifisch an den AAA.f-Server weiterzuleiten.

Während der Registrierung werden, wie in Abbildung 42 dargestellt, die Registrierungsanforderungen vom AAA-Manager empfangen, welcher diese zum entsprechenden AAA-Server weiterleitet. Nach erfolgreicher Authentisierung und Autorisierung gemäß Kapitel 4.4.1 empfängt der AAA-Manager die Antwort, welche entweder positiv oder negativ ausfallen kann. Bei einer negativen Antwort wird kein Benutzerprofil zum Zugangsrouten gesendet. Die Registrierung ist somit implizit abgelehnt. Bei positiver Registrierung erhält der AAA-Manager ein Benutzerprofil nach Abbildung 43, in dem alle zur Dienstbereitstellung erforderlichen Daten enthalten sind. Diese Daten werden dann über den Profile Client im Benutzer-Repository (siehe Abbildung 54) abgelegt. Dieser Prozess ist schematisch in Abbildung 56 dargestellt.

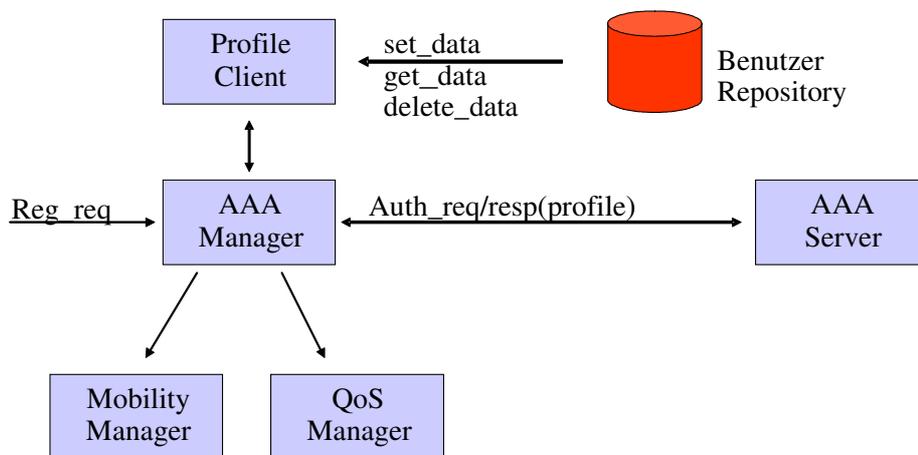


Abbildung 56: Registrierungsprozess beim Zugangsrouten

6.4.4 Skalierbarkeit der Registrierung

Die Anzahl der Registrierungen, welche bei einem Zugangsrouten eintreffen, hängt u.a. von der Anzahl der sich im Einzugsbereich eines Zugangsrouters befindlichen Benutzer ab. Es soll zunächst folgende Beispielrechnung gemacht werden:

Sei ein Zellradius einer Netzzelle, welche nach Moby Dick dem Einzugsbereich eines Zugangsrouters entspricht, mit 300 Meter angenommen. In einem solchen Gebiet (z.B. Fußballstadion) sollen sich 60000 Personen aufhalten. Geht man weiter davon aus, dass sich 4 Betreiber (UMTS-Markt in Deutschland) diese Kunden „gerecht“ aufteilen, dann kann man 15000 Benutzer in dieser Netzzelle annehmen. Geht man weiter davon aus, dass sich ein solches Stadion binnen 30 Minuten füllt, dann registrieren sich pro Minute 500 bzw. pro Sekunde ca. 9 Benutzer. Dies hat ein zu verarbeitendes Datenvolumen zur Folge, welches in aktuell implementierten Systemen nicht als technische Herausforderung betrachtet wird. Systeme wie z.B. elektronische Handelsplattformen können mehrere 100000 Transaktionen in einer Sekunde zuverlässig abwickeln, was bei den oben angenommenen Zahlen einem fiktiven Zuschauer-volumen von 800 Millionen Personen im Fußballstadion entspräche.

6.4.5 QoS-Manager

Der QoS-Manager ist für die Ausführung von Regeln zuständig. Diese Regeln organisieren die Freigabe der Netzressourcen in der Dienstgüte gemäß dem Vertrag des Benutzers mit dem Betreiber. Sollte sich das Budget eines Benutzers mit Prepaid-Vertrag einer Sitzung aufbrauchen, wird dies der AAA-Manager dem QoS-Manager berichten, welcher dann den entsprech-

enden Zugang sperrt. Somit kann der QoS-Manager als Instanz bezeichnet werden, welche sicherstellt, dass die konkrete, benutzerspezifische Autorisierung entsprechend des Vertrags mit dem Benutzer ausgeführt wird.

6.4.6 Skalierbarkeit des QoS-Managers

Die in Kapitel 6.4.4 gemachten Überlegungen treffen qualitativ auch auf den QoS-Manager zu, so dass auch hier keinerlei zusätzlichen Vorkehrungen getroffen werden müssen.

6.4.7 Mobility-Manager

Der Mobility-Manager basiert auf den in Moby Dick entwickelten FAST-Ansatz nach [81], der jedoch mit spezifischen Erweiterungen versehen wurde. Dies sind der Kontext-Transfer der benutzerspezifischen Informationen wie Benutzerklasse und Sitzungs-ID und ggf. Restguthaben (vergleiche hierzu auch Kapitel 7.5.2 und Kapitel 7.6.2.2) beinhaltet. Beide Parameter werden in die Signalisierung nach Abbildung 44 integriert, um den Meter im neuen Zugangsroutern zeitnah benutzerspezifisch konfigurieren zu können. Bei Übertragung der Sitzungs-ID wird gemäß Abbildung 49 die Netzknotten-ID entsprechend aktualisiert.

6.4.8 Skalierbarkeit des Mobility-Managers

Nach Abbildung 44 wird ein Zugangsroutern pro Benutzer und Netzwechsel mit 7 Nachrichten beansprucht, welche möglichst in Echtzeit verarbeitet werden müssen. Dies ist auf Basis der in Kapitel 6.4.4 gemachten Überschlagsrechnung ebenfalls problemlos verarbeitbar.

6.5 Integration des Gesamtsystems

6.5.1 Nutzungsdatenerfassung und nachgelagerte Weiterverarbeitung

Generell kann die Übertragung der Nutzungsdaten entweder proaktiv, oder aber durch explizite Anforderungen initiiert werden. Wie in Abbildung 57 gezeigt, generiert der Meter (NeTraMet) kontinuierlich Daten, welche vom Meter-Manager (NeMaCDB) gelesen werden. Dieser legt sie im Accounting-Repository (ACC DB) gemäß Abbildung 54 ab. Der AAA-Manager holt sich diese Daten aus dem Repository ab. Das Abfrageintervall richtet sich nach der Benutzerklasse des Benutzers und wird in Kapitel 7.3.7 ausführlicher diskutiert.

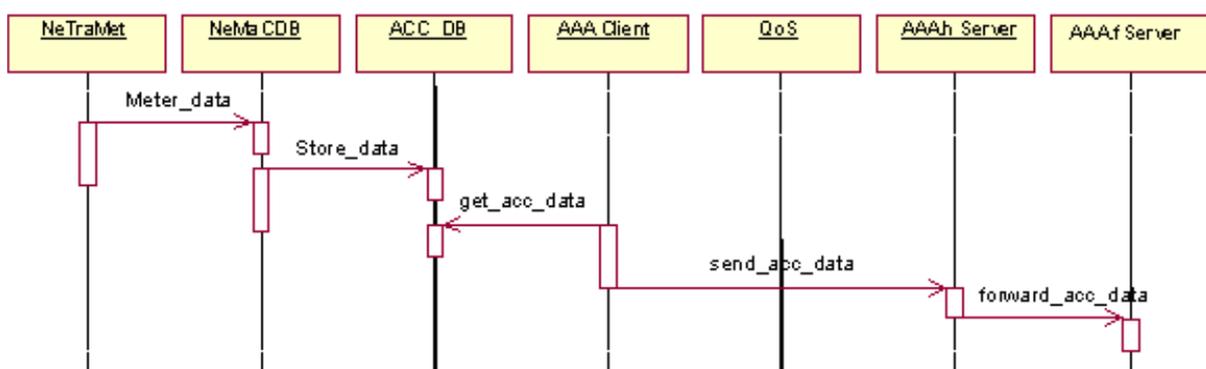


Abbildung 57: Datenfluss im Accountingprozess

Auch hier soll wieder von dem Szenario aus Kapitel 6.4.4 ausgegangen werden, in dem 15000 Personen in einer Zelle angenommen wurden. Natürlich ist es nicht realistisch, dass alle Personen gleichzeitig aktiv sind, da dies die physikalischen Möglichkeiten einer Zugangszelle sprengen würde. Folgt man dem „Always-On“ Paradigma, so muss das Netz jedoch in der Lage sein, jedem Zugriffswunsch auf die Ressourcen je nach aktueller Ressourcenbelegung und vertraglicher Ausgestaltung der Beziehung mit dem Netzbetreiber Folge leisten zu kön-

nen. Hier wäre periodisch die konsumierte Datenmenge an den AAA-Server zu berichten. Im ungünstigsten Fall hat jeder der 15000 Benutzer einen Teil der zur Verfügung stehenden Ressourcen konsumiert. Die absolute Größe dieses Teils ist hierbei unerheblich, da die zu übertragene Datenmenge hiervon unabhängig ist. Folglich müssen 15000 Datensätze in einem vordefinierten Zeitintervall übertragen werden. Ein solcher Datensatz basierend auf dem Vorschlag aus Tabelle 11 hat eine Größe von 212 byte, bei Verwendung einer Dienstklasse. Pro zusätzlicher Dienstklasse erhöht sich dieser Wert um 100 byte.

Für die weitere Berechnung wird ein Datenvolumen von 350 byte angenommen (2 angenommene Dienstklassen; 38 byte Reserve), welches dann durch überlagerte Sicherungsmechanismen (IPsec und andere Fehlerkorrekturmechanismen) 500 byte nicht überschreiten sollte.

Tabelle 11: Dimensionierung der Nutzungsdatenberichte

Allgemeine Attribute	
Sitzungs-ID (vergl. Abbildung 50)	34 byte
Zielparameter (vergl. Tabelle 9)	18 byte
Sendeparameter (vergl. Tabelle 9)	18 byte
Nutzungsattribute	
Start Zeit (vergl. Kapitel 6.3.5.2)	8 byte
Stop Zeit (vergl. Kapitel 6.3.5.2)	8 byte
Class-ID (vergl. Kapitel 6.4.1)	1 byte
Dienstklasse (vergl. Abbildung 52)	1 byte
Nutzungsmenge (vergl. Kapitel 6.3.6)	100 byte pro Dienstklasse
Authentisierungs-Transaktions-Referenz	24 byte

6.5.2 Skalierbarkeitsbetrachtungen des Gesamtsystems

Das Übertragen von 15000 Datensätze nach Tabelle 11 in einem Zeitintervall von 1 Minute (hier wird angenommen, dass sich ausschließlich Benutzer mit Prepaid-Verträgen im Einzugsbereich einer Zelle befinden) erfordert eine Übertragungskapazität von 60 Mbit/Minute, welche im Extremfall innerhalb einer Sekunde angerufen werden müssen. Somit kann diese Schnittstelle von einer gebräuchlichen 1 Gbit/s Ethernet-Netzschnittstelle inklusive des Datenvolumens der Zelle problemlos bewältigt werden. Geht man von Bandbreiten im Kernnetz von 10 Gbit/s aus, so könnten hiermit 166 Zellen mit je 15000 Benutzern, also 2,5 Mio. Personen von einem AAA-Server versorgt werden, was für ein Land wie der Bundesrepublik Deutschland ca. 25 AAA-Server zur Folge hätte.

Setzt man an dieser Stelle die Berechnung aus Kapitel 6.4.2 fort und nimmt 9 Registrierungen pro Zelle und Sekunde an, so ergeben sich bei 166 Zellen ca. 1500 Registrierungen pro Sekunde am AAA-Server. Dies ist von Systemen, welche im Finanzbereich eingesetzt werden problemlos zu bewältigen, so dass hier keine konzeptuellen Beschränkungen zu erwarten sind.

6.6 Zusammenfassung

In diesem Kapitel wurden die drei Kernforderungen dieser Arbeit - Zugangsrouterarchitektur, Signalisierungskonzept und Sitzungsmanagement - detailliert vorgestellt. Jegliche Transitionsstrategien und Migrationskonzepte von der existierenden Infrastruktur zu der vorgeschlagenen Lösung werden nicht betrachtet. Es ist jedoch nicht zu erwarten, dass die vorgeschlagenen Mechanismen nicht in die aktuelle Internet Architektur integrierbar wären, da sie weitestgehend in existierende Protokolle und Standards der IETF eingebettet sind und somit parallel zum existierenden Internet betrieben werden könnten.

Die Zugangsroutearchitektur wurde spezifiziert und bezüglich der Skalierbarkeit untersucht. Die detaillierte Signalspezifikation ist im Anhang dargestellt. Dadurch, dass die hier vorgeschlagene Lösung konzeptionell eine voll verteilte Architektur vorsieht, in der selbst zentrale Instanzen wie der *AAA.h*-Server relativ viele Aufgaben an die Zugangsrouters delegieren, skaliert die vorgeschlagene Lösung sehr gut. Aktuell verfügbare PCs können die Aufgaben eines Zugangsrouters sehr leicht übernehmen. Der Grund hierfür liegt darin, dass ein Zugangsrouters wegen der geringen Zellgröße relativ wenig Benutzer bedienen muss.

Architektonisch ist diese Lösung fundamental von der 3GPP/3GPP2-Lösung aus Kapitel 2 verschieden, welche tendenziell mehr Intelligenz in die *RNCs* verlagert, und die *Node B's* relativ einfach hält. Die hier eingesetzte Architektur verfolgt den Ansatz von einer komplexen Funktionalität der *Node B's*, welche in Moby Dick Zugangsrouters genannt werden, jedoch in Konsequenz den *RNC* in dieser Form entfallen lassen. Im Moby Dick-Projekt wurde ein Teil der *RNC*-Funktionalität in die *QoS*-Broker implementiert, was aber in dieser Arbeit ausdrücklich nicht erfolgen soll. Auch diese Funktionalität wurde verteilt in die Zugangsrouters implementiert.

Die Signalisierung ist ebenfalls in aktuelle Lösungen der IETF eingebettet, so dass auch hier sichergestellt ist, dass die Vorschläge mit überschaubarem Aufwand umgesetzt werden könnten. Das Sitzungsmanagement hingegen ist neu. Auch hier wurde sichergestellt, dass eine Einbettung in aktuelle IETF-Konzepte sichergestellt ist. Weiter wurden die Nutzungsdatenberichte und die prinzipielle Struktur des Dienstangebote von den *IPDR*-Spezifikationen abgeleitet.

7 Validierung und Verifikation des Gesamtsystems

7.1 Einleitung

In diesem Kapitel erfolgt die Validierung und Verifikation der bislang gemachten architektonischen Vorschläge. Methodisch erfolgt dies mittels einer Spezifikation, die durch formale Methoden validiert und verifiziert wird. Konkret wird die aus der ITU stammende SDL/MSK-Sprache (siehe Kapitel 7.2) eingesetzt, welche aus dem Bereich der Telekommunikation kommt und speziell während der Spezifikationsphase von 2G und 3G verbreitet eingesetzt wurde, und nun mit UML migriert. Im Umfeld des Internets hat sich diese formale Protokollentwicklung noch nicht durchgesetzt. Die prinzipielle Vorgehensweise ist jedoch grundsätzlich nicht neu und wurde z.B. in [80] im Rahmen eines internationalen Forschungsprojektes angewendet.

Als Kernbeitrag dieses Kapitels wird ein Weg vorgezeichnet, der einem Netzanbieter eines Mobilens Internets ermöglicht, durch eigene Parametrisierung den für sich optimalen Kompromiss zwischen Dienstvielfalt inklusive Ausprägung der QoS-Unterstützung auf der einen Seite, und Abrechnungsaufwand auf der anderen Seite festzulegen. Eine unter rein ökonomischen Gesichtspunkten vernünftige Gewichtung dieser sich gegenseitig negativ beeinflussenden Aspekte ist aus betriebswirtschaftlichen Gründen für einen Netzbetreiber zwingend notwendig, um kommerziell erfolgreich zu sein, da hier Fixkosten reduziert werden können. Mittels des aufgezeigten methodischen Vorgehens entsteht ein Werkzeug, welches hilft, gewinnmaximiert (transaktionskostenminimiert) Produktbündel zu definieren, welche kommerziell am Markt angeboten werden. Weiter können konkrete Aussagen über den Grad an QoS-Verhandlung bzw. QoS-Granularität gemacht werden, die ein Netzbetreiber einem Benutzer aus betriebswirtschaftlicher Sicht anbietet.

Es ist zu erwarten, dass einem Benutzer die Möglichkeit, Dienstgüteparameter mit dem Netz ggf. dynamisch zu verhandeln, nur dann angeboten wird, wenn der inkrementelle Umsatz aus diesem Angebot die inkrementellen Kosten, welche zusätzlich durch den Abrechnungsprozess bzw. den zusätzlichen Investitionen in die Abrechnungsinfrastruktur entstehen, übersteigt.

7.2 SDL und MSC

In der Telekommunikation besteht historisch die Notwendigkeit herstellerunabhängige Standards für Kommunikationsprotokolle und dienste zu erstellen. Dies führte innerhalb der ITU zur Entwicklung der formalen Modellierungssprachen SDL (System Description Language) und MSC (Message Sequence Chart), welche in [62] bzw. [63] spezifiziert sind. Historisch grenzen sich diese Sprachen zu UML [23] dadurch ab, dass bei SDL/MSK eine formale Semantik zugrunde liegt, welche die Bedeutung der einzelnen Sprachkonstrukte definiert.

Formal wurden für beide Sprachen sowohl eine graphische Repräsentation, als auch eine textuelle Notation erstellt. Die textuelle Notation spielt in dieser Arbeit keine wesentliche Rolle und wird somit nicht weiter behandelt.

Generell wird MSC vorwiegend zur Anforderungsdefinition, Testfallspezifikation und Dokumentation eingesetzt, während SDL überwiegend in der Spezifikations- und Implementierungsphase von komplexen Protokollabläufen eingesetzt wird, welche mehrere zeitgleich ablaufende Prozesse beinhalten, die mittels Nachrichten Informationen austauschen. Der Einsatz von SDL zur Leistungsbewertung von Kommunikationsprotokollen wurde u.a. in [115] eingehend erläutert. SDL ist hierarchisch aufgebaut und erlaubt somit das Abstrahieren bestimmter Nachrichtengehalte durch das formale Verbergen bestimmter Informationen und

unterstützt die Spezifikation zeitkritischer Abläufe. Durch formale Regeln sind Werkzeuge zur Validierung und Verifizierung der Spezifikation verfügbar, welche unter Verwendung von formalen Schnittstellen die SDL-Spezifikation in die Lage versetzen, mit der Umgebung zu kommunizieren.

In der Version SDL 2000 wurde begonnen, SDL mit UML zusammenzuführen. Hierfür wurden abstrakte Zustandsmaschinen (Abstract State Machine – ASM) eingeführt. Die grundsätzliche Motivation dieser Zusammenführung liegt in der Notwendigkeit, SDL mit objektorientierten Modellierungssprachen anzureichern und in der Anforderung an SDL, mittelfristig automatische Software-Generierung zu unterstützen.

SDL-Systeme bestehen aus mehreren simultan ablaufenden Prozessen, welche untereinander Nachrichten austauschen. Jeder Prozess ist in SDL mittels einer Zustandsmaschine beschrieben. SDL verfügt über eine Grammatik, welche die Entwicklung von Werkzeugen zur Simulation und Validierung der komplexen Systeme mit dem Ziel unterstützt, Fehler möglichst vor der eigentlichen Implementierung der Systeme entdecken zu können.

MSCs wurden als graphische Beschreibungssprache spezifiziert, um das Zusammenspiel in vorwiegend asynchron ablaufenden Vorgängen und somit das Kommunikationsverhalten zwischen verschiedenen Instanzen bzw. Systemkomponenten zu beschreiben. Die Kommunikation wird hierbei durch den Austausch von Nachrichten spezifiziert. Eine MSC-Variante wurde in UML als *Sequence Diagrams* integriert.

MSCs werden sehr häufig im Zusammenhang mit SDL eingesetzt, um z.B. Anforderungen, welche aus einem MSC abgeleitet werden, dazu zu verwenden, SDL-Modelle zu validieren. In der Telekommunikationsindustrie hat sich MSC als normierte Sprache zur Beschreibung von Beispielabläufen durchgesetzt. Sie wird dabei in allen Phasen - von der Anforderungsanalyse bis zum Test - eingesetzt. In Abbildung 42 ist ein einfacher MSC dargestellt. Generell können bei einem MSC neben den Kommunikations- und Timerereignissen auch Aktions- und Instanzerzeugungsereignisse definiert werden. MSC abstrahiert von den Details einer Komponente und fokussiert sich auf den Informationsaustausch zwischen den Hauptkomponenten eines Systems. Nun soll auf einige MSC-Basiskonstrukte näher eingegangen werden.

7.2.1 MSC-Basiskonstrukte

Innerhalb von MSC wurden folgende grundlegenden Sprachkonstrukte für die Spezifikation von Nachrichtenflüssen festgelegt: *Instance*, *Message*, *Environment*, *Action*, *TimerStart*, *Timeout*, *TimerStop*, *Create*, *Stop* und *Condition*. Im Folgenden soll auf diese Sprachkonstrukte näher eingegangen werden.

7.2.1.1 Instance und Message

Komponenten des Typs *Instance* werden benutzt, um untereinander oder mit der Systemumgebung asynchron Nachrichten auszutauschen. In der graphischen Form werden *Instances* als vertikale Linien dargestellt. Innerhalb des *Instance*-Kopfes wird typischerweise der Instanzname spezifiziert.

Eine *Message* wird durch Pfeile dargestellt, welche entweder horizontal, oder aber (um einen Zeitverlauf anzuzeigen) geneigt sein können. Eine *Message* definiert mittels des Pfeilanfangs das Senden der Nachricht und mittels der Pfeilspitze das Verarbeiten der Nachricht. Typischerweise werden die Pfeile mit Namen und der Nachricht und deren Parametern beschriftet.

7.2.1.2 Environment

Die Diagrammfläche eines MSC wird durch einen rechteckigen Rahmen begrenzt. Der Diagrammrahmen definiert die Systemumgebung. *Environment*-Nachrichten, die aus der Systemumgebung kommen oder an die Systemumgebung gesendet werden, beginnen und enden auf diesem Rahmen.

7.2.1.3 Action

Zusätzlich zur Nachrichtenkommunikation können *Actions* von Instanzen spezifiziert werden. Eine *Action* wird durch ein Rechtecksymbol dargestellt, das beliebigen Text enthalten kann.

7.2.1.4 Create und Stop

Die Konstrukte *Create* und *Stop* werden für die dynamische Erzeugung (*Create*) und Terminierung (*Stop*) von Instanzen eingesetzt. Ein *Create* wird durch einen gestrichelten Pfeil mit optionalen Parametern beschrieben welcher an der Erzeuger-Instanz beginnt und am Kopf der erzeugten Instanz endet. Eine Instanz kann sich selbst durch eine *Stop*-Aktion terminieren. Ein *Stop* wird graphisch durch ein Kreuz am Ende der Instanzachse spezifiziert.

7.2.1.5 Timer

Zur Beschreibung von Timern bietet MSC die Sprachkonstrukte *Timer-Start*, *Timeout* und *Timer-Stop* an. *Timer-Start* spezifiziert das Setzen, *Timeout* den Ablauf und *Timer-Stop* das Zurücksetzen eines Timers. In MSC ist ein Timer immer einer Instanz zugeordnet.

7.2.1.6 Condition

Eine *Condition* beschreibt einen Zustand, der sich auf eine Menge der im MSC enthaltenen Instanzen bezieht. Graphisch werden *Conditions* durch Sechsecke dargestellt, welche die Instanzen, auf die sich die Bedingung bezieht, überdecken. *Conditions* werden zur Beschreibung von wichtigen Systemzuständen benutzt.

7.2.2 Telelogic TAU

Im Rahmen der Spezifikation des Gesamtsystems wurde das Produkt TAU der Firma Telelogic in der Version 2.5 eingesetzt [124]. Ausgehend von einem Komponentenmodell, welches eine syntaktische Validierung der Spezifikation automatisch vollzieht, ist es möglich, aus dem Modell heraus die MSCs dynamisch zu generieren und die erforderliche Logik der Komponenten mittels externer Interaktion zu steuern. Somit ist eine dynamische Validierung der modellierten und spezifizierten Prozessabläufe unter qualitativer Berücksichtigung von Laufzeiteffekten möglich.

TAU bietet die Möglichkeit, eine Architektur mittels UML 2.0 zu definieren, welche dann detailliert beschrieben werden kann. TAU unterstützt SDL-2000. Somit können UML-Diagramme direkt in SDL umgesetzt werden und in einem integrierten Designprozess benutzt werden. TAU bietet die Möglichkeit, das Verhalten der Komponenten mittels externen Programmierschnittstellen (C, C++) in eine Systemspezifikation mit einzubinden. Im Anschluss daran ist es möglich, das Gesamtsystem zu validieren, indem dynamisch zur Verifikation ein MSC generiert wird.

7.3 Modellierung

7.3.1 Ziel der Modellierung

Will man verschiedene Wertschöpfungsketten und Dienstbündel im kommerziellen Mobil Internet vergleichen und einer Leistungsbewertung bzw. Aufwandsabschätzung unterziehen,

so müssen die Auslöser von Benutzerverwaltungsprozeduren, der Ablauf dieser Prozeduren und die sie ausführenden und unterstützenden Instanzen geeignet modelliert werden. Dies geschieht in der vorliegenden Arbeit mit Hilfe eines Lastmodells, eines Teilnehmer- und Betreibermodells, eines Vertragsmodells und eines Netzmodells. Aus dem Teilnehmer/Betreibermodell ergibt sich, welche Benutzerverwaltungsprozeduren wie häufig und in welcher Weise angestoßen werden. Mit Hilfe des Lastmodells wird beschrieben, in welcher Weise diese Prozeduren ablaufen und welche Last in Form von Datenaufkommen und zu übertragenden Nachrichten an welchen Stellen anfällt. Im Netzmodell schließlich wird dargestellt, in welcher Weise die anfallende Last verarbeitet wird, so dass schließlich im Rahmen einer Leistungsbeurteilung charakteristische Größen, wie zum Beispiel zulässige Dienstvielfalt und optimale Dienstkomposition, ermittelt werden können.

7.3.2 Teilnehmer- und Betreibermodell

Das Teilnehmermodell ist in Abbildung 58 dargestellt und lehnt sich an die in Kapitel 4.8 vorgestellte Architektur an. Im Modell ist ein Heimatbetreiber (Betreiber A in Abbildung 58) dargestellt, welcher den direkten Zugang zum Benutzer hat und welcher über Roaming-Abkommen diesem Benutzer Dienste außerhalb des eigenen Netzbereiches anbietet (Betreiber B und C in Abbildung 58). Die topologische Lage der Kommunikationsgegenstelle (CN, vergl. Kapitel 2.1.4) ist für die weiteren Überlegungen nicht relevant. Jeglicher Einfluss von kommerziell betriebenen Transfernetzen wird in diesem Modell nicht berücksichtigt.

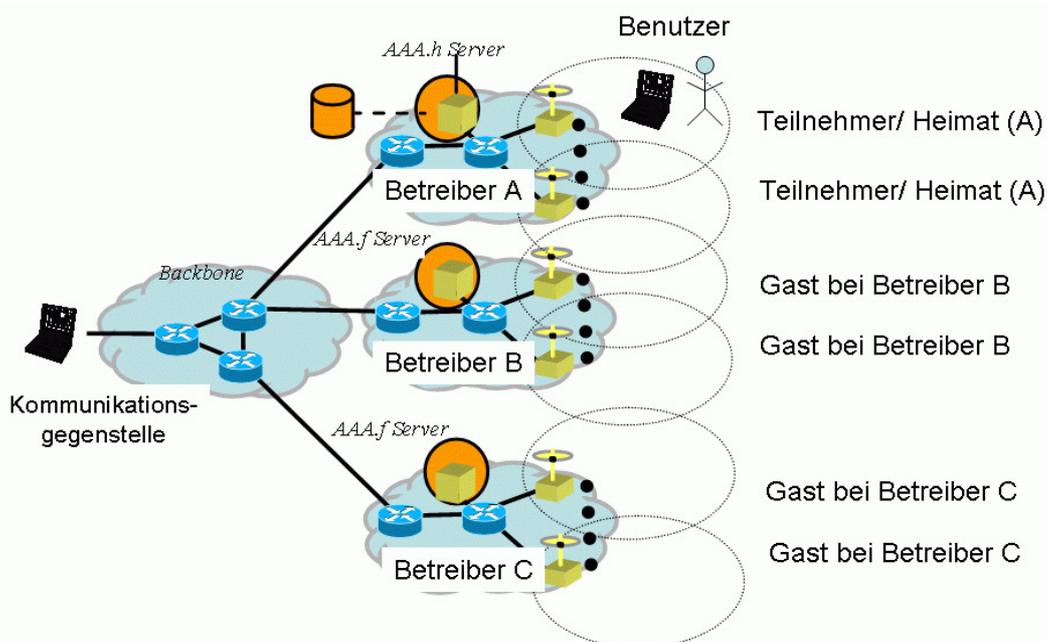


Abbildung 58: Teilnehmer- und Betreibermodell

7.3.3 Netzmodell

Bei der Spezifikation und Leistungsbewertung des hier vorgeschlagenen hierarchischen Konzeptes zur Nutzungsdatenerfassung und der nachgelagerten Weiterverarbeitung wurde auf das im Bereich 2G und 3G verwendete Manhattan-Grid-Referenzmodell [80] zurückgegriffen, welches exemplarisch in Abbildung 59 dargestellt ist.

Generell wird beim Manhattan-Grid-Modell davon ausgegangen, dass Netzübergänge nur entlang der Straßen (Kanten) möglich sind. Im konkreten Modell sind 9 vertikale Straßenzüge und 5 horizontale Straßenzüge definiert. An jedem Schnittpunkt zweier Straßen ist ein Zu-

gangsrouter positioniert, der in die vier Richtungen jeweils das halbe Gebiet bis zum nächsten Knoten abdeckt.

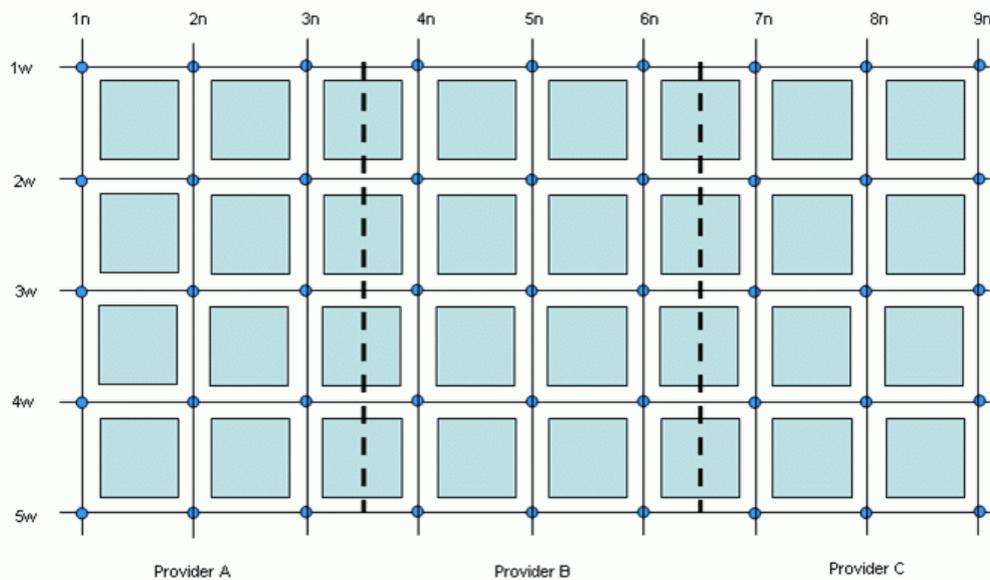


Abbildung 59: Verwendetes Manhattan-Grid-Modell

Um auch administrative Netzübergänge geeignet berücksichtigen zu können, wurde das Gesamtmodell analog zu Kapitel 7.3.2 auf drei Netzbetreiber aufgeteilt. Diese Aufteilung geschieht in einer Weise, dass jeder der Netzbetreiber seine Dienste nur in einem Teilgebiet anbietet, dort jedoch exklusiv. Im Modell hat somit jeder Betreiber 22 Netzübergänge innerhalb seines administrativen Bereiches und 5 Übergänge zum angrenzenden administrativen Bereich. Zur Validierung und Verifikation des Modells bewegt sich das mobile Endgerät/der Benutzer im Modell zunächst auf vorgegebenen Wegen durch den Manhattan-Grid. Hierfür wurden 4 Wege definiert, welche im Anhang B.2 beschrieben sind.

7.3.4 Mobilitätsmodell

In dieser Arbeit wurde neben fest vorgegebenen und somit replizierbaren Netzübergangssequenzen das Random-Walk-Modell [95] verwendet, bei dem sich die Endsysteme/Benutzer im Raum bewegen können. Dieses Modell wird oft dazu eingesetzt, das Verhalten von Fußgängern zu modellieren, welche häufig unkontrolliert stehen bleiben, um dann die Bewegung in beliebiger Richtung fortsetzen. Die Wahrscheinlichkeit, dass ein Teilnehmer einen Zellewechsel vollzieht, wird für die weiteren Betrachtungen in dieser Arbeit gemäß Abbildung 60 festgelegt.

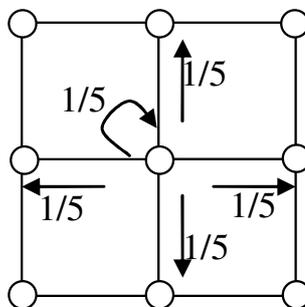


Abbildung 60: Übergangswahrscheinlichkeiten

7.3.5 Lastmodell

Das Lastmodell beschreibt das genaue Kommunikationsverhalten eines Teilnehmers. In dieser Arbeit repräsentiert das Lastmodell die Nutzdaten, welche vom Benutzer generiert wurden.

Diese Nutzdaten stoßen dann den Prozess der Nutzungsdatenerfassung entsprechend dem strategischen Erfassungskonzept an. Die Verteilung der Nutzdaten innerhalb eines Intervalls der Nutzungsdatenerfassung ist im Kontext dieser Arbeit generell nicht relevant. Somit kann in erster Näherung ein Benutzer innerhalb eines Erfassungsintervalls entweder aktiv oder inaktiv sein. Weiter ist festzuhalten, dass der Aufwand an die Abrechnungsinfrastruktur unabhängig von der Intensität der Nutzlast des Benutzers ist. Sobald eine abrechnungsrelevante Aktion von Benutzer ausgeht, wird der Prozess der Nutzungsdatenerfassung angestoßen.

Je nach Strategie der Nutzungsdatenerfassung und nachgelagerten Weiterverarbeitung kann bei Inaktivität eines Benutzers innerhalb eines abrechnungsrelevanten Intervalls ggf. kein Nutzungsdatenprotokoll im System erzeugt werden. Um die Leitungsfähigkeit des Systems zu validieren und zu verifizieren, kann vereinfacht davon ausgegangen werden, dass nur ein Teilnehmer in jeder Zelle und innerhalb jedes Abrechnungsintervalls aktiv ist. Hieraus folgt, dass vom mobilen Endgerät permanent abrechnungsrelevante Daten in Höhe eines normierten Einheitswertes versendet/empfangen werden. Somit ergibt sich aus dem Lastmodell eine normierte konstante Datenquelle mit einer Nutzungsdatenrate von $1/\text{Intervall}$ und Dienstklasse.

7.3.6 Vertragsmodell

Generell wird in der durchgeführten Validierung und Verifikation des Gesamtsystems aus Kapitel 6 vom Falle eines Heimatbetreibers ausgegangen, welcher mit weiteren Betreibern ein Roaming-Abkommen unterhält. Dieses Abkommen ist in Anlehnung an die gängige 2G-Praxis derart ausgestaltet, dass ein Dienst über alle drei Netze hinweg angeboten wird. Prinzipiell wäre es denkbar, dass die einzelnen Dienstklassen im Netz eines Roaming-Partners entsprechend den lokalen Netzgegebenheiten umgesetzt werden. Dies soll jedoch im Rahmen dieser Arbeit nicht berücksichtigt werden. Die Details des Vertrages zwischen Benutzer und Heimatnetzbetreiber werden in Anlehnung an Kapitel 5.5 vereinfacht gemäß Tabelle 12 definiert. Es gibt zwei verschiedene Dienstklassen, deren Struktur in Tabelle 12 dargestellt ist. Die Bezahlform wurde als normaler Festvertrag und als Prepaid angenommen. Basierend auf Kapitel 6.5.1 ergibt sich ein Nutzungsdatenbericht von 500 byte.

Tabelle 12: Zur Validierung gewählte Verkehrsklassen

Klasse 1	Max Bandbreite 1	Max. Datenvolumen 1
Klasse 2	Max Bandbreite 2	Max. Datenvolumen 2

7.3.7 Strategien zur Verwaltung der Nutzungsdaten

Ein Kernproblem des Mobilens Internets im Kontext dieser Arbeit ist das Nichtvorhandensein einer expliziten Abmeldungsprozedur, wenn einen Benutzer bzw. sein Endsystem vom, administrativen Netz kontrolliert trennt. Ursache dieser Trennung kann sowohl ein plötzliches „Verschwinden“ des Endsystems z.B. durch geplantes oder auch nicht geplantes Abschalten, oder auch ein Netzübergang nach Kapitel 6.2.1.4 sein. Somit stellt sich die Frage, wie ein idealer Zeitpunkt definiert werden kann, zu dem Nutzungsdateninformationen zur weiteren Datenkonsolidierung weitervermittelt werden. Prinzipiell sind hierfür vier Strategien möglich, auf die nun näher eingegangen wird.

7.3.7.1 Zeitgesteuerte Weiterleitung der Nutzungsdaten

Bei dieser Strategie wird über eine Zeitsteuerung in äquidistanten Intervallen die in der auswärtigen Domäne erfassten Nutzungsdaten an die Heimatdomäne weitergeleitet.

7.3.7.2 Mengengesteuerte Weiterleitung der Nutzungsdaten

Bei dieser Strategie werden die erfassten Nutzungsdaten erst dann an das Heimatnetz des Benutzers weitergeleitet, wenn das konsumierte Datenvolumen einen bestimmten Schwellwert

überschritten hat. Zunächst erscheint diese Strategie sehr plausibel, da der Schwellwert Kernbestandteil des Preismodells nach Kapitel 5.5 ist. Jedoch müssen geeignete Strategien und Konzepte entwickelt werden, welche verhindern, dass das Weiterleiten der Berichte ausbleibt, sollte der Schwellwert nicht (ggf. nie) erreicht werden.

7.3.7.3 Betreibergesteuerte Weiterleitung der Nutzungsdaten

Bei dieser Methode könnte ein Netzbetreiber die erfassten Nutzungsdaten so lange in seinem eigenen Netz konsolidieren, wie sich der Teilnehmer im entsprechenden Netz befindet. Dieses Intervall wird idealerweise mit einem maximalen zulässigen Zeitobergrenze abgeschlossen, welches in der Größenordnung von Tagen ist. Diese Strategie minimiert das Datenaufkommen der anfallenden Nutzungsdaten erheblich.

7.3.7.4 Ereignisgesteuerte Weiterleitung der Nutzungsdatenberichte

Bei dieser Strategie wird vorgeschlagen, die Nutzungsdatenberichte in Koordination mit dem Intervall zur Neuregistrierung nach Kapitel 6.2.1.2 weiterzuleiten. Hier wäre dann die Frage zu diskutieren, ob die Weiterleitung zusätzlich an die Mobile IP-Registrierung nach Kapitel 2.1.4 gebunden sein kann, da sich die in Kapitel 6.2.1.2 beschriebene Registrierung nicht auf administrative Domänen, nicht jedoch auf Netzzellen beschränkt.

7.3.7.5 Bewertungskriterien der vorgestellten Strategien

Generell sind bei der Evaluierung der vorgeschlagenen strategischen Ansätze verschiedene Bewertungskriterien festzuhalten, welche diese Evaluierung wesentlich beeinflussen. Dies sind:

- Datenmenge des Nutzungsdanenaufkommens im Heimatbetreiber
- Aufwand des Nutzungsdatenmanagements im auswärtigen Netz
- Sicherstellung der Liquidität eines Betreibers
- Minimierung des Signalisierungsverkehrs

Selbstverständlich ist die Sicherstellung der Liquidität ein Kernkriterium. Ein Betreiber muss zu einem bestimmten Zeitpunkt in der Lage sein, eine Rechnung erstellen zu können (setzt man kein Prepaid voraus). Diese Garantie kann strenggenommen nur von einem zeitgesteuerten Mechanismus (Kapitel 7.3.7.1) erbracht werden.

Bewegt sich ein Benutzer nicht, so wird bei einer betreiber- bzw. ereignisgesteuerten Weiterleitung kein Datenbericht innerhalb des Berichtsintervalls generiert, was zu einer erheblichen Reduktion der Nutzungsdaten führt.

Eine mengengesteuerte Weiterleitung der Nutzungsdaten bei Benutzern mit relativ wenig Nutzungsdanenaufkommen ist vorteilhaft, wenn sichergestellt ist, dass die Nutzungsdatenberichte grundsätzlich erzeugt werden.

7.4 Referenzberechnung

7.4.1 Nutzungsdatenaufkommen

In diesem Abschnitt soll, basierend auf dem modellierten Beispielszenario, das fiktive Datenaufkommen im Mobil Internet ohne die im Rahmen dieser Arbeit gemachten Modifikationen berechnet und mit einer Architektur mit den vorgeschlagenen Modifikationen verglichen werden.

Generell entsteht bei einem Nutzungsdatenerfassungskonzept nach existierendem Standard, also ohne die im Rahmen dieser Arbeit vorgeschlagenen Verbesserungen, ein Nutzungsdatenaufkommen nach Gleichung (1), wobei der Parameter $AIntervall$ das Erfassungsintervall darstellt.

$$G_Nutzungsdatenaufkommen = \sum_{i=1}^{Nutzer} \frac{Nutzungsdatensatz}{AIntervall(i)} \quad (1)$$

Das Erfassungsintervall muss generell sehr klein gewählt werden, will man in dieser Infrastruktur Prepaid-Dienste anbieten. In der Praxis und für die weitere Abschätzung des Nutzungsdatenaufkommens sei hier ein $AIntervall$ von 1 Sekunde angenommen. Bei einem Nutzungsdatenbericht von 500 byte würde sich hier pro Benutzer und Tag ein Datenvolumen von 43,2 Mbyte pro Benutzer ansammeln, was bei einer monatlichen Abrechnung zu einer Datenmenge von 1,3 Gbyte ansteigen würde.

In dieser Beispielrechnung ist der Aufwand für die „Gäste“ im Netz, also das Datenvolumen der Benutzer, für die der AAA.h-Server die Rolle des AAA.f-Servers einnimmt, noch nicht berücksichtigt. Weiter wurde in dieser Betrachtung von einer permanenten Aktivität aller Benutzer („Always-On“) ausgegangen, was in aktuellen Telekommunikationsszenarien als (noch) undenkbar erscheint. Jedoch ist es heute keine Seltenheit, dass z.B. Arbeitsplatz-PCs im Büro permanent eingeschaltet sind. Weiter ist genau dies erforderlich, setzt man den Trend hin zur SIP-basierten Telefonie konsequent um, da die Endsysteme permanent im Netz registriert sein müssen.

Ein Umsetzen der vorgeschlagenen Konzepte und Mechanismen würde zu einem Nutzungsdatenaufkommen nach Gleichung (2) führen. Hier würden Flat-Rate-Benutzer keinerlei Nutzungsdaten erzeugen. Prepaid-Benutzer würden dann Nutzungsdaten entsprechend Gleichung (1) erzeugen (Hier kann der Nutzungsdatensatz an die benutzerspezifischen Anforderungen angepasst werden und ist ggf. kleiner; dies wird in der weiteren Betrachtung vernachlässigt). Weiter wird in den folgenden Betrachtungen angenommen, dass das Abrechnungsintervall $AINTP$ dem Abrechnungsintervall $AIntervall$ aus Gleichung (1) entspricht. Traditionelle Postpaid-Kunden würden in einem deutlich reduzierten Ausmaß zum Gesamtnutzungsdatenaufkommen beitragen. Der entsprechende Reduktionsfaktor K drückt sich in der jeweiligen Datenreduktionsstrategie nach Kapitel 7.3.7 aus, welche mit dem Standardvorgehen verglichen wird.

$$Nutzungsdatenaufkommen = \sum_{i=1}^{Prepaid} \frac{Nutzungsdatensatz(i)}{AINTP(i)} + \sum_{j=1}^{Flat} 0 + \sum_{k=1}^{PostPaid} \frac{Nutzungsdatensatz(k)}{K(k)} \quad (2)$$

Die Ergebnisse der Vergleichsberechnung des Nutzungsdatenaufkommens sind in Tabelle 13 dargestellt. Der jeweilige Gewinn der letzten Spalte drückt die resultierende Datenreduktion gegenüber einer Berechnung nach Gleichung (1) aus. Das Beispiel verdeutlicht, dass der Anteil der „Flat-Kunden“ direkt und linear das Nutzungsdatenaufkommen reduziert. Der Reduktionsfaktor verringert dann weiter das Nutzungsdatenaufkommen der Postpaid-Benutzer. Die Tabelle zeigt weiter, dass sich hier bei großem Anteil von Postpaid-Benutzern eine weitere Reduktion des Nutzungsdatenaufkommens erzielen lässt. Der Beitrag hier ist jedoch negativ exponentiell, so dass ein relativ geringer Aufwand bereits signifikant zur Datenreduktion beiträgt.

Tabelle 13: Ergebnisse bei 50 % Anteil von Flat-Rate-Benutzern

Anteil Prepaid [%]	Anteil Postpaid [%]	Anteil Flat [%]	Reduktionsfaktor K	Reduzierter Anteil [%]
25	25	50	1	50
25	25	50	2	37,5
33	17	50	2	41,5
17	33	50	2	33,5
25	25	50	100	25,25
33	17	50	100	33,17
17	33	50	100	17,33
25	25	50	∞	25
33	17	50	∞	33
17	33	50	∞	17

Ist der Anteil an Postpaid-Kunden signifikant höher und spielen Flat-Rate bzw. Prepaid keine wesentliche Rolle beim Kundenstamm eines Netzbetreibers, ergibt sich ein Ergebnis nach Tabelle 14. Auch dieses Beispiel zeigt, dass sich bei hohem Anteil von Postpaid-Benutzern über einen Reduktionsfaktor ein signifikanter Gewinn beim Nutzungsdatenaufkommen ergibt. Um sich die absoluten Dimensionen des jeweiligen Gewinns zu verdeutlichen, soll folgende Beispielrechnung gemacht werden:

Tabelle 14: Ergebnisse bei 90 % Anteil von Postpaid-Benutzer

Anteil Prepaid [%]	Anteil Postpaid [%]	Anteil Flat [%]	Reduktionsfaktor	Reduzierter Anteil [%]
5	90	5	1	95
5	90	5	2	50
5	90	5	10	14
5	90	5	100	5,09
5	90	5	∞	5

Der Parameter $A_{\text{Intervall}}$ bzw. A_{INIP} nach (1) und (2) muss wegen der Existenz von Prepaid-Kunden sehr klein gewählt werden und wird hier mit 1 Sekunde angenommen. Bei einem Nutzungsdatenbericht von 500 byte würde sich pro Benutzer und Tag ein Datenvolumen von 43,2 Mbyte pro Benutzer ansammeln, was bei einer monatlichen Abrechnung zu einer Datenmenge von 1,3 Gbyte ansteigen würde. Bei 27 Millionen Benutzern (T-Mobile in Deutschland) und der Annahme dass alle dem „Always-On“ Paradigma folgen, würde dies zu einem Nutzungsdatenaufkommen von 32 Petabyte pro Monat führen. Hier wären „Gäste“, also Nutzer eines Roaming-Partners im Netz noch nicht berücksichtigt.

7.4.2 Bandbreitenabschätzung

Die Übertragung von 32 Petabyte mit einer Netzanbindung von 10 Gbit/s würde 27 Millionen Sekunden – also ca. 11 Monate dauern. Die Übertragung dieser Datenmengen wird von der aktuellen Netzinfrastruktur nicht ausreichend unterstützt. Weiter ist das Verarbeiten der Daten zur finalen Rechnungserstellung zusätzlich rechenintensiv und somit sehr teuer. Dies ist als wesentliche Hürde bei der weiteren Kommerzialisierung des Mobilens Internets zu betrachten.

Nimmt man jedoch die aus Kapitel 6.5.2 erwähnten 25 AAA-Server in einem Netz mit 27 Millionen beheimateten Kunden an, welche dann jeweils für einen Teil der Benutzer die Rolle des AAA.h einnehmen, und nimmt man weiter einen Reduktionsfaktor der zu verarbeitenden Datensätze von 70 %, dann würden 9 Petabyte auf 25 AAA-Server verteilt werden müssen.

Dies hätte ein Datenvolumen von 380 Terrabyte pro AAA-Server zur Folge. Diese Datenmenge wäre nach obiger Beispielrechnung in 4 Tagen über eine 10 Gbit/s Schnittstelle übertragbar.

7.5 Parametrisierung im Gesamtmodell

7.5.1 Grundsätzliche Überlegungen

Generell können im Kontext dieser Arbeit die Benutzer in die Kategorie *Prepaid*, *Postpaid* oder *Flat-Rate* eingruppiert werden. Die in Kapitel 7.4 vorgestellten einführenden Referenzbetrachtungen verdeutlichen, dass durch die hier vorgestellten Vorschläge einer benutzerspezifischen Nutzungsdatenerfassung eine erhebliche Datenreduktion bei Flat-Rate Kunden haben, welche direkt und linear das Nutzungsdatenaufkommen reduziert.

Es ist offensichtlich, dass bei Postpaid-Benutzern der Datenreduktionsfaktor eine wesentliche Rolle spielt und je nach Gesamtkundenverteilung einen entsprechenden Beitrag zur Datenreduktion beim gesamten Nutzungsdatenaufkommen beiträgt. Bezüglich der Parametrisierung dieses Parameters für Postpaid-Benutzer kann aus den überschlägigen Betrachtungen aus Kapitel 7.4 festgehalten werden, dass mit einem relativ geringen Faktor, eine signifikante Datenreduktion erreicht werden kann. Ein wesentliches Kriterium ist die Festlegung des Erfassungsintervalls, welches idealerweise ohne signifikanten Mehraufwand in die Gesamtarchitektur integriert werden sollte. Detaillierte Vorschläge hierzu werden in Kapitel 8.2 gegeben.

Da auch bei diesen Vertragsmodellen kein weiterer Optimierungsspielraum gegeben ist, fokussieren sich die weiteren Überlegungen auf den Fall *Prepaid*.

7.5.2 Parametrisierung bei Benutzern mit Prepaid-Vertrag

Bei Benutzern mit Prepaid-Vertrag wurde bei einer Betrachtung des Nutzungsdatenaufkommens gemäß Kapitel 7.4 kein potentieller Beitrag zu einer Reduktion des erfassten Datenvolumens beigetragen. Das generelle Problem bei dieser Benutzerkategorie ist die Anforderung an das Gesamtsystem, die erfassten Nutzungsdaten in Echtzeit mit dem Restguthaben eines Benutzers abzugleichen, um unentgeltlichen Ressourcenverbrauch auszuschließen [77]. Generell wird das Restguthaben durch *Credits* repräsentiert, welche dann basierend auf dem Preismodell einen entsprechenden Ressourcenverbrauch repräsentieren. In Kapitel 7.4 wurde angenommen, dass die *Credits* im AAA.h-Server verwaltet werden. Bei dieser Strategie ergibt sich kein weiteres Optimierungspotential, so dass diese Strategie im Grundsatz in Frage gestellt wird.

Je weiter jegliche nachgelagerte Weiterverarbeitung der Nutzungsdaten in Richtung der Nutzungsdatenentstehung verschoben wird, desto höher der mögliche Konsolidierungsgewinn, welcher die Menge der über das Netz übertragenen Nutzungsdaten reduziert. Analog hierzu kann festgehalten werden, dass dieser Zusammenhang auch für die Verwaltung der *Credits* Gültigkeit hat. Je stärker das Restguthaben Richtung des „Konsumenten“ delegiert wird, desto effektiver arbeitet das Gesamtsystem, da sich hier ebenfalls ein Konsolidierungsgewinn einstellen kann. Diese Maßnahme erfordert jedoch zusätzlichen Signalisierungsaufwand und eine besondere Betrachtung bei administrativen Betreiberübergängen. Setzt man diese Überlegungen konsequent um, so würde man sowohl die Restguthabenverwaltung, als auch Aufbereitung der Nutzungsdaten im Endsystem durchführen wollen. An dieser Stelle ist zu diskutieren, wer in B3G-Netzen Besitzer des Endsystems sein wird. Eine strengere Einbindung des Endsystems in das Betreibermodell des Netzbetreibers könnte eine solche Realisierung zulassen. Dies hätte jedoch zur Folge, dass die im Rahmen dieser Arbeit angenommene Entkopplung von Benutzer und Endgerät neu überdacht werden müsste. Somit beschränken sich die

Überlegungen im Rahmen dieser Arbeit auf die Netzkomponenten, welche absehbar im operativen Besitz des Netzbetreibers sein werden und schliessen diese Möglichkeit aus. Es ergibt sich für das Verwalten der *Credits* ein Optimierungsraum nach Abbildung 61, auf den im Folgenden näher eingegangen wird.

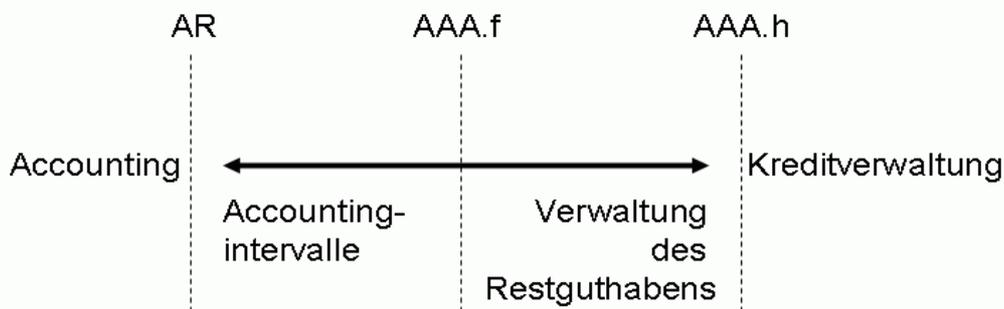


Abbildung 61: Optimierungsraum der Kreditverwaltung

7.5.2.1 Restguthabenverwaltung im Heimatnetz

Bei dieser Strategie werden alle *Credits* (vergl. Kapitel 7.4) in der Heimatdomäne (AAA.h-Server) verwaltet. Basierend auf den Nutzungsdatenbereichen werden die verfügbaren *Credits* kontinuierlich neu berechnet. Im Falle eines totalen Verbrauchs der *Credits* ist diese neue Situation in Echtzeit an die entsprechende Netzinstanz, welche den aktuellen Ressourcenkonsum kontrolliert, bzw. an eine administrative Instanz einer auswärtigen Domäne, zu senden. Dies erfordert eine Information des Zugangsrouters (QoS-Manager nach Abbildung 54 bzw. Anhang C). Hieraus folgt, dass diese Information (in Echtzeit) ggf. über Betreiber Grenzen hinweg kommuniziert werden muss. Weiter ist das Nutzungsdatenaufkommen sehr hoch.

7.5.2.2 Restguthabenverwaltung im AAA.f-Server

Als weitere grundsätzliche Möglichkeit könnten die *Credits* im AAA.f-Server verwaltet werden. Konkret würde ein Betreiber eine Anzahl von *Credits* vorab bei der Registrierung an die auswärtige Domäne „überweisen“, welche dort entsprechend verwaltet werden. Bei Unterschreiten des Restguthabens unter einen Schwellwert müssten weitere *Credits* „in Vorleistung“ bei der Heimatdomäne angefordert werden, welche dann entweder in der auswärtigen Domäne verbraucht, oder unverbraucht zurückgegeben werden. In jedem Fall müssten die Zugangsrouten bei Implementierung dieser Strategie weiterhin die Nutzungsdaten für diese Kunden in Echtzeit an den AAA.f-Server weiterleiten. Da es sich beim AAA.f bzw. AAA.h-Server um eine Rollenbezeichnung handelt, ist der Mehrwert dieser Strategie auf eine reduzierte Netzlast im Kernnetz zwischen den involvierten Betreibern begrenzt. Geht man weiter davon aus, dass ein Benutzer Ressourcen vorwiegend im Heimatnetz konsumiert, ist der Gewinn, welcher aus dieser Strategie resultiert, sehr begrenzt. Diese Strategie wird in dieser Arbeit nicht weiter verfolgt.

7.5.2.3 Restguthabenverwaltung im Zugangsrouten

Prinzipiell könnten die *Credits* vom AAA.f-Server weiter an den Zugangsrouten delegiert werden. Grundsätzlich ist diese Strategie mit der Strategie aus Kapitel 7.5.2.2 kombinierbar. Es könnte z.B. ein größerer „Kreditgutschein“ vom AAA.h an den AAA.f (und somit an die auswärtige Domäne) delegiert werden. Im AAA.f könnten dann kleinere „Kreditgutscheine“ weiter an den Zugangsrouten delegiert werden, welche diese dann direkt verwalten.

Der Vorteil dieses grundsätzlichen Verfahrens liegt in der Möglichkeit, die Nutzungsdaten bei deren Entstehung benutzerspezifisch konsolidieren zu können. Hier kann sowohl die Netzlast, als auch das Nutzungsdatenaufkommen minimiert werden. Weiter ist der erforderliche Ver-

kehr auf dem Netz nicht mehr zeitkritisch und kann vom Betreiber auf ein geeignetes Zeitfenster verschoben werden. Je nach Dimensionierung der *Credits* kann ein Datenreduktionsfaktor erreicht werden, welcher prinzipiell in seiner Effektivität und somit in seinem Beitrag zur Reduktion der zu erfassenden Nutzungsdaten mit einer Größenordnung beiträgt, welche mit den Prepaid-Benutzern vergleichbar ist.

Diese Strategie wurde im Modell validiert und im Gesamtsystem verifiziert. Die Ergebnisse werden im folgenden Kapitel dargestellt und diskutiert.

7.6 Analyse der Ergebnisse

7.6.1 Fokus der Analyse

Die durchgeführte Modellierung verfolgt zwei Ziele, ein „statisches Ziel“ und ein „dynamisches Ziel“. Das „statische Ziel“ fokussiert sich auf die formale Spezifikation ausgewählter Komponenten aus Kapitel 6, womit eine semantische und syntaktische Überprüfung der Spezifikation erfolgen soll, welche weiter mittels statischer Tests das Gesamtmodell verifiziert. Die einzelnen Tests wurden fehlerfrei durchlaufen und sind im Anhang B beschrieben.

Das „dynamische Ziel“ der Modellierung besteht in der quantitativen Analyse des Nutzungsdatenaufkommens basierend auf den vorgestellten Vorschlägen, welches bei angenommenem Verkehrsverhalten im System erzeugt wird. Da das Datenaufkommen der Nutzungsdaten ausschliesslich von einem Benutzer abhängig ist, fokussiert sich die Betrachtung auf einen Benutzer. Skalierbarkeitsbetrachtungen wurden hierzu in Kapitel 6 durchgeführt.

Mit dieser Analyse wird eine Datenbasis geschaffen, welche es erlaubt, den Zusammenhang zwischen der absoluten Anzahl der angebotenen Dienste und der Granularität des Abrechnungssystems für verschiedene Bewegungsverhalten und Quellverhalten der mobilen Teilnehmer auf Basis der Szenarien Registrierung (siehe Kapitel 6.2.1.2 bzw. Kapitel 6.2.1.3) und Netzübergang nach Kapitel 6.2.1.4 zu ermitteln.

7.6.2 Validierung der Spezifikation und Durchführung der Leistungsbewertung

Ziel der Leistungsbewertung ist die Minimierung der Transaktionskosten eines Netzbetreibers im Mobil Internet, welche in dieser Arbeit durch die Parameter *Daten/Verkehrsaufkommen* beim Prozess der Nutzungsdatenerfassung repräsentiert werden, um entsprechende Erkenntnisse und Handreichungen für zukünftige Betreiber eines Mobil Internet bereitzustellen zu können.

7.6.2.1 Validierung und qualitative Leistungsbewertung des Gesamtkonzeptes

Um die Spezifikation in einem aus dem Modell abgeleiteten Simulator in einem ersten Schritt zu validieren, wurden vier verschiedene Wege durch das Modell nach Kapitel 7.3.3 definiert, welche vom Endgerät durchlaufen werden. Ziel dieser Wege ist das Abdecken aller repräsentativen Fälle. Diese Wege sind detailliert im Anhang B beschrieben und wurden für Benutzer aller drei Kategorien aus Kapitel 7.5.1 fehlerfrei durchlaufen. Die Auswahl der Wege wurde so gewählt, dass folgende Schlüsselszenarien abgedeckt sind:

- Registrierung in Auswärtiger Domäne und Übergang in Heimatdomäne
- Registrierung in Auswärtiger Domäne und Übergang in andere Auswärtige Domäne
- Registrierung in Heimatdomäne und Übergang in Auswärtige Domäne
- Registrierung in Heimatdomäne und Übergang innerhalb der Heimatdomäne
- Netzübergang zwischen Heimatdomäne und Auswärtiger Domäne

- Netzübergang zwischen zwei verschiedenen Auswärtigen Domänen

Das gewählte Teilnehmer- und Betreibermodell erfolgte nach dem Modell aus Kapitel 7.3.2 und bzgl. des Lastmodells wurde in Anlehnung an die Überlegungen aus Kapitel 7.3.5 eine konstante Netzlast angenommen. Die Intensität spielt in der weiteren Analyse keine Rolle.

Neben einer fehlerfreien Syntax, welche von TAU implizit eingefordert wird, konnte mittels dieser Vorgehensweise das System für diese relevanten Szenarien validiert werden. Ausgewählte Zustandsmaschinen sind in Anhang C dargestellt.

7.6.2.2 Dynamische Validierung der Spezifikation

In Fortsetzung an dieses Vorgehen wurde das Endsystem mittels des Random-Walk-Modells nach Kapitel 7.3.4 durch den „Manhattan-Grid“ bewegt. Hier wurde ein Benutzer mit Prepaid-Vertrag und beschränkten Budget und ein Benutzer mit Festvertrag simuliert. Analog zu Kapitel 7.6.2.1 wurden das Teilnehmer und Betreibermodell nach Kapitel 7.3.2, das Lastmodell aus Kapitel 7.3.5 und das Vertragsmodell aus Kapitel 7.3.6 angewendet. Bei mehrmaligem Durchlaufen konnte somit die spezifizierte Architektur validiert werden. Das System befand sich während ca. 1000 Netzübergängen nie in einem undefinierten Zustand.

7.6.2.3 Diskussion der Konzepte für Postpaid-Benutzer

Um das Datenaufkommen der Nutzungsdaten bei Postpaid im Kontext dieser Arbeit entsprechend zu reduzieren, ist der Reduktionsfaktor und somit eine Komprimierung analog zu dem Überlegungen aus Kapitel 7.5.2 an den drei Stellen Zugangsroutern, AAA.f und AAA.h möglich. Dies ist schematisch in Abbildung 62 dargestellt. Hier sei T das Intervall, in dem die erfassten Rohdaten aus dem Meter exportiert werden (siehe hierzu Abbildung 55). Bei diesen Benutzern ergibt sich folgender Datenreduktionsfaktor pro Zeiteinheit:

- Nutzungsdanenaufkommen im Zugangsroutern (AR): T
- Reduktionsintervall im Zugangsroutern (AR): T_{AR}
- Reduktionsintervall im AAA.f: $T_{AAA.f}$

Somit reduziert sich bei Umsetzung der im Rahmen dieser Arbeit gemachten Vorschläge das Datenaufkommen im AAA.f um den Faktor T_{AR}/T und zusätzlich im AAA.h um den Faktor $T_{AAA.f}/T_{AR}$ pro Benutzer und separat abzurechnenden Dienst. Grundsätzlich kann der Faktor $T_{AAA.f}/T$ auch voll im Zugangsroutern realisiert werden. Daraus folgt jedoch nicht, dass ein Zugangsroutern diese Information direkt zum AAA.h kommuniziert.

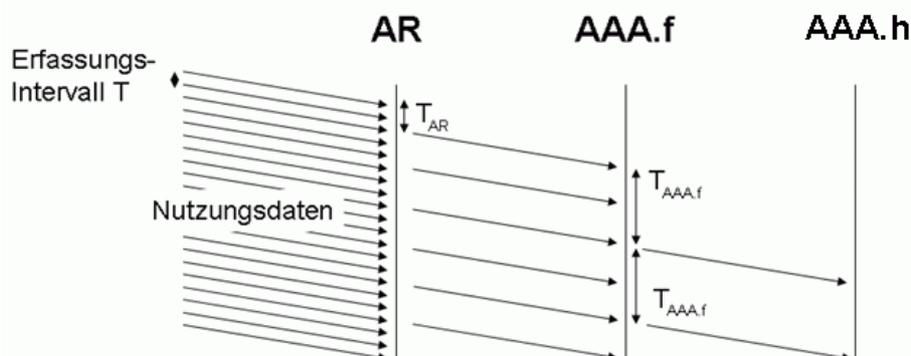


Abbildung 62: Komprimierung der Nutzungsdatenerfassung bei Postpaid-Benutzern

Da keine zwingende Notwendigkeit besteht, die Nutzungsdaten in Echtzeit weiterzuvermitteln, kann der Komprimierungsgrad aus Sicht eines Netzbetreibers frei dimensioniert werden. Der hieraus resultierende Effizienzgewinn lässt sich leicht ermitteln. In der Praxis wird ein Netzbetreiber jedoch als Obergrenze die Intervalle der Rechnungserstellung festlegen müssen,

welche in der Regel monatlich erfolgt. Wie nun ein „auswärtiger Betreiber“ die Daten zwischen AR und AAA.f konsolidiert, muss dem Betreiber der Heimatdomäne nicht kommuniziert werden. Hier ist lediglich zu erwarten, dass die Intervalle $T_{AAA.f}$ im Rahmen eines SLAs als Bestandteil des Roaming-Abkommens definiert werden.

Mengengesteuerte Datenkonsolidierung

Bei dieser Methode wird der Zugangsrouten so konfiguriert, dass das Intervall T_{AR} variabel an eine Datenmenge geknüpft wird. Diese Methode ist prinzipiell leicht zu implementieren und erlaubt eine benutzerspezifische Datenreduktion des Nutzungsdatenaufkommens im AAA.f. Weiter kann diese Methode problemlos auf das Intervall $T_{AAA.f}$ ausgeweitet werden.

Eine ausschließlich mengengesteuerte Vorgehensweise bringt das Problem mit sich, dass bei Benutzern mit sehr geringem Verkehrsaufkommen die definierte Nutzungsdatenmenge nie erreicht wird. Somit ist diese Methode in jedem Fall mit einer zeitgesteuerten Vorgehensweise zu kombinieren.

Grundsätzlich kann der Komprimierungsgrad basierend auf 4 verschiedenen Vorschlägen definiert werden, auf die nun kurz eingegangen wird. Generell ist in diesem Zusammenhang festzuhalten, dass die Möglichkeit der Datenkonsolidierung grundsätzlich vom Sitzungskonzept abhängig ist und der Ort der Datenkonsolidierung hier aus Sicht des Nutzungsdatenaufkommens eine untergeordnete Rolle spielt, jedoch eine Lösung, welche die Datenkonsolidierung an einem zentralen Punkt durchführt, mehr Netzlast erzeugt und einen höheren Verarbeitungsaufwand am zentralen Ort der Datenkonsolidierung erfordert.

Zeitgesteuerte Datenkonsolidierung

Bei einer zeitgesteuerten Datenkonsolidierung ist sichergestellt, dass die Nutzungsdatenberichte vorhersagbar zur jeweils nächsten Instanz weitergeleitet werden. Das „Always-On“ Paradigma, des Internets führt dazu, dass Benutzer zwar registriert, aber ggf. sehr lange Zeit inaktiv sind. Eine rein zeitgesteuerte Datenkonsolidierung würde unnötig Nutzungsdatenberichte generieren.

Ereignisgesteuerte Datenkonsolidierung

Bei der ereignisgesteuerten Datenkonsolidierung würde ein Nutzungsdatenbericht bei Verlassen einer administrativen Domäne erzeugt werden.

Der Vorteil dieses Vorgehens besteht darin, dass der Prozess der Nutzungsdatenerfassung aller sich in einer auswärtigen Domäne befindlichen Benutzer effizient vollzogen wird. Da davon auszugehen ist, dass die Netzbetreiber eine maximierte Flächenabdeckung anstreben, würde diese Maßnahme vorwiegend Netzbetreibern zugute kommen, welche über eine regional begrenzte Netzabdeckung verfügen und deren Benutzer sich vorwiegend in auswärtigen Netzen bewegen.

Empfohlenes Vorgehen

Der nun vorgestellte Algorithmus (Abbildung 63) soll basierend auf der vorherigen Diskussion die optimierte Datenkonsolidierung beschreiben. Da die aufbereiteten Nutzungsdatenberichte prinzipiell an den AAA.h-Server gesendet werden müssen, und die Registrierung (sowohl Mobile IP, als auch AAA nach Kapitel 4.3) an ein Zeitintervall gebunden ist, ist zunächst zu prüfen, inwieweit diese bereits spezifizierte Signalisierung erweitert werden kann. Als Redundationsfaktor (vergl. Tabelle 14) soll mindestens ein Wert von 100 angestrebt werden. Dies würde zu einem Neuregistrierungsintervall von 100 Sekunden führen. Hier ist

der im Projekt Moby Dick vorgeschlagene Wert von 5 Minuten akzeptabel, kann jedoch von den Netzbetreibern benutzerspezifisch und variabel festgelegt werden.

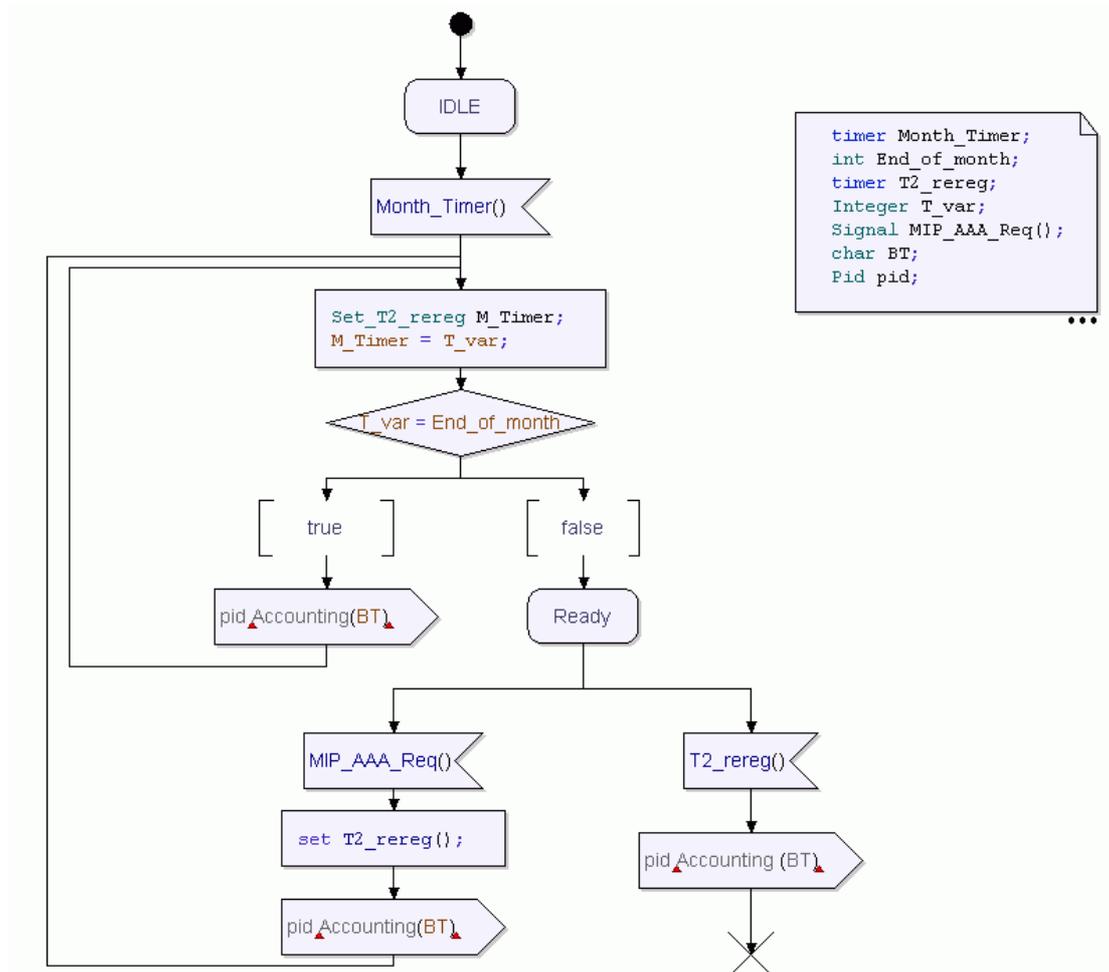


Abbildung 63: Vorgeschlagenes Vorgehen für die Datenkonsolidierung bei Postpaid-Benutzern

7.6.2.4 Diskussion der Konzepte für Prepaid-Benutzer

Nach Kapitel 7.5.2.3 fokussiert sich die hier folgende Analyse auf eine Kreditverwaltung im Zugangsrouter. Bei der Umsetzung dieses Konzeptes sind drei Prozesse zu betrachten. Dies sind die Verwaltung des Restbudgets eines Benutzers, welches einen in Echtzeit stattfindenden Abgleich mit den Nutzungsdaten erfordert, das Weiterreichen der verbleibenden *Credits* an den neuen Zugangsrouter und das Anfordern neuer *Credits* im Falle eines zur Neige gehenden Restguthabens.

Die Verwaltung des Restbudgets wird durch die Architektur aus Kapitel 6.4 unterstützt und hat den Vorteil, dass die verteilt erfassten Daten nicht mehr in Echtzeit weiterverarbeitet werden müssen. Im Falle eines Netzzellenwechsels innerhalb eines administrativen Bereiches ist das Weiterleiten des Restguthabens einfach in die FHO-Prozedur nach Kapitel 6.2.1.3 integrierbar. Das Anfordern neuer *Credits* in der Heimatdomäne kann einfach mittels Schwellwerten gesteuert werden. Eine negative Antwort von der Heimatdomäne im Falle eines totalen Konsums des verfügbaren Guthabens muss jedoch weiterhin in Echtzeit an den Zugangsrouter gesendet werden, jedoch ist der hierfür erforderliche Aufwand aus Sicht des Datenvolumens bzw. der Netzlast vernachlässigbar.

Somit ist im Folgenden das detaillierte Verfahren bei einem Netzübergang zwischen Betreibern näher zu betrachten. Bei diesem Szenario ist zwischen drei verschiedenen Fällen zu unterscheiden, welche den ersten drei Wegen aus Kapitel 7.6.2.1 bzw. Anhang B2 entsprechen. Der Übergang zwischen zwei „auswärtigen“ Domänen stellt den allgemeinsten Fall dar, welcher die Anforderungen für die beiden anderen Fälle mit berücksichtigt. Dieser Fall wird im Folgenden detaillierter betrachtet.

Abbildung 64 beschreibt diesen allgemeinen Fall, bei dem aus organisatorischen Gründen (vergl. Kapitel 6.2.1.4) das Restbudget nicht direkt von der Domäne *B* an die Domäne *C* übermittelt werden kann. Es ist eine indirekte Übermittlung des Restbudgets über die Heimatdomäne *A* erforderlich. Weiter kann bei diesem Prozess nicht davon ausgegangen werden, dass der Domäne, welche den Prozess auslösen könnte (Domäne *B*) das Ziel (Domäne *C*) bekannt ist. Implizit wird eine solche Abmeldung durch das Neuregistrierungsintervall nach Kapitel 2.1.4.1 realisiert. Diese Neuregistrierung erfolgt jedoch transparent zur Neuregistrierung nach Kapitel 4.4.

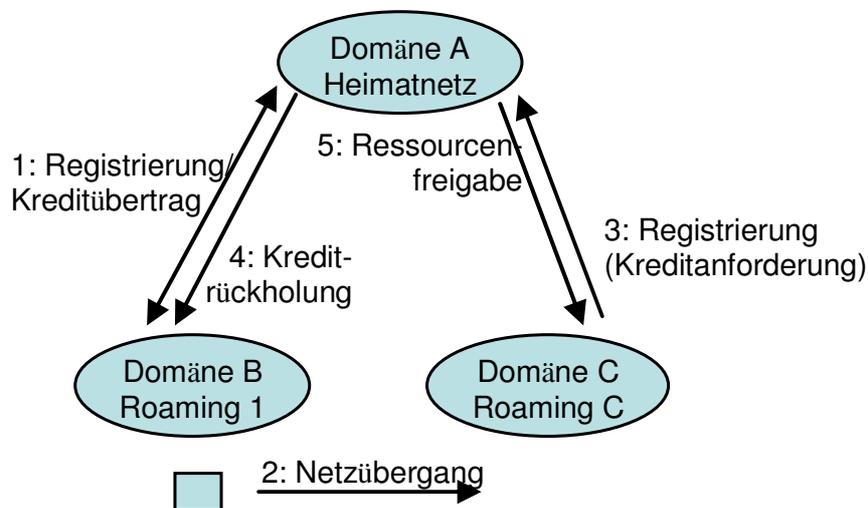


Abbildung 64: Delegationsproblem im Prepaid-Fall

Die generelle Frage hierbei ist, ob die *Credits* eines Benutzers, welcher eine administrative Domäne verlassen hat, diese bei Erkennung dieses Vorgangs zurückgibt, oder von der Heimatdomäne zurückgefordert werden. Wegen der nichtvorhandenen Abmeldeprozedur bei Mobile IP wird folgendes Vorgehen vorgeschlagen:

- 1.) Die *Credits* verbleiben in der auswärtigen Domäne, solange das Endsystem dort registriert ist.
- 2.) Mit Ablauf des Timers für die Neuregistrierung werden die *Credits* an die Heimatdomäne zurückgegeben.
- 3.) Im Falle eines Netzübergangs nach Kapitel 6.2.1.4 werden die verbleibenden *Credits* von der Heimatdomäne bei der auswärtigen Domäne, bei der das Endsystem vor dem Netzübergang angemeldet war, angefordert.

Dieser Algorithmus erfordert im Wesentlichen zwei neue Mechanismen. Dies ist zunächst im AAA.h/HA die Information bzgl. der Domäne, bei der das Endsystem registriert ist. Weiter muss der AAA-Manager aus 6.4.3 das *Credit-Management* durchführen.

Dieser Algorithmus soll nun basierend auf der Validierung von Weg 2 aus Anhang B2 (vergl. Abbildung 65) näher untersucht werden. Bei diesem Weg kommt es im Prinzip zu einem permanenten Netzübergang zwischen zwei „auswärtigen“ Domänen. Die hier vorgestellten

Vorschläge führen zu einem erhöhten Signalisierungsaufwand, da die Credits permanent über die Heimatdomäne in der jeweils anderen auswärtigen Domäne angefordert werden müssten.

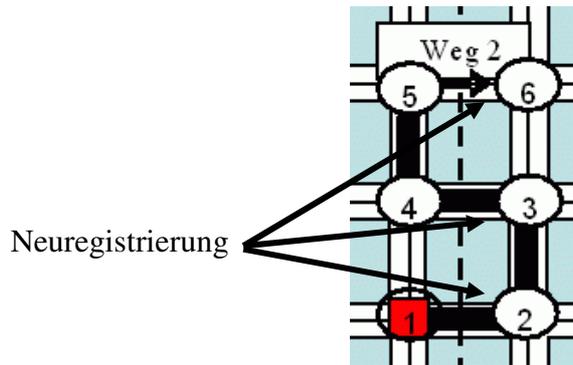


Abbildung 65: Kreditverwaltung bei Weg 2 und Delegation

Dieser Signalisierungsaufwand drückt sich in einer zeitlichen Verzögerung aus, welche dazu führt, dass dieses Szenario basierend auf aktueller Mobile IP-Technologie voraussichtlich nicht unterbrechungsfrei vollzogen werden kann.

Es sei nun die Frage aufgeworfen, inwieweit dieser Nachteil wirklich das Gesamtkonzept nachhaltig beeinträchtigt, da dieses Szenario aus Sicht eines Netzbetreibers ohnehin nicht aktiv unterstützt werden sollte.

7.6.3 Quantitative Reduktion der Nutzungsdaten

Die hier vorgestellten Vorschläge führen zu einer Reduktion der Nutzungsdaten nach Gleichung (3).

$$\text{Nutzungsdatenaufkommen} = \sum_{i=1}^{\text{Prepaid}} \frac{\text{Nutzungsdatensatz}(i)}{K \text{Pre}(i)} + \sum_{j=1}^{\text{Flat}} 0 + \sum_{k=1}^{\text{PostPaid}} \frac{\text{Nutzungsdatensatz}(k)}{K \text{Post}(k)} \quad (3)$$

Dies führt zu einem normierten *Reduktionsfaktor_N* der Nutzungsdatenmenge gegenüber Gleichung (1) aus Kapitel 7.4.1

$$\text{Reduktionsfaktor}_{N} = \frac{\sum_{i=1}^{\text{Prepaid}} \frac{1}{K \text{Pre}(i)} + \sum_{i=k}^{\text{Postpaid}} \frac{1}{K \text{Post}(k)}}{100} \quad (4)$$

Hiermit ergibt sich basierend auf Tabelle 13 und vereinheitlichtem Reduktionsfaktor K eine konsolidierte Reduktion der entstehenden Nutzungsdaten nach Tabelle 15. In diesem Fall ist die Verteilung zwischen Prepaid und Postpaid-Kunden ohne weiteren Einfluss auf das Nutzungsdatenaufkommen.

Zieht man basierend auf diesen Ergebnissen die Beispielrechnung aus Kapitel 7.4.2 heran, so reduziert sich das Nutzungsdatenaufkommen im Beispiel um den Faktor 15, was zu einer Datenmenge von ca. 25 Terrabyte pro AAA-Server entspricht. Dies würde die Referenz-Datenübertragung aus Kapitel 7.4.2 von 4 Tagen auf 6 Stunden reduzieren.

Generell wurde bei dieser Betrachtung angenommen, dass der Reduktionsfaktor $K \text{Pre}(i)$ voll zur Geltung kommt, d.h. der Benutzer des mobilen Endsystems nach minimal 100 Sekunden den Netzbetreiber wechselt. Diese Annahme kann im Rahmen dieser Arbeit nicht formal bestätigt werden, sollte aber in der Praxis das hier vorgestellte Konzept nicht in Frage stellen, da davon auszugehen ist, dass ein Netzbetreiber grundsätzlich seine Klientel solange als möglich

in seinem Netz behalten will und im Falle einer lückenhaften Netzabdeckung (wie im Prinzip im Modell angenommen) eine Föderierung mit anderen Netzbetreibern eingehen wird.

Tabelle 15: Quantitative Nutzungsdatenreduktion

Anteil Prepaid [%]	Anteil Postpaid [%]	Anteil Flat [%]	Reduktionsfaktor K	Datenreduktion gegenüber (1) [%]
25	25	50	1	50
25	25	50	2	75
25	25	50	100	2
25	25	50	∞	0

Weiter soll an dieser Stelle festgehalten werden, dass ein Datenreduktionsfaktor von 100 im Falle der Benutzer mit Prepaid-Vertrag grundsätzlich sehr einfach zu realisieren ist. Nimmt man hier einen „Always-On“-Szenario an und ein Benutzer, welcher sich permanent in der Heimatdomäne aufhält, so würden bei Umsetzung der vorgestellten Vorschläge Prepaid-Benutzer prinzipiell keinerlei Nutzungsdaten generieren, was faktisch dem Flat-Rate Fall entspricht. Dies würde aufkommenden die Nutzungsdaten aus Tabelle 15 nochmals erheblich reduzieren.

Die hier gemachten Vorschläge würden gegenüber der Architektur aus Kapitel 6 noch die Anforderung beisteuern, dass die Heimatdomäne jeweils über konkrete Informationen über den genauen Aufenthaltsort des Benutzers hat. Der Informationsgehalt einer CoA ist hierfür nicht ausreichend. Hier sollte administrative Informationen hinzugefügt werden, wie z.B. Domänen ID, und entsprechender AAA.f. Man könnte zu dieser Forderung optional noch detailliertere Informationen wie z.B. Zugangsrouten und dessen topologische Lage hinzufügen, jedoch gibt es organisatorische Argumente, welche ggf. gegen diesen Schritt sprechen, da ein Roaming-Partner ggf. kein strategisches Interesse hat, das detaillierte Bewegungsverhalten der Benutzer in seinem Netz Konkurrenten zugänglich zu machen. Dies wird in jedem Fall Bestandteil eines Roaming-Abkommens sein.

8 Zusammenfassung und Ausblick

Die fundamentalen Unterschiede zwischen historischer Sprachkommunikation in Verbindung mit leitungsvermittelnden Kommunikationsprinzipien und den paketbasierten Datennetzen sind so groß, dass eine Migration dieser beiden Netze in seiner Ganzheit bislang nicht ausreichend durchdacht wurde.

Im Rahmen dieser Arbeit wurde ein Aspekt, der der Nutzungsdatenerfassung zum Zwecke der Rechnungserstellung, in einem Mobil Internet eingehend betrachtet. Ausgehend von den technologischen Grundlagen der Kommunikationsnetze wurden die fundamentalen Unterschiede dieser beiden Kommunikationsprinzipien (Leitungsvermittlung und Paketvermittlung) herausgearbeitet. Dies sind die Art der Ressourcenallokierung, welche bei leitungsvermittelnden Netzen auf Reservationsprinzipien beruht, und bei den Datennetzen auf Basis eines verbindungslosen und konkurrierenden Ressourcenzugriffs. Weiter ist bei den Sprachnetzen eine Trennung zwischen Daten und Signalisierung vorzufinden. Speziell die Signalisierung ist bei leitungsvermittelnden Netzen sehr mächtig und nicht direkt mit den Signalisierungsmechanismen in den Datennetzen zu vergleichen, welche teilweise im IP-Header und teilweise über weitere Protokolle realisiert wird.

Neben diesen rein technischen Unterschieden, welche eine Migration dieser beiden Netze beeinflussen, sind als weiterer wesentlicher Aspekt dieser Migration die Gegebenheiten in einem kommerziellen und offenen Marktumfeld zu betrachten, welche in Kapitel 3 herausgearbeitet wurden. Auch hier sind fundamentale Unterschiede zwischen den beiden erwähnten Netzen vorzufinden. Weiter sind bestimmte Vorgehensweisen aus marketingpolitischer Sicht wie z.B. Preisdifferenzierung und Dienstbündelung in Märkten etabliert, welche dazu führen, dass benutzerindividuell Dienste angeboten und abgerechnet werden. Weiter wird in ein dynamisches Marktumfeld eine Flexibilität bzgl. der Fähigkeit neue, oft auch zeitlich begrenzte Preismodelle und Dienstbündel zu implementieren, gefordert, welche aktuelle Datennetze in ihrer aktuellen Architektur nicht einfach umsetzen können.

Nach dem Erarbeiten der technischen und ökonomischen Rahmenbedingungen wurde ein Überblick über ausgewählte Mechanismen, Konzepte und Architekturen gegeben, welche potentiell den Migrationsprozess zum kommerzialisierten Mobil Internet stimulieren. Hier wurde ausgehend von der Moby Dick-Architektur ein Überblick über Konzepte gegeben, welche diese gewählte Referenzarchitektur geeignet in Richtung eines kommerziellen Mobil Internet erweitern können. So bieten sich innerhalb der IETF vor allen die Mechanismen und Konzepte im erweiterten Umfeld der RTFM/AAA-Arbeitsgruppen an, einen wesentlichen Beitrag zu einer effektiven Nutzungsdatenerfassung zum Zweck der späteren Rechnungserstellung zu leisten. Nach genauerer Analyse dieser Mechanismen und Konzepte wurden offene Fragen und Schwachstellen aufgedeckt, welche die aktuell observierte Migration hin zu einem reinen IP-basierten Mobilkommunikationsnetz, dem Mobil Internet, aufhalten könnten. Konkret bringen die Echtzeitanforderungen während eines Netzübergangs, sowohl innerhalb eines administrativen Netzes, als auch zwischen administrativen Netzen, einen erhöhten Aufwand an Signalisierung und höhere zeitkritische Anforderungen mit sich, welche aktuell nicht ausreichend befriedigt werden. Diese zeitkritischen Anforderungen werden zusätzlich durch den Druck, verschiedene Preismodelle zu unterstützen und gleichzeitig den Abrechnungsprozess kostenminimiert durchzuführen, verstärkt.

Dies führt zum Zusammenhang zwischen einem implementierten Preismodell und dem dafür erforderlichen Gesamtaufwand der Nutzungsdatenerfassung samt nachgelagerter Weiterverarbeitung der erfassten Nutzungsdaten für die Rechnungserstellung. Im weiteren Verlauf

dieser Arbeit wurde ein exemplarisches Preismodell für ein zukünftiges kommerzielles Mobiles Internet entwickelt, anhand dessen die weiteren Optimierungsvorschläge geeignet bewertet werden können.

Der Kernansatz hierbei folgt aus der Feststellung, dass in einem kommerziellen Netz eine benutzerorientierte Sichtweise bzgl. der Nutzungsdatenerfassung implementiert sein muss, um die entsprechenden Modelle aus der Preispolitik kostenminimiert anwenden zu können. Diese benutzerorientierte Sichtweise ist bislang im Internet nicht ausreichend berücksichtigt.

Aus der abstrakten Forderung einer benutzerorientierten Nutzungsdatenerfassung, welche die nachgelagerten Prozesse entsprechend kostenminimiert unterstützt, wurden auf Basis der Szenarien *Registrierung* und *Netzübergang* drei Kernforderungen aufgestellt. Dies sind eine Netzknotenarchitektur, ein Sitzungskonzept und eine Signalisierungsarchitektur, welche in optimierter Zusammenarbeit untereinander im Rahmen eines Gesamtkonzeptes die verteilt entstehenden Nutzungsdaten verteilt erfasst und so aufbereitet, dass die anschließende Abrechnung kostenminimiert in einem aus Benutzersicht förderierten Netz implementiert werden kann.

Im weiteren Verlauf der Arbeit wurden die oben erwähnten drei Kernforderungen spezifiziert und validiert. Hierzu wurden mittels SDL und MSC formale Methoden verwendet, welche bislang im Internet zur Spezifikation und Validierung nicht großflächig eingesetzt wurden, jedoch in der Telekommunikationsbranche sehr weit verbreitet sind. Beide Methoden wurden so integriert, dass die Spezifikation zunächst automatisch auf Konsistenz geprüft und dann basierend auf definierten Testzyklen die Interaktion zwischen den Komponenten dynamisch mittels MSCs validiert wird. Bei der Spezifikation wurde darauf geachtet, dass die neuen Mechanismen nahtlos in die aktuelle Mobile Internet-Architektur auf Basis der gewählten Referenzarchitektur integriert werden können und prinzipiell mit den Lösungen der IETF kompatibel sind.

Nach Validierung und Verifikation des Gesamtsystems wurden einige Parametrisierungen und Optimierungen diskutiert, welche zu konkreten Empfehlungen an kommerzielle Betreiber eines Mobiles Internets führen.

Am Ende dieser Arbeit steht zunächst auf abstrakter Ebene ein Beitrag zur Bewusstseinssteigerung im Mobilem Internet, dass der Aspekt Abrechenbarkeit eines Dienstes bislang völlig vernachlässigt wurde. Weiter ist eine methodische Vorgehensweise aufgezeigt worden, welche den Betreiber einer Dienstinfrastruktur bei der Definition seiner Dienstbündel dahingehend unterstützt, dass die aus der Dienstkomponierung resultierenden Anforderungen an die Abrechnungsinfrastruktur vorab abgeschätzt werden können. Dies ermöglicht einem Betreiber die angebotenen Dienste in ihrer Komposition auf Basis einer gegebenen Leistungsfähigkeit der Abrechnungsinfrastruktur optimieren zu können. Die validierte und mittels einiger Modelle verifizierte Spezifikation in Kombination mit den Skalierungsbetrachtungen geben hier umfangreiche Handreichungen bei der Dimensionierung der Abrechnungsinfrastruktur in einem kommerziellen Mobilem Internet.

8.1 Zusammenfassung der Ergebnisse

Im Kapitel 6 wurden drei Kernforderungen aufgestellt, diese sind: Signalisierungsarchitektur, Sitzungskonzept und Knotenarchitektur für einen Zugangsrouten.

Im Rahmen der Spezifikation wurde nachgewiesen, dass die im Kapitel 6 vorgestellten Mechanismen basierend auf den getroffenen Annahmen qualitativ umsetzbar sind. Die Validie-

rung erfolgte anhand ausgewählter Wege, welche die Nachvollziehbarkeit dieses Ergebnis unterlegen. Weiter wurde mittels des erfolgreich durchgeführten Random-Walk-Durchlaufes keinerlei Anhaltspunkte gefunden, welche eine Schwachstelle in der Spezifikation vermuten lassen.

Die in Kapitel 6 und 7 angestellten Überlegungen bzgl. der Skalierbarkeit der Gesamtlösung zeigen, dass diese Gesamtlösung wegen der verteilten Architektur in großen Netzen effektiv zum Einsatz kommen könnte.

Als weiteres Ergebnis dieser Arbeit lässt sich festhalten, dass eine benutzerorientierte Abrechnungsinfrastruktur ermöglicht, benutzerspezifische Abrechnungsmechanismen zu implementieren, welche an die benutzerindividuellen Preismodelle angepasst sind. Dies reduziert den Aufwand und somit die Kosten des Abrechnungsprozesses.

Es wurde aufgezeigt, dass die drei grundsätzlichen Abrechnungsschemata (Prepaid, Postpaid und Flat-Rate) unterschiedliche Anforderungen mit sich bringen, welche bei einem grundsätzlich benutzerspezifischen Vorgehen effektiv in einem Netz implementiert werden können.

Als Kernergebnis bei der Betrachtung der Flat-Rate-Benutzer ist festzuhalten, dass die vorgestellte Architektur die Nutzungsdatenerfassung bei Flat-Rate-Benutzern unterdrücken kann, d.h. die Entstehung von Nutzungsdaten wird grundsätzlich vermieden, was als Folge keinen komplexen und rechenintensiven Prozess bei der Aufbereitung der Nutzungsdaten erfordert.

Als Kernergebnis bezüglich des Abrechnungsverfahrens bei Postpaid-Benutzern ist festzuhalten, dass eine frühe Komprimierung der Nutzungsdaten (also bei deren Entstehung) den Gesamtaufwand signifikant reduziert. Diese Komprimierung erfolgt idealerweise im Zugangsrouten.

Als Kernergebnis bezüglich des Abrechnungsverfahrens bei Prepaid-Benutzern wird eine verteilte Kreditverwaltung vorgeschlagen, welche die *Credits* an das auswärtige Netz delegiert. Dort ist eine Weiterreichung der *Credits* zum Zugangsrouten vorteilhaft. Dieser Vorschlag erlaubt eine Integration der Signalisierung in die FHO-Prozeduren der Mobilitätsverwaltung, hat jedoch leichte Nachteile beim Übergang in ein anderes administratives Netz. Dieser rein technische Nachteil ist jedoch aus Marktsicht eher als Vorteil zu bewerten. Es soll an dieser Stelle die Vermutung angestellt werden, dass durch regulatorischen Druck hier weitere Maßnahmen implementiert werden müssen, welche jedoch in dieser Arbeit nicht berücksichtigt wurden. Dies könnte dann ggf. eine verteilte Vorhaltung der *Credits* zwischen AR, AAA.h und AAA.f erfordern, um bestimmte regulatorisch geforderten SLAs beim Netzübergang zwischen zwei administrativen Netzen einhalten zu können.

8.2 Empfehlungen an Netzbetreiber

Selbstverständlich steht am Ende dieser Arbeit die Empfehlung an einen Netzbetreiber, die Abrechnungspolitiken benutzerspezifisch innerhalb seines Netzes zu organisieren. Im Falle von Roaming bzw. Föderierung sollte dieser Prozess entsprechend netzübergreifend sein. Die Spezifikation und der verteilte Ansatz der vorgestellten Systemarchitektur lassen dies in einem kommerziell betriebenen Netz zu. Grundsätzlich können die entsprechenden Parameter, mit welchen das Nutzungsdatenaufkommen reduziert werden bzw. die Kreditverwaltung effizienter gestaltet werden kann, individuell konfiguriert werden. Grundsätzlich sollte in Betracht gezogen werden, die verschiedenen Registrierungsvorgänge untereinander zu koppeln. Vor allem könnten durch Integration von Preismodellen und Netzprotokollen Synergieeffekte geschaffen werden. Konkret könnte man die verschiedenen Registrierungsvorgänge bzw. die

bereits vorhandenen Signalisierungsnachrichten auf den verschiedenen Ebenen konsolidieren. Konkret erfolgt auf Vermittlungsebene die Mobile IP-Registrierung (vergl. Kapitel 2.1.4.2), ggf. eine Neuregistrierung (vergl. Kapitel 4.4) und die Registrierung des Benutzers im Heimatnetz. Diese Vorgänge wurden bereits in Kapitel 6.2.1.2 integriert. Hier könnte man benutzerspezifisch und von jeweiligen Vertragsmodell abgeleitet diese Signalisierungsnachrichten um Parameter der Kreditverwaltung so erweitern, dass keine zusätzliche Verkehrslast auf dem Kernnetz erzeugt wird.

8.3 Ausblick

Als Ausblick auf mögliche weiterführende Arbeiten ist zunächst festzuhalten, dass im Rahmen dieser Arbeit angenommen wurde, dass die realen Nutzungsdaten erfasst werden sollen. Grundsätzlich ist es denkbar, über z.B. statistische Methoden nicht mehr die realen Daten in ihrer Gänze erfassen zu müssen, sondern die Nutzungsdatenerfassung entsprechend auf repräsentative Proben zu reduzieren. Grundsätzlich ist diese Maßnahme je nach regulatorischen Vorgaben ebenso denkbar und man könnte im Rahmen einer weiterführenden Arbeit untersuchen, ob und wie diese Maßnahmen mit der hier vorgestellten Systemarchitektur kombinierbar sind. Auch in diesem Fall würde ein benutzerorientiertes Vorgehen zu weiteren Verbesserungen führen, da ggf. verschiedene statistische Algorithmen auf die entsprechenden Benutzergruppen/Vertragsmodelle angepasst werden müssten.

Abschließend soll festgehalten werden, dass sich diese Arbeit auf Parameter welche sich ausschließlich auf Schicht 3 und Schicht 4 nach ISO/OSI beziehen. Folgt man dem aktuellen Trend hin zu höherwertigen Diensten, welche sich dann an Inhalten und Transaktionen orientieren, so ist aus heutiger Sicht zu erwarten, dass sich mittelfristig die Parameter auf Netz und Transportschicht mit Parametern aus höheren Schichten, welche dann Inhalte repräsentieren, vermischen werden [13]. Somit könnte in einer weiterführenden Arbeit untersucht werden, wie Sitzungskonzepte, welche z.B. Inhalte und komplexe Transaktionen mit einbeziehen, in eine solche Gesamtarchitektur integriert werden könnten. Hier stoßen die verfügbaren Nutzungsdatenerfassungskonzepte der IETF an Ihre Grenzen.

Generell ist in diesem Zusammenhang zu prüfen, wie die Nutzungsdatenerfassung mittelfristig von Parametern der Schicht 3 entkoppelt werden kann. In diesem Kontext soll auf die „HIP“-Aktivitäten innerhalb der IETF hingewiesen werden [94]. Somit könnte die Authentifizierung völlig von den IP-Adressen des Endsystems entkoppelt werden. In [33] wurde ein solcher Ansatz verfolgt.

Um die Abrechnung effektiver gestalten zu können, sollte dann die Heimadresse Bestandteil der Sitzungs-ID sein. Grundsätzlich könnte man diese derartig abändern, jedoch gibt es aktuell in der IETF Ansätze, die dieses Problem grundsätzlich adressieren. So ist der Ansatz, die Adressierung von der Verkehrslenkung grundsätzlich zu trennen, um die IP-Adresse ausschliesslich zur Wegführung zu verwenden und die Schicht 4 Sitzungen an einen entkoppelten Identifikator, den sogenannten HIT (Host Identity Tag) [93] zu koppeln, ein Vorhaben, welches auch die Effektivität des Abrechnungsprozesses wesentlich verbessern könnte. Konkret wird hier vorgeschlagen den HIT mit der Sitzungs-ID zu synchronisieren und somit die Registrierung in verschiedenen auswärtigen Netzen länger und zeitlich überlappend aufrechterhalten können. Dies hätte den Vorteil bzw. kann als technische Voraussetzungen formuliert werden, dass die Benutzer dynamischer zwischen verschiedenen administrativen Netzen ohne Einschränkung der Netzqualität wechseln können. Weiter wäre es einem Benutzer so möglich, sich gleichzeitig mit verschiedenen Anwendungen in verschiedenen Netzbereichen zu registrieren.

Literaturverzeichnis

- [1] 3GPP: Technical Specification Group RAN, 3G TS 25.301 V3.4.0, *Radio Interface Protocol Architecture*
- [2] 3GPP: Technical Specification Group RAN, 3G TS 25.302 V3.4.0, *Services Provided by the Physical Layer*
- [3] 3GPP: TS 23.127; *Virtual Home Environment (VHE) / Open Service Access (OSA)*; Stage 2.
- [4] 3GPP TS 29.078; *Customised Applications for Mobile network Enhanced Logic (CAMEL) Phase 3*; CAMEL Application Part (CAP) specification.
- [5] ABOBA, B.; VOLLBRECHT, J.; „*Proxy Chaining and Policy Implementation in Roaming*“, IETF RFC 2607, Juni 1999
- [6] ABOBA, B.; BEADLES, M.; „*The Network Access Identifier*“, IETF RFC 2486, Januar 1999
- [7] ABOBA, B.; CALHOUN, P.; GLASS, S.; PERKINS, C.; PATIL, B.; MITTON, D.; ZORN, G.: „*Criteria for Evaluating AAA Protocols for Network Access*“, IETF RFC 2989, November 2000
- [8] BAGNULO, M.; SOTO, I, GARCIAMARTINEZ, A.; AZCORRA, A.: „*Random generation of interface identifiers*“, draftsotomobileiprandomiids00.txt, Januar 2002
- [9] BLACK, D.; BLAKE, S.; CARLSON, M.; DAVIES, E.; WANG, Z., WEISS, W.: „*An Architecture for Differentiated Services*“, IETF RFC 2475, Dezember 1998
- [10] BLUNK, L.; VOLLBRECHT, J.: „*PPP Extensible Authentication Protocol (EAP)*“, IETF RFC 2284, März 1998
- [11] BRADEN, R.; CLARK, D.; SHENKER, S.: „*Integrated Services in the Internet Architecture: an Overview*“, IETF RFC 1633, Juni 1994
- [12] BRADEN, R. (Editor): „*Resource ReSerVation Protocol (RSVP)*“, IETF RFC 2205, September 1997
- [13] Braun, A. M.; *Qualitätsaspekte multimodaler Kommunikation*, Dissertation der ETH Zürich Nr. 15101, Zürich, 2003
- [14] BROWNLEE, N., FULTON, R.: „*Kawaihiko and the thirdquartile day*“; *IEEE Communications Magazine*, Vol 38, No. 8 , Aug 2000, pp. 162 –168
- [15] BROWNLEE, N.; „*Using NeTraMet for production traffic measurement*“, *Proceedings of IEEE/IFIP International Symposium on Integrated Network Management*, Seattle, 2001, pp. 213 –226
- [16] BROWNLEE, N: „*Network Management and realtime Traffic Flow. Measurement*“, *Journal of Network and Systems Management*, Vol. 6, No. 2, 1998, pp 223227
- [17] BROWNLEE, N.: „*New Zealand Experiences with Network Traffic Charging*“, *Journal of Electronic Publishing*, Vol 2, Issue 1, May 1996
- [18] BROWNLEE, N.: „*Internet Pricing in Practice*“, *Internet Economics*, MIT Press, 1995, pp 7790
- [19] BROWNLEE, N.; MILLS, C.; RUTH, G.: „*Traffic Flow Measurement: Architecture*“, IETF RFC 2722, Oktober 1999
- [20] BROWNLEE, N.: „*Traffic Flow Measurement: Experiences with NeTraMet*“, IETF RFC 2123, März 1997
- [21] BROWNLEE, N.: „*Traffic Flow Measurement: Meter MIB*“, IETF RFC 2720, Oktober 1999
- [22] BROWNLEE, N.; BLOUNT, A.: „*Accounting Attributes and Record Formats*“, IETF RFC2924, September 2000

- [23] BOOCH, G.; RUMBAUGH, J.; JACOBSON, I.: *The Unified Modeling Language User Guide*, AddisonWesley, 1999
- [24] CALHOUN, P.; ZORN, D.; SPENCE, D.; MITTON, D.: „Diameter Network Access Server Application“, IETF draftietf-aaa-diameter-nasreq17.txt, Juli 2004
- [25] CALHOUN, P.; LOUGHNEY, J.; GUTTMANN, E.; ZORN, G.; ARKKO, J.: „Diameter Base Protocol“, IETF RFC 3588, IETF, September 2003
- [26] CARLE, G.; ZANDER, S.; TSEBY, T.: „Policy-based Accounting“, IETF RFC3334, September 2002
- [27] CHARZINSKI, J.: „HTTP/TCP Connection and Flow Characteristics“, *ACM Performance Evaluation*, Vol. 42 No. 23, Sep. 2000, pp. 149162
- [28] Cisco Systems Internet Seite; http://www.cisco.com/warp/public/cc/pd/iosw/ioft/neflct/tech/napps_wp.htm, Aufrufdatum: 15.6.2005.
- [29] CLARK, D. D.: „Internet Cost Allocation and Pricing“, *Internet Economics*, MIT Press, pp 215278, 1995
- [30] COTTON, S. (Editor): „Network Data Management – Usage (NDMU) For IP-Based Services“, Version 3.1.1, *IPDR Konsortium*, Oktober 2002
- [31] COTTON, S. (Editor): „Network Data Management – Usage (NDMU) For IP-Based Services Service Specification – DOCSIS™1.1 Service Flow Metering“, Version 3.1 – A. 0, *IPDR Konsortium*, September 2002
- [32] COTTON, S. (Editor): „Network Data Management – Usage (NDMU) For IP-Based Services Service Specification – Voice over IP (VoIP)“, Version 3.1 – A0.2, *IPDR Konsortium*, August 2002
- [33] DAIDALOS: <http://www.istdaidalos.org>, Abrufdatum 1.1.2006
- [34] DARPA Internet Program Protocol Specification, „Internet Protocol“, IETF RFC791, September 1981.
- [35] DEERING, S.; HINDEN, R.: „Internet Protocol, Version 6 (IPv6) Specification“, IETF RFC 1883, Dezember 1995
- [36] DEERING, S.: „ICMP Router Discovery Messages“, IETF RFC 1256, September 1991
- [37] DEERING, S.; HINDEN, R.: „Internet Protocol, Version 6 (IPv6)“, IETF RFC 2460, December 1998
- [38] DEUTSCH, P.: „GZIP file format specification version 4.3“, IETF RFC 1952, Mai 1996
- [39] EINSIEDLER, H.; AGUIAR, R. L.; JAEHNERT, J.; JONAS, K.; LIEBSCH, M.; SCHMITZ, R.; PACYNA, P.; GOZDECKI, J.; PAPIR, Z.; MORENO, J. I.; SOTO, I.: „The Moby Dick Project: a Mobile Heterogenous AllIP Architecture“, *Proceedings of the Mobile Summit 2001*, Sitges, Spanien, September 2001
- [40] ETSI: European Telecommunication Standards Institute, www.etsi.org, Abrufdatum 1.1.2006
- [41] FANKHAUSER, G.; STILLER, B.; VÖGTLI, C.; PLATTNER, B.: „Reservationbased Charging in an Integrated Services Network“, *Proceedings of 4th INFORMS Telecommunications Conference*, Boca Raton, Florida, March 1998
- [42] FIELDING, R.; IRVINE, UC.; GETTYS, J.; MOGUL, J.; FRYSTYK, H.; MASINTER, L.; LEACH, P.; BERNERSLEE, T.: „Hypertext Transfer Protocol – HTTP/1.1“, IETF RFC 2616, Juni 1999
- [43] GARCIAMARTIN, M. (Editor): „Diameter Session Initiation Protocol (SIP) Application“, IETF draftietf-aaa-diameter-sippapp07, März 2005
- [44] GROSSMANN, D.: „New Terminology and Clarifications for Diffserv“, IETF RFC 3260, April 2002

- [45] GUPTA, A.; STAHL, D. O.; WHINSTON, A. B.: „A Priority Pricing Approach to Manage MultiService Class Networks in RealTime“, *Proceedings of MIT Workshop on Internet Economics*, März 1995
- [46] HAKALA, H.; MATTILA, L.; KOSKINEN, P.; STURA, M.; LOUGHNEY, J.; „*Diameter CreditControl Application*“, IETF RFC 4006, August 2005
- [47] HANDELMANN, S.; STIBLER, S.; BROWNLEE, N.; RUTH, G.: „*New Attributes for Traffic Flow Measurement*“, IETF RFC 2724, Oktober 1999
- [48] HANDLEY, M.; JACOBSON, V.: „*SDP: Session Description Protocol*“, IETF RFC2327, April 1998
- [49] HANDLEY, M.; SCHULZRINNE, H.; SCHOOLER, E.; ROSENBERG, J.: „*SIP: Session Initiation Protocol*“, IETF RFC2543, März 1999
- [50] HASAN; JAEHNERT, J; ZANDER, S.; STILLER, B.: „Authentication, Authorisation, Accounting and Charging for the Mobile Internet“, *ETH ZÜRICH, TIK Report No. 114*, Juni 2001
- [51] HEINANEN, J.; BAKER, F.; WEISS, W. WROCLAWSKI, J.: „*Assured Forwarding PHB Group*“, IETF RFC 2597, Juni 1999
- [52] HINDEN, R.; DEERING, S.: „*IP Version 6 Addressing Architecture*“, IETF RFC 1884, Dezember 1995
- [53] HINDEN, R.; DEERING, S.: „*IP Version 6 Addressing Architecture*“, IETF RFC 2373, Juli 1998
- [54] HINDEN, R.; O'DELL, M.; DEERING, S.: „*An IPv6 Aggregatable Global Unicast Address Format*“, IETF RFC 2374, Juli 1998
- [55] IANA: <http://www.iana.org>, Abrufdatum: 19.03.2005
- [56] IEEE: „Wireless LAN, MAC, and PHY Specifications,“ IEEE Document P802.11/D6.1.97/5, June 1997, AlphaGraphics #35, 10201 N. 35th Ave., Phoenix, AZ 85051
- [57] IETF: <http://www.ietf.org>, Abrufdatum: 19.03.2005
- [58] ISO/IEC 138881: *Information Technology Security techniques Nonrepudiation Part 1: General*; ISO/IEC, 1997.
- [59] ISO/IEC 138882: *Information Technology Security techniques Nonrepudiation Part 2: Mechanisms using symmetric techniques*; ISO/IEC, 1998
- [60] ISO/IEC 138883: *Information Technology Security techniques Nonrepudiation Part 3: Mechanisms using asymmetric techniques*; ISO/IEC, 1997
- [61] ISOC: <http://www.isoc.org>, Abrufdatum: 19.03.2005
- [62] ITUT SG 10: *Message Sequence Chart (MSC)*, Rec. Z.120, Geneva, November 1999
- [63] ITUT SG 10: *Specification and Description Language (SDL)*, Rec. Z 100, Geneva, November 1999
- [64] ITUT: *Specification of TMN Applications at the Q3 Interface: Call Detail Recording*, Recommendation Q.825, Genf, 1998
- [65] ITUT: *The international public telecommunication numbering plan* ; Recommendation E 164, Genf, 1997
- [66] JACOBSON, V.; NICHOLS, K.; PODURI, K.: „An Expedited Forwarding PHB“, IETF RFC 2598, Juni 1999
- [67] JÄHNERT, J.: „Costefficient metering and Accounting for 4G networks“, *Proceedings of the 18th International Teletraffic Congress*, Berlin, September 2003
- [68] JÄHNERT, J.: „Problem Statement: Metering and Accounting in the fullIP 4G environment“, *Proceedings of the third International Workshop on Internet Charging and QoS Technology*, Munich, September 2003

- [69] JÄHNERT, J.: *Preisstrategien für drahtlosen Internetzugang der nächsten Generation*, Diplomarbeit an der Fachhochschule für Technik und Wirtschaft, Berlin, Mai 2002
- [70] Jain Initiative: <http://java.sun.com/products/jain/>, Aufrufdatum 16.5.2005.
- [71] JOHNSON, D.; PERKINS, C.: "Mobility Support in IPv6," *Proceedings of ACM Mobicom 96*, ACM, Nov. 1996, pp. 2737.
- [72] KAUSAR, N.; BRISCOE, B.; CROWCROFT, J.: „A Charging Model for Sessions on the Internet“, *Proceedings of the Fourth IEEE Symposium on Computers and Communications*, Red Sea, Egypt, June 1999
- [73] KENT, S.; ATKINSON, R.: „*Security Architecture for the Internet Protocol*“, IETF RFC2401, November 1998.
- [74] KLOPFENSTEIN, B. C.: „Internet Economics: An Annotated Bibliography“, *Journal of Media Economics*, Vol. 11, No. 1, 1998, pp. 38 48
- [75] KOTLER, P.; BLIEMEL, F: *Marketing Management*, SchäfferPoeschel Verlag, 1999
- [76] KRISTOL, D.; MONTULLI, L.: „*HTTP State Management Mechanism*“, IETF RFC 2965, Oktober 2000
- [77] KURTANSKY, P.; STILLER, B.: „Time Intervalbased PrepaidCharging of QoS-enabled IP Services“, *Proceedings of the first Workshop on Internet and Network Economics*, Hongkong, Dezember 2005
- [78] DE LAAT, C.; GROSS, G.; GOMMANS, L.; VOLLBRECHT, J.; SPENCE, D.: „*Generic AAA Architecture*“, IETF RFC 2903, August 2000
- [79] LE, F.; BASAVARAJ, P.; PERKINS, C.; FACCIN, S.: „*Diameter Mobile IPv6 Application*“, IETF Internet draft draftleaaadiametermobileip604.txt, November 2004
- [80] LIANG, Y.; CHRIST, P.; JÄHNERT, J.; SARMA, A.; Melia, T.: „Taming Monsters like Daidalos“, *Proceedings of the ISSRE04 Workshop on Integrated reliability with Telecommunications and UML Languages*, Rennes, October 2004
- [81] LIEBSCH, M.; PEREZCOSTA, X.; SCHMITZ, R.; SARMA, A; JÄHNERT, J.; TESSIER, S.; WETTERWALD, M.; SOTO, I: „Solutions for IPv6based mobility in the EU project Moby Dick“, *Proceedings of the World Telecom Congress*, Paris, September 2002
- [82] LIOR, A.; YEGANI, P.; CHOWDHURY, K.; TSCHOFENIG, H.; PASHALIDIS, A.: „PrepaidExtensions to Remote Authentication DialIn User Service (RADIUS)“, IETF draftliorradiusprepaidextensions07.txt, Februar 2005
- [83] MACKIEMASON, J. K.; VARIAN, H. R.: “Economic FAQs About the Internet”, *Internet Economics*, MIT Press, 1995, pp 2762
- [84] MACKIEMASON, J. K.; MURPHY, L.; MURPHY, J.: „Responsive Pricing in the Internet“, *Internet Economics*, MIT Press, 1995, pp 279304
- [85] MACKIEMASON, J. K.; VARIAN, H. R.: „Pricing the Internet“, *Economics Working Paper Archive EconWPA*, 1995, pp. 269314
- [86] MACKIEMASON, J. K.: “A Smart Market for Resource Reservation in a Multiple Quality of Service Information Network”, Internal Report, *University of Michigan*, USA, September 1997
- [87] MACKIEMASON, J. K.; MURPHY, L.; MURPHY, J.: „Pricing Congestible Networks“, *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, 1995, pp. 11411149
- [88] MCKNIGHT, L. W.; BAILEY, J.P.: *Internet Economics*, *The MIT Press*, Cambridge, Mass.,1997
- [89] MEFFERT, H.: *Marketing*, Gabler Verlag, ISBN 3409690174, 9. Auflage, Oktober 2000

- [90] MILGROM, P.: "Auctions and Bidding: a Primer", *Journal of Economic Perspectives*, Vol. 3, Issue 3, 1989, pp. 322
- [91] MONTENEGRO, G.; GUPTA, V.: „Sun’s SKIP Firewall traversal for Mobile IP“, IETF RFC2356, Juni 1998
- [92] MOORE, K.; FREED, N.: „*Use of HTTP State Management*“, IETF RFC2964, Oktober 2000
- [93] MOSKOWITZ, R., NIKANDER, P.; JOKELA, P.; HENDERSON, T.: „Host Identity Protocol“, IETF draftietfhipbase04, Oktober 2005
- [94] MOSKOWITZ, R., NIKANDER, P.: „*Host Identity Protocol Architecture*“, IETF draftietfhiparch03, August 2005
- [95] NAIN, P.; TOWSLEY, D.; LIU, B.; LIU, Z.: „Properties of Random Direction Models“, *INRIA technical report RR5284*, July 2004
- [96] N.N., Information Science Institute California: „*Transmission Control Protocol*“, IETF RFC 793, September 1981
- [97] NICHOLS, K.; BLAKE, S.; BAKER, F.; BLACK, D.: „*Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers*“, IETF RFC 2474, Dezember 1998
- [98] ODLYZKO, A.: „A modest proposal for preventing Internet congestion“, *AT&T Labs Internal Report*, Florham Park, NJ, September 1997
- [99] ODLYZKO, A.: „The economics of the Internet: Utility, utilization, pricing, and Quality of Service“, *AT&T Labs Internal Report*, Florham Park, NJ, Juli 1998
- [100] ODLYZKO, A.: „The Internet and other networks: Utilization rates and their implications“, *AT&T Labs Internal Report*, Florham Park, NJ, September 1998
- [101] PATEL, G.; DENNET, S.: "The 3GPP and 3GPP2 Movements Toward an All-IP Mobile Network", *IEEE Personal Communication Magazine*, Aug. 2000
- [102] PANA: <http://www.ietf.org/html.charters/panacharter.html>, Abrufdatum 16.5.2005
- [103] PARLAY GROUP: <http://www.parlay.org/>, Aufrufdatum 16.5.2005
- [104] PERKINS, C.: „*IP Mobility Support*“, IETF RFC 2002, Oktober 1996
- [105] PERKINS, C.: „*Mobile IP Authentication, Authorization, and Accounting Requirements*“, IETF RFC 2977, Oktober 2000
- [106] PERKINS, C.: *Mobile IP: Design Principles and Practice*, AddisonWesley Longman, Reading, Mass., 1998
- [107] PIGOU, A.C.: *The Economics of Welfare*, London, Macmillan, 1920
- [108] POSTEL, J.: „*User Datagram Protocol*“, IETF RFC 768, August 1980
- [109] REGTP: <http://www.regtp.de>, Abrufdatum: 19.03.2005
- [110] REICHEL, P.; FANKHAUSER, G.; STILLER, B.: „Auction Models for Multi-Provider Internet Connections“, *10. GI/ITG Fachtagung Messung, Modellierung und Bewertung von Rechen und Kommunikationssystemen (MMB 99)*, Trier, September, 1999
- [111] RIGNEY, C.; RUBENS, A.; SIMPSON, W. WILLENS, S.: „*Remote Authentication Dial In User Service (RADIUS)*“, IETF RFC 2138, April 1997
- [112] RIGNEY, C.: „*RADIUS Accounting*“, IETF RFC 2139, April 1997
- [113] RIVEST, R. L.: „*The MD5 MessageDigest Algorithm*“, IETF RFC 1321, April 1992
- [114] ROMKEY, J.: „*Nonstandard for transmission of IP datagrams over serial lines: SLIP*“, RFC 1055, Juni 1988
- [115] SCHEFCZIK, P.; MITSCHLETHIEL, A.; SOELLNER, M.: „On MSCBased Performance Simulation“, *Proceedings of the third international workshop on Software and performance*, Rom, Juli 2002

- [116] SCHULZRINNE, H.; RAO, A.; LANPHIER, R.: „*Real Time Streaming Protocol (RTSP)*“, IETF RFC2326, April 1998
- [117] SHAPIRO, C.; VARIAN, H. R.: *Information Rules: A Strategic Guide to the Network Economy*, ISBN 087584863X, Harvard Business School Press, 1999
- [118] SHENKER, S.; PATRIDGE, C.; GUERIN, R.: „*Specification of Guaranteed Quality of Service*“, IETF RFC2212, September 1997
- [119] SIMSON, W. (Editor): „*The PointtoPoint Protocol (PPP)*“, STD 51, IETF RFC 1661, Juli 1994
- [120] SMITH, C; COLLINS, D.: *3G Wireless Networks*, McGrawHill Telecom, ISBN 0071363815, 2002.
- [121] STEPHAN, E., PALET, J.; „*Remote Network Monitoring (RMON) Protocol Identifiers for IPv6 and Multi Protocol Label Switching (MPLS)*“, IETF, RFC 3919, Oktober 2004
- [122] STEVENS, R.: *TCP/IP Illustrated, Volume 1: The Protocols*, AddisonWesley, 1994, ISBN 0201633469
- [123] STEWART, R; XIE, Q.; MORNEAULT, K.; SHARP, C.; SCHWARZBAUER, H.; TAYLOR, T.; RYTINA, I.; KALLA, M.; ZHANG, L.; PAXSON, V.: „*Stream Control Transmission Protocol*“, IETF RFC 2960, Oktober 2000
- [124] TELELOGIC: <http://www.telelogic.com/products/tau/index.cfm>, Abrufdatum 16.10.2004
- [125] THOMSON, S.; NARTEN T.: „*IPv6 Stateless Address Autoconfiguration*“, IETF RFC 1971, August 1996
- [126] UMTS Auktion, Internet Seite
<http://www.umtswebpage.de/UMTSAuktion/umtsauktion.html>, Abrufdatum 11.4.2007
- [127] VARIAN, H.: *Grundzüge der Mikroökonomik*, Oldenburg Verlag, ISBN 3486255436, 2001
- [128] VOLLBRECHT, J.; CALHOUN, P.; FARRELL, S.; GOMMANS, L.; GROSS, G; DE BRUIJN, B.; DE LAAT, C.; HOLDREGE, M.; SPENCE, D.: „*AAA Authorization Application Examples*“, IETF RFC 2905, August 2000
- [129] WALKE, B.; SEIDENBERG, P; Althoff, M. P.: *UMTS – The Fundamentals*, John Wiley & Sons, ISBN 047084557, 2003
- [130] WIEDNEBELING, S.: *Markt und Preistheorie*, Springer Verlag, ISBN 3540636269, 1997
- [131] WROCLAWSKI, J.: „*Specification of the ControlledLoad Network Element Service*“, IETF RFC 2211, September 1997
- [132] XIREN, C.: „Internet Pricing With a game Theoretical Approach: Concepts and Examples“, *IEEE/ACM Transactions on Networking*, Vol 10, No 2, April 2002
- [133] Zahariadis, T.B.: „*Migration toward 4G wireless communications*“, *IEEE Wireless Communication Magazine*, Volume 11, Issue 3, June 2004.

Anhang A: Signalspezifikation

In diesem Abschnitt werden die im Modell versendeten Nachrichten inklusive der Parameter beschrieben. Diese Spezifikation beschränkt sich nur auf die im Kontext dieser Arbeit relevanten Nachrichten. So werden z.B. nur diejenigen Nachrichten des Netzübergangs nach 6.2 in der nachfolgenden Tabelle aufgeführt, welche in direktem Zusammenhang mit den in Kapitel 6 dargestellten Verbesserungsvorschlägen stehen.

A.1 Nachrichtenformate zwischen Komponenten

Die folgende Tabelle beschreibt alle Nachrichten, welche im direkten Zusammenhang mit den im Rahmen dieser Arbeit gemachten Verbesserungsvorschlägen stehen und zwischen physikalischen Komponenten ausgetauscht werden. Abgesehen von den Nachrichten in direktem Zusammenhang mit dem Fast Handover nach 6.2.1.3 (*HO_Init* und *Ho_Ack*) wird vorgeschlagen, alle anderen Nachrichten mittels des Diameter/PANA Protokolls unter voller Kompatibilität zu den in Kapitel 4.4.1 und Kapitel 4.4.3 vorgestellten Standards zu transportieren.

Tabelle 16: Nachrichten zwischen Netzelementen

Name	Start	Ziel	Parameter
Flow	Input	AR	IP Quelladresse IPsrc (char) IP Zieladresse IPdst (char) Codepoint DSCP (char) Gesendete Daten Tooctets (int) Empfangene Daten Fromoctets (int)
RegReq	MT	AR	IP Quelladresse CoA (char) Benutzerkennung UID (char) Passwort – PWD (char) Benutzerdomäne – DU (char)
RegReqRel	AR	AAA.f	IP Quelladresse CoA (char) Benutzerkennung UID (char) Passwort – PWD (char) Benutzerdomäne – DU (char)
RegReqFWD	AAA.f	AAA.h	IP Quelladresse CoA (char) Benutzerkennung UID (char) Passwort – PWD (char)
RegResp	AAA.h	AAA.f	Sitzungs-ID Heimatnetz – SIDHD (char) Class-ID – CID (char) Credits – C (int)
RegRespFWD	AAA.f	AR	Sitzungs-ID Auswärtiges Netz SIDFD (char) Sitzungs-ID Heimatnetz – SIDHD (char) Class-ID – CID (char) Credits – Credits (int)
RegConf	AR	MT	Registration Confirmation – conf (char)
AccData	AR	AAA.f	Sitzungs-ID (komplett) SID CDR (Flow Parameter)
AccDataFWD	AAA.f	AAA.h	Sitzungs-ID (komplett) SID CDR (Flow Parameter)
CreditReq	AR	AAA.f	Sitzungs-ID (komplett) SID
CreditReqFWD	AAA.f	AAA.h	Sitzungs-ID (komplett) SID
CreditResp	AAA.h	AAA.f	Sitzungs-ID (komplett) – SID Credits – C (int)

CreditRespFWD	AAA.f	AF	Sitzungs-ID (komplett) – SID Credits – C (int)
HO_Init	oAR	nAR	Context Sitzungs-ID (komplett) SID Class-ID – CID (char) Credits – C (int)
Ho_Ack	nAR	oAR	HO Confirmation – conf (char)
Credit_Return	AR	AAA.f	Sitzungs-ID (komplett) – SID Unused Credits – C (int)
Credit_DReturn	AAA.f	AAA.h	Sitzungs-ID (komplett) – SID Unused Credits – C (int)

A.2 Nachrichtenformate innerhalb des Zugangsrouters

Die folgende Tabelle beschreibt alle Nachrichten, welche im direkten Zusammenhang mit den im Rahmen dieser Arbeit gemachten Verbesserungsvorschlägen stehen und in direktem Zusammenhang mit der vorgeschlagenen Architektur des Zugangsrouters stehen. Auf die detaillierte Implementierung wird im Rahmen dieser Arbeit nicht näher eingegangen.

Tabelle 17: Nachrichten innerhalb des Zugangsrouters

Name	Start	Ziel	Parameter
Meter_Data	Meter (NeTraMet)	Meter Manager (NeMaCDB)	IP Quelladresse - IPsrc (char) IP Zieladresse - IPdst (char) Codepoint - DSCP (char) Gesendete Daten Tooctets (int) Empfangene Daten - Fromoctets (int)
Meter_FWD	Meter Manager (NeMaCDB)	Meter Repository	IP Quelladresse - IPsrc (char) IP Zieladresse - IPdst (char) Codepoint - DSCP (char) Gesendete Daten Tooctets (int) Empfangene Daten - Fromoctets (int)
Meter_Start	Metering Steuereinheit	Metering	IP Quelleadresse - IPsrc (char) ClassID - CID (char)
Meter_Stop	Metering Steuereinheit	Metering	IP Quelleadresse - IPsrc (char) ClassID - CID (char)
Get_Data	Metering Steuereinheit	Accounting Repository	IP Quelladresse - IPsrc (char) IP Zieladresse - IPdst (char) Codepoint - DSCP (char) Gesendete Daten - Tooctets (int) Empfangene Daten - Fromoctets (int) Sitzungs-ID (komplett) - SID
Del_Data	Metering Steuereinheit	Accounting Repository	Sitzungs-ID (komplett) - SID
Get_UID	AAA Manager	Benutzer Repository	IP Quelladresse - IPsrc (char) Class-ID - CID (char) Sitzungs-ID (komplett) - SID
Set_UID	AAA Manager	Benutzer Repository	IP Quelladresse - IPsrc (char) Class-ID - CID (char) Sitzungs-ID (komplett) - SID
Sitzung_Stop	AAA Manager	QoS Manager	IP Quelladresse - IPsrc (char) IP Zieladresse - IPdst (char)

FHO	AAA Manager	Mobility Manager	Context Class-ID - CID (char) Sitzungs-ID (komplett) - SID Credits - C (int)
-----	----------------	---------------------	---

Anhang B: Testprotokolle

Die Tests sind aufgeteilt in statische und dynamische Tests. Die statischen Tests dienen dazu, Grundfunktionalitäten zu überprüfen, welche dann die Basis für weiterführende dynamische Tests bilden. Somit ist keine der Testspezifikationen für sich als komplett zu betrachten. In der Summe wird jedoch davon ausgegangen, dass diese Tests die relevanten Fälle abdecken und somit die Funktion des vorgeschlagenen Systems für diese Fälle nachgewiesen wird.

B.1 Statische Tests

Tabelle 18: Registrierung 1.0

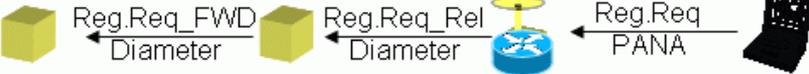
T Name: Testzweck: Testkonfiguration / Randbedingungen:	<p>Registrierung 1.0 Dieser Test dient zur Überprüfung der Registrierung in der Heimatdomäne</p> <p>AAA.h Server AAA.f Server AR MN</p>  <pre> sequenceDiagram participant MN as MN participant AR as AR participant AAAf as AAA.f Server participant AAAh as AAA.h Server MN->>AR: Reg.Reg PANA AR->>AAAf: Reg.Reg Rel Diameter AAAf->>AAAh: Reg.Reg FWD Diameter </pre>
Test Beschreibung:	
Testaufbau:	MN, AR, AAA.f, AAA.h
Testprozedur:	Ein MN sendet eine PANA Registrierung zum AR. Dieser leitet die Registrierung weiter zum AAA.f, welcher diese wiederum je nach UID zum AAA.h weiterleitet
Bewertungskriterien:	Erfolg: AAA.f erhält die Registrierung. Bei UID aus Heimatdomäne wird Nachricht weitergeleitet Bei UID aus Auswärtiger Domäne wird Nachricht nicht weitergeleitet, sondern lokal übergeben Bei anderer UID wird Nachricht verworfen
Ergebnis:	AAA.f erhält die Registrierung. Bei UID aus Heimatdomäne wird Nachricht weitergeleitet Bei UID aus Auswärtiger Domäne wird Nachricht nicht weitergeleitet, sondern lokal übergeben Bei anderer UID wird Nachricht verworfen

Tabelle 19: Registrierung 2.0

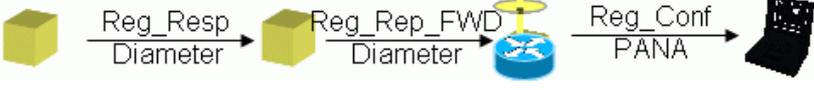
T Name: Testzweck: Testkonfiguration / Randbedingungen:	<p>Registrierung 2.0 Dieser Test dient zur Überprüfung des Benutzernamen/Passwort</p> <p>AAA.h Server AAA.f Server AR MN</p>  <p>Je bei AAA.f und AAA.h sind die Benutzer mit UID, PWD hinterlegt. Bei positiver Antwort wird das Profil zurückgegeben (Siehe Bild), bei negativer Antwort ein NACK</p>
Test Beschreibung:	
Testaufbau:	MN, AR, AAA.f, AAA.h.
Testprozedur:	Ein MN sendet ein Registrierungspaket mit UID/PWD über den AR zum AAA.f und ggf. dieser weiter zum AAA.h.
Bewertungs- kriterien:	Erfolg: Passwort ist richtig, positive Antwort (Antwort eines Profils. Misserfolg: Keine Antwort des Profils.
Ergebnis:	Passwort ist richtig, positive Antwort

Tabelle 20: Registrierung 2.1

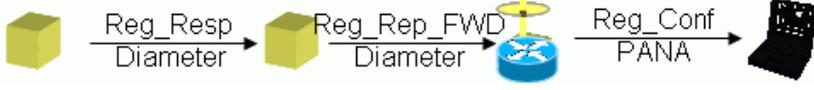
T Name: Testzweck: Testkonfiguration / Randbedingungen:	<p>Registrierung 2.1 Dieser Test dient zur Überprüfung des Benutzernamen/Passwort</p> <p>AAA.h Server AAA.f Server AR MN</p>  <p>Je bei AAA.f und AAA.h sind die Benutzer mit UID, PWD hinterlegt. Bei positiver Antwort wird das Profil zurückgegeben (Siehe Bild), bei negativer Antwort ein NACK</p>
Test Beschreibung:	
Testaufbau:	MN, AR, AAA.f, AAA.h.
Testprozedur:	Ein MN sendet ein Registrierungspaket mit falschem Benutzer/Passwort zum über den AR zum AAA.f.
Bewertungs- kriterien:	Erfolg: Passwort ist falsch, negative Antwort (Antwort einer NAK). Misserfolg: Antwort des Profils.
Ergebnis:	Negative Antwort

Tabelle 21: Registrierung 3.0

T Name:	Registrierung 3.0
Testzweck:	Dieser Test dient zur Überprüfung des Roamings
Testkonfiguration / Randbedingungen:	<p>AAA.h Server AAA.f Server AR MN</p> <p>AAA.h Server</p> <p>Beim AAA.h ist ein Benutzer <Benutzer> mit Passwort <Passwort> eingetragen; dem AAA.f ist eine Domäne <fdomain> zugewiesen, dem AAA.h ebenfalls <hdomain></p>
Test Beschreibung:	
Testaufbau:	Ein MN, ein AR, AAA.f und AAA.h.
Testprozedur:	Ein MN sendet ein Registrierungs paket mit benutzer@hdomain1/Passwort zum über den AR und AAA.f zum AAA.h.
Bewertungs-kriterien:	Erfolg: AAA.f schickt NAK an AR. Misserfolg: AAA.f kein NAK an AR.
Ergebnis:	AAA.f Registrierung an AAA.h

Tabelle 22: Registrierung 3.1

T Name:	Registrierung 3.1
Testzweck:	Dieser Test dient zur Überprüfung des Roamings
Testkonfiguration / Randbedingungen:	<p>AAA.h Server AAA.f Server AR MN</p> <p>AAA.h Server</p> <p>Die Heimatdomäne des Benutzers ist dem AAA.f Server als RoamingPartner nicht bekannt. A</p>
Test Beschreibung:	
Testaufbau:	Ein MN, ein AR, AAA.f und (zwei) AAA.h.
Testprozedur:	Ein MN sendet ein Registrierungs paket mit benutzer@hdomain1/Passwort zum über den AR zum AAA.f
Bewertungs-kriterien:	Erfolg: AAA.f schickt NAK an AR. Misserfolg: AAA.f leitet Anfrage an falschen AAA.h weiter.
Ergebnis:	AAA.f schickt NAK an AR

Tabelle 23: Registrierung 4.0

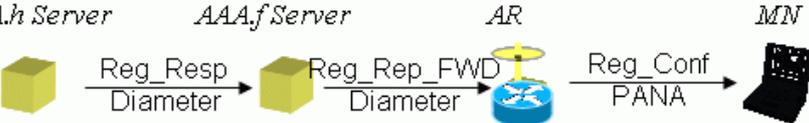
T Name: Testzweck: Testkonfiguration / Randbedingungen:	<p>Registrierung 4.0 Dieser Test dient zur Überprüfung des Benutzerprofil-Transfers beim Zugangsrouter</p> <p>AAA.h Server AAA.f Server AR MN</p>  <p>Die Antwort eines erfolgreichen Authentisierungsprozesses ist ein Benutzerprofil, welches vom AAA.h zum AAA.f gesendet wird. Der AAA.f behält ein Teil des Benutzerprofils lokal und leitet einen weiteren Teil an den Zugangsrouter weiter (u.a. QoS Parameter wie Bandbreite,...)</p>
Test Beschreibung:	
Testaufbau:	Ein MN, ein AR, AAA.f und AAA.h.
Testprozedur:	AAA.h beantwortet mit dem Benutzerprofil. AAA.f behält Benutzerprofil 1 aus Abbildung 43 und leitet Benutzerprofil 2 aus Abbildung 43 zum AR weiter. QoS-Manager im AR stellt korrekte Parameter ein.
Bewertungskriterien:	Erfolg: AR stellt Bandbreite ein. Und AAA.f hat Benutzerprofil 1 Misserfolg: AR stellt keine Bandbreite ein. / AAA.f hat kein Benutzerprofil 1.
Ergebnis:	AR stellt Bandbreite ein; Benutzerprofil 1 in AAA.f

Tabelle 24: Registrierung 5.0

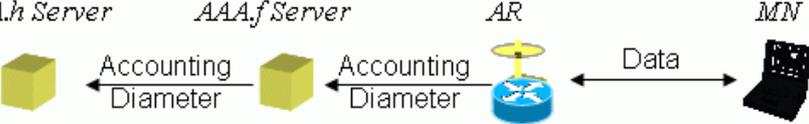
T Name: Testzweck: Testkonfiguration / Randbedingungen:	<p>Registrierung 5.0 Dieser Test dient zur Überprüfung der korrekten Funktion der Nutzungsdatenerfassung</p> <p>AAA.h Server AAA.f Server AR MN</p>  <p>Ein Benutzer bekommt nach erfolgreicher Authentisierung Netzzugang und konsumiert Ressourcen. Accounting Daten werden im AR generiert und je nach Konfiguration weitergeleitet.</p>
Test Beschreibung:	
Testaufbau:	Ein MN, ein AR, AAA.f und AAA.h.
Testprozedur:	Ein MN (erfolgreich authentisiert, Postpaid, keine Flat-Rate) sendet 10 Pakete.
Bewertungskriterien:	Erfolg: Nutzungsnachweis in AR Misserfolg: Nutzungsnachweis nicht in AR.
Ergebnis:	Nutzungsnachweis in AR

Tabelle 25: Registrierung 5.1

T Name: Testzweck: Testkonfiguration / Randbedingungen:	<p>Registrierung 5.1 Dieser Test dient zur Überprüfung der korrekten Funktion der Nutzungsdatenerfassung</p> <p>AAA.h Server AAA.f Server AR MN</p>  <p>Ein Benutzer bekommt nach erfolgreicher Authentisierung Netzzugang und konsumiert Ressourcen. Accounting Daten werden im AR nicht generiert, da Flat-Rate Benutzer.</p>
Test Beschreibung:	
Testaufbau:	Ein MN, ein AR, AAA.f und AAA.h.
Testprozedur:	Ein MN (erfolgreich authentisiert, Flat-Rate) sendet Daten. Meter wird nicht konfiguriert. Keine Datenerfassung nachweisbar
Bewertungs- kriterien:	Erfolg: Nutzungsnachweis nicht in AR Misserfolg: Nutzungsnachweis in AR
Ergebnis:	Nutzungsnachweis nicht in AR

Tabelle 26: Registrierung 6.0

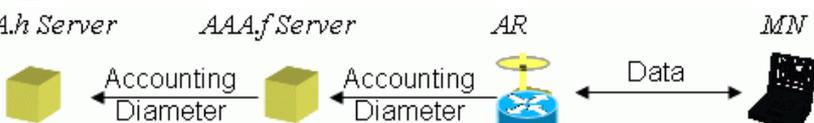
T Name: Testzweck: Beschreibung: Testkonfiguration / Randbedingungen:	<p>Registrierung 6.0 Dieser Test dient zur Überprüfung des korrekten Arbeitens des Accounting-Intervalls im Zugangsrouter. Hier wird getestet, ob sich der AAA Manager im AR korrekt verhält</p> <p>AAA.h Server AAA.f Server AR MN</p>  <p>Benutzer 1 ist Prepaid Benutzer, Benutzer 2 Postpaid Benutzer (beide keine Flat-Rate). Jeder Benutzer sendet 10 Pakete aus. Accounting-Intervall beider Benutzer ist verschieden</p>
Test Beschreibung:	
Testaufbau:	Ein MN, ein AR, AAA.f und AAA.h.
Testprozedur:	Benutzer 1 authentisiert sich und sendet 10 Pakete aus. Es werden 10 Accounting-Berichte erzeugt. Benutzer 2 Authentisiert sich und sendet 10 Pakete aus. E wird nur ein konsolidierter Accounting-Bericht erzeugt.
Bewertungs- kriterien:	Erfolg: je nach Benutzer Anzahl und Inhalt der Berichte korrekt Misserfolg: Anzahl und Inhalt der Berichte nicht korrekt.
Ergebnis:	Anzahl und Inhalt der Berichte korrekt

Tabelle 27: Registrierung 6.1

T Name: Testzweck: Beschreibung: Testkonfiguration / Randbedingungen:	<p>Registrierung 6.1 Dieser Test dient zur Überprüfung des korrekten Arbeitens des Accounting-Intervalls im AAA.f Server. Hier wird getestet, ob sich der AAA.f Server korrekt verhält</p> <p>AAA.h Server AAA.f Server AR MN</p>  <p>Benutzer 1 ist Prepaid Benutzer, Benutzer 2 Postpaid Benutzer (beide keine Flat-Rate). Jeder Benutzer sendet 100 Pakete aus. Accounting-Intervall beider Benutzer ist verschieden. Benutzer 2 verhält sich nach Test „Registrierung 6.0“</p>
Test Description:	
Testaufbau:	Ein MN, ein AR, AAA.f und AAA.h.
Testprozedur:	Benutzer 1 authentisiert sich und sendet 100 Pakete aus. Es werden 100 Accounting-Berichte erzeugt (Messpunkt im AAA.h). Benutzer 2 Authentisiert sich und sendet 100 Pakete aus. Es werden 10 konsolidierte Accounting-Bericht im AAA.f und ein Bericht im AAA.h erzeugt.
Bewertungskriterien:	Erfolg: je nach Benutzer Anzahl und Inhalt der Berichte korrekt Misserfolg: Anzahl und Inhalt der Berichte nicht korrekt
Ergebnis:	Anzahl und Inhalt der Berichte korrekt

B.2 Dynamische Tests

Das Ziel der dynamischen Tests ist aufbauend auf den statischen Tests die Funktionalität der Spezifikation des Gesamtsystems für diese gewählten Szenarien nachzuweisen. In diesem Abschnitt soll zunächst die Auswahl der dynamischen Szenarien begründet werden. Konkret gehen die folgenden dynamischen Testszzenarien von folgender Annahme aus:

Im Modell sind drei Netzbetreiber (A, B, C) aktiv, welche Netzdienste anbieten. Untereinander unterhalten diese Anbieter ein Roaming-Abkommen. Im Modell gibt es einen Benutzer des mobilen Endsystems, welcher bei Provider C unter Vertrag steht. Das Netz des Providers C ist somit das Heimatnetz.

Um die Validierung in diesem Stadium der Arbeit reproduzierbar zu machen, wurden vier vordefinierte Wege festgelegt, welche von einem Benutzer mit seinem mobilen Endsystem durchlaufen werden. Das erzeugte Verkehrsprofil ist statisch und erfolgt nach Kapitel 7.3.5. In der Simulation wurden drei Netzbetreiber vorgesehen, um beide Übergänge, also zwischen Heimatnetz und auswärtigem Netz und zwischen zwei verschiedenen auswärtigen Netzen in beiden Richtungen validieren und verifizieren zu können.

Bevor nun die einzelnen Ergebnisse der Simulationswege vorgestellt werden, soll deren Plausibilität begründet werden:

Weg 1:

Das Endsystem wird im auswärtigen Netz (Betreiber A) eingeschaltet und registriert sich dort. Anschließend erfolgen zwei Netzübergänge im registrierten Netz, ein Netzübergang zum Betreiber B und weitere zwei Netzübergänge bei Betreiber B. Anschließend erfolgt ein Netzübergang ins Heimatnetz von Betreiber C. Auch hier erfolgen weitere zwei Netzübergänge.

Weg 2:

Das Endsystem wird wiederum im auswärtigen Netz des Betreibers A registriert und es erfolgen alternierende Netzübergänge zwischen zwei verschiedenen auswärtigen Netzen (Betreiber A und Betreiber B).

Weg 3:

Hier wird das Endsystem im Netz des Betreibers C, also dem Heimatnetz, registriert. Es erfolgen nun analog zu Weg 2 alternierende Netzübergänge zum Netz des Betreibers B und zurück. Im Gegensatz zu Weg 2 werden hier jetzt Übergänge zwischen dem Heimatnetz und einem auswärtigen Netz vollzogen.

Weg 4:

Analog zu Weg 1 wird das Endsystem im Heimatnetz registriert und es erfolgen nach je zwei Netzübergängen im aktuellen Betreibernetz ein Übergang ins Netz B und Netz A.

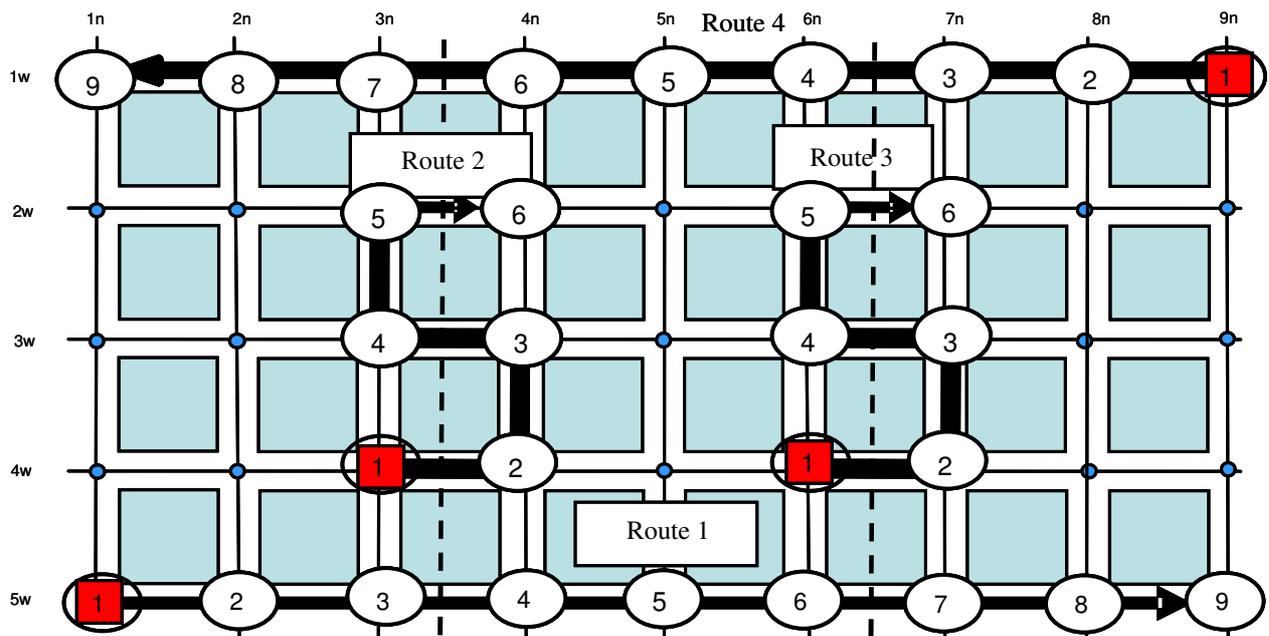


Abbildung 66: Simulationswege des Endsystems

Fiktiv wird das mobile Endsystem von verschiedenen Benutzern verwendet, welche über verschiedene Vertragsmodelle an den Heimatbetreiber gebunden sind. Diese Benutzer sind in Tabelle 28 dargestellt.

Tabelle 28: Verträge der verschiedenen Benutzer

Benutzer	Vertragsart	Verbleibendes Budget
Benutzer 1	Prepaid	2
Benutzer 2	Prepaid	5
Benutzer 3	Prepaid	7
Benutzer 4	Prepaid	10
Benutzer 5	Postpaid	
Benutzer 6	Postpaid Flat	

Benutzer 1 durchläuft Weg 1 und das verbleibende Budget ist im Netz des Betreiber A aufgebraucht.

Benutzer 2 durchläuft ebenfalls Weg 1, jedoch ist das verbleibende Budget erst bei Betreiber B aufgebraucht.

Benutzer 3 durchläuft ebenfalls Weg 1. Hier ist das verbleibende Budget erst im Netz des Betreibers C, also dem Heimatnetz aufgebraucht.

Benutzer 4 durchläuft Weg 1. Das verbleibende Budget reicht aus, um den kompletten Weg abzulaufen.

Bei Benutzer 5 fallen die entsprechenden Nutzungsdatenberichte konsolidiert im Heimatnetz an, da es sich hierbei um einen Postpaid-Vertrag handelt.

Benutzer 6 hat einen Flat-Rate-Vertrag. Es werden keine Nutzungsdaten erfasst.

Anhang C: SDL – Spezifikation

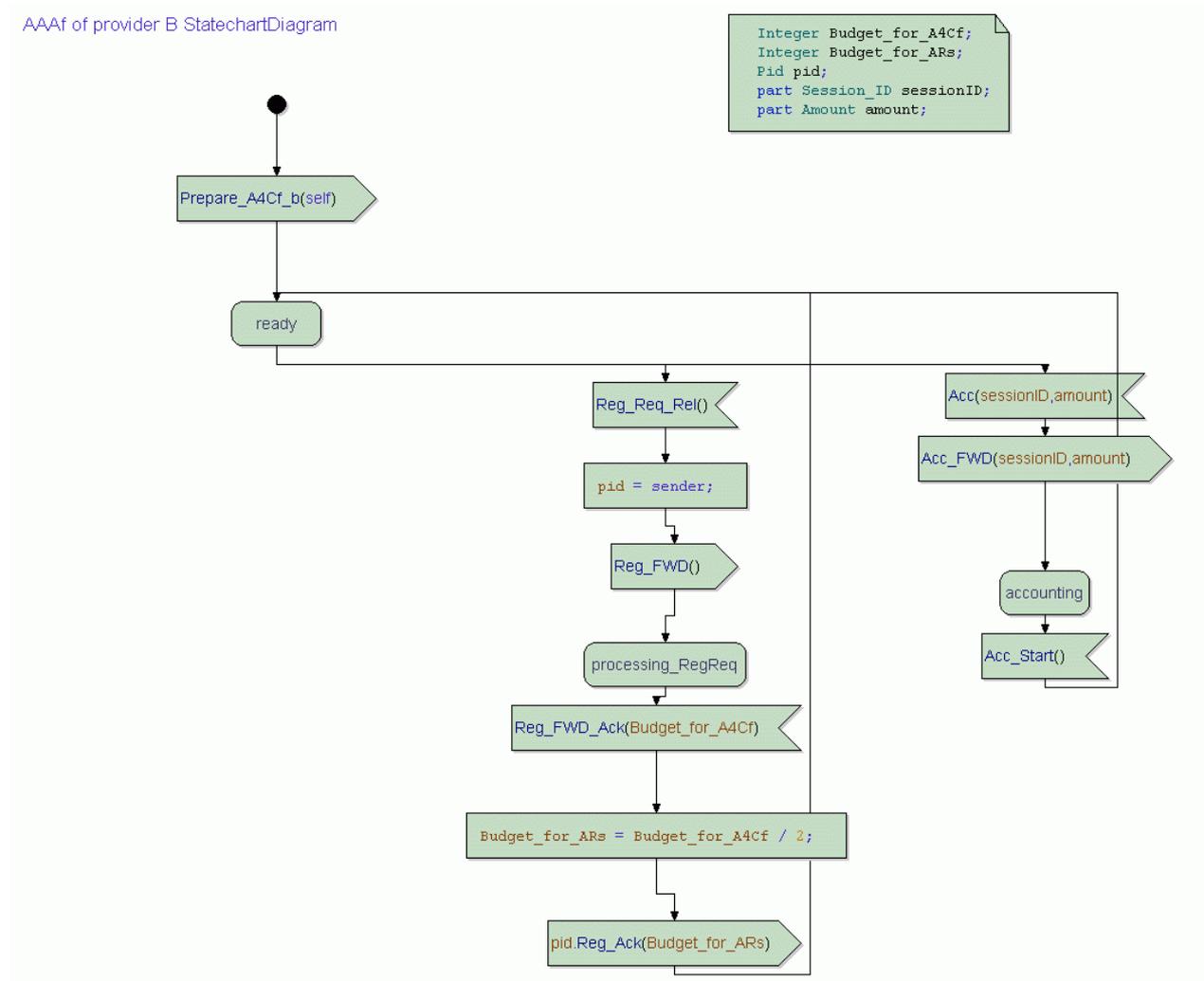


Abbildung 67: AAA-Server in auswärtiger Domäne (Beschränkung auf Accounting)

Die folgenden vier Seiten zeigen das modellierte SDL-Diagramm des Zugangsrouters von Provider A (auswärtige Domäne) aus dem modellierten Manhattan-Grid-Modell. Wegen der besseren Lesbarkeit wurde das Diagramm ohne die Einführung neuer Konnektoren bzw. Zustände an den jeweiligen Stellen abgeschnitten.

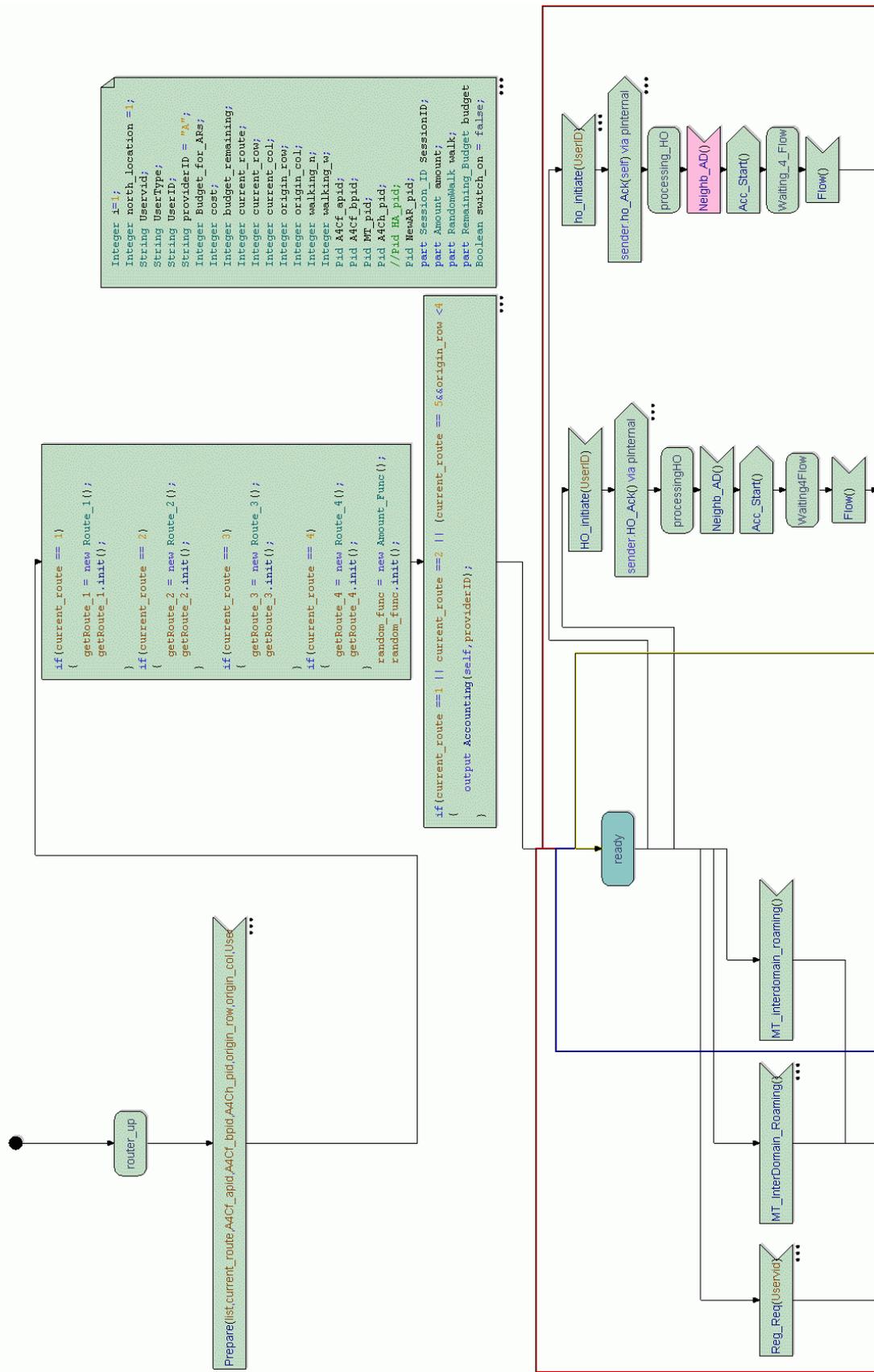
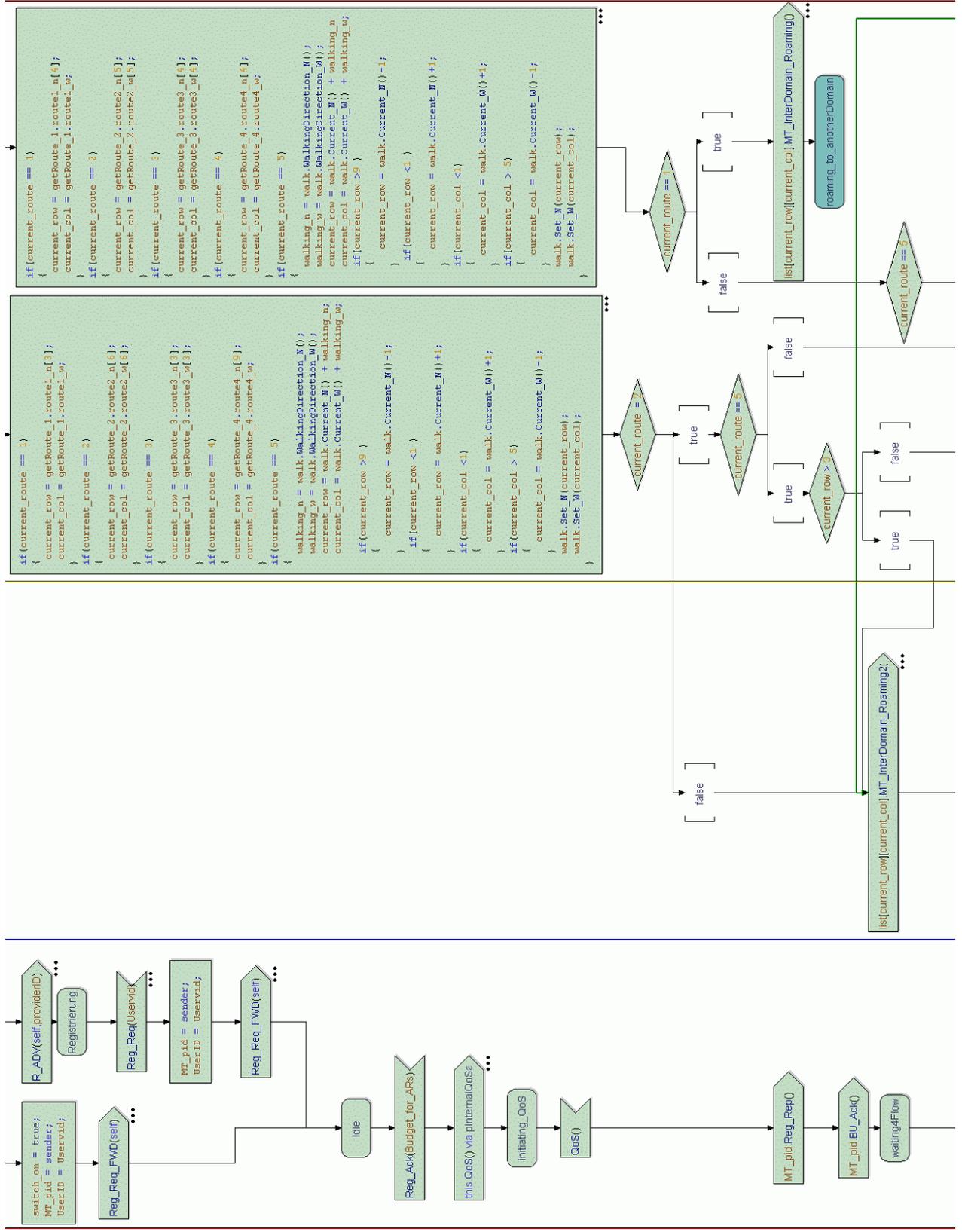


Abbildung 68: Zugangsroutermmodellierung (1)



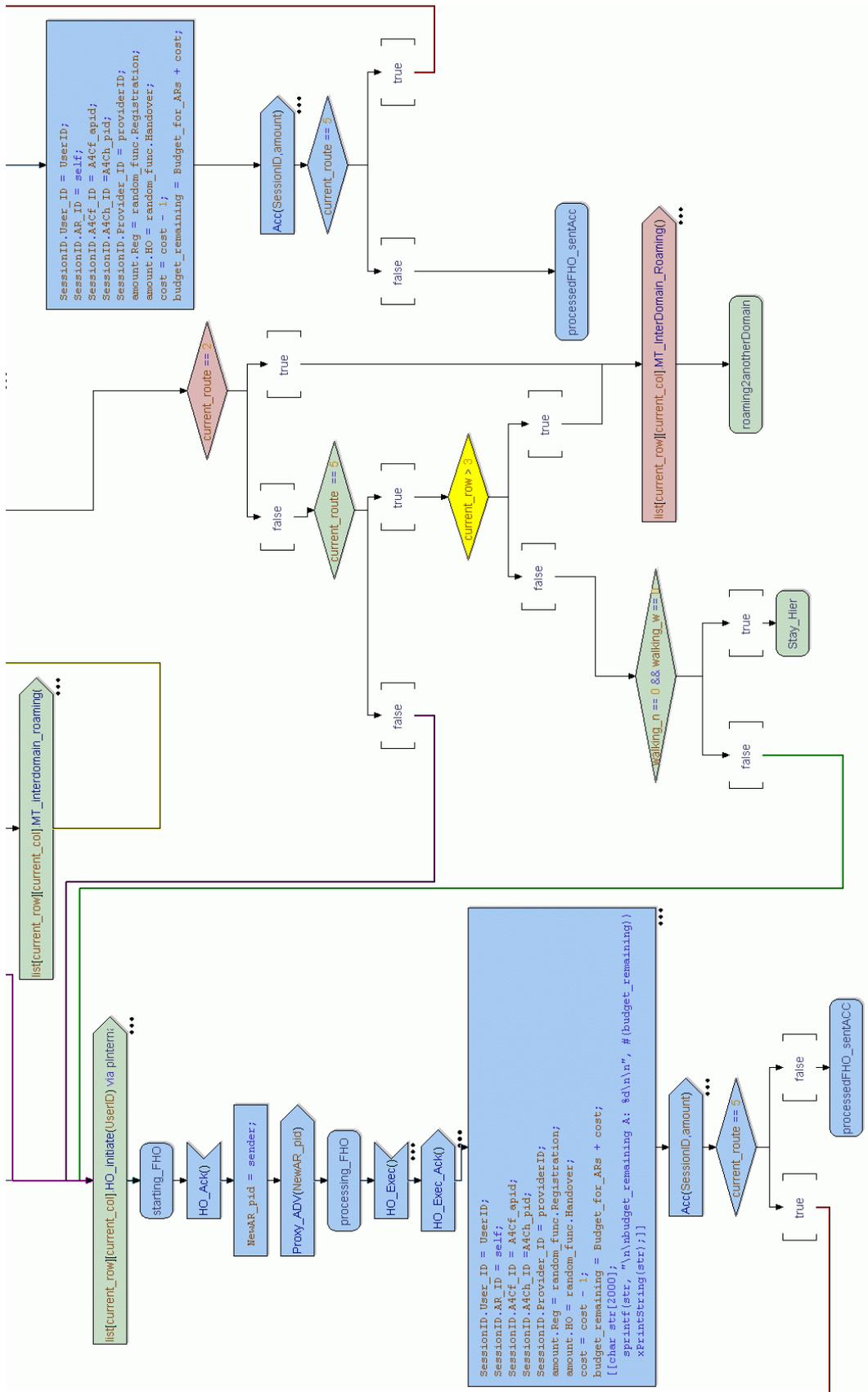


Abbildung 68: Zugangsroutermodellierung (4)

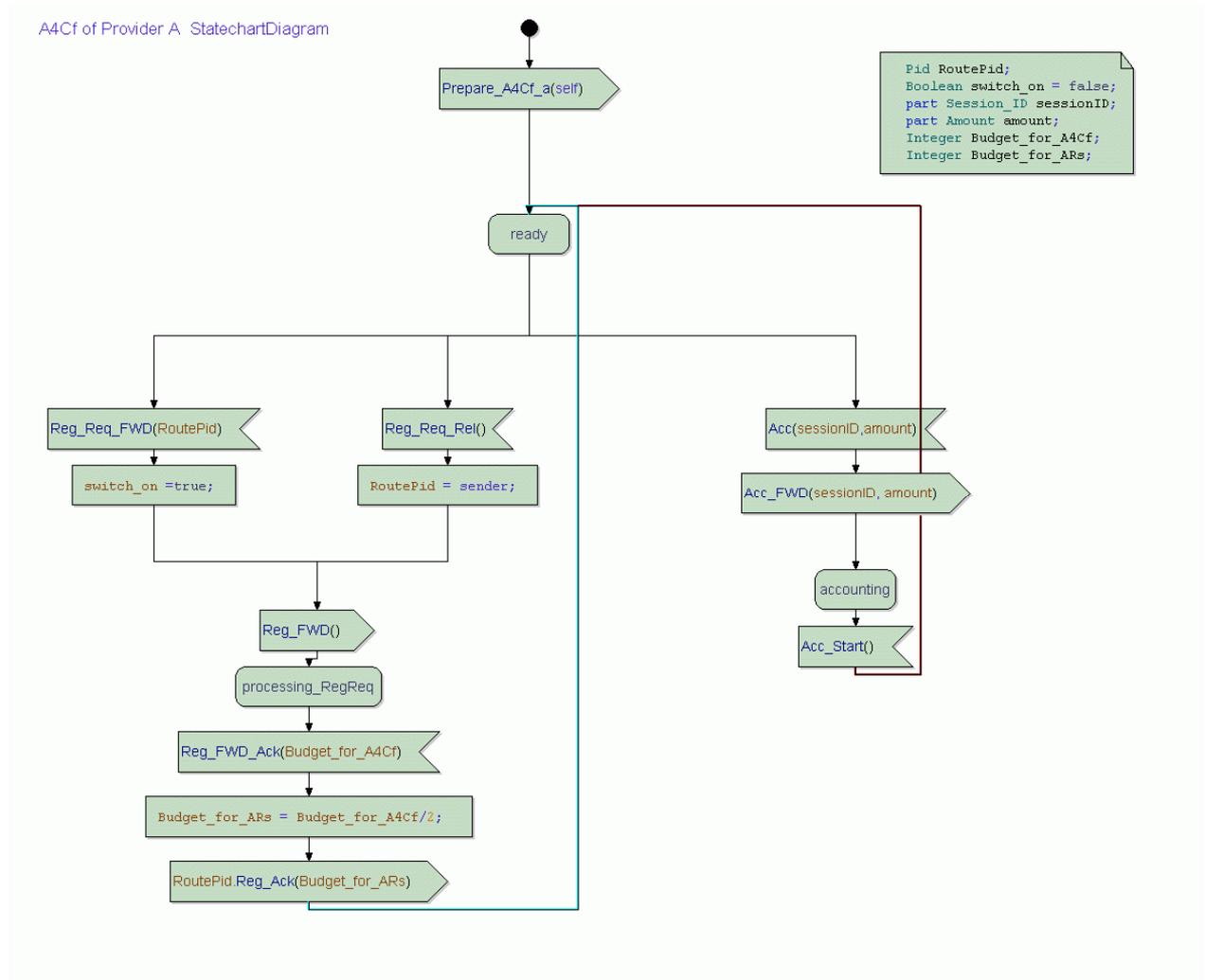


Abbildung 69: AAA-Server in Heimatdomäne

Die folgenden 4 Seiten beschreiben wieder analog zu Abbildung 68 das Zustandsdiagramm des Zugangsrouters der Heimatdomäne in vier Teile aufgeteilt, welche als zusammenhängendes Diagramm zu Lesen sind.

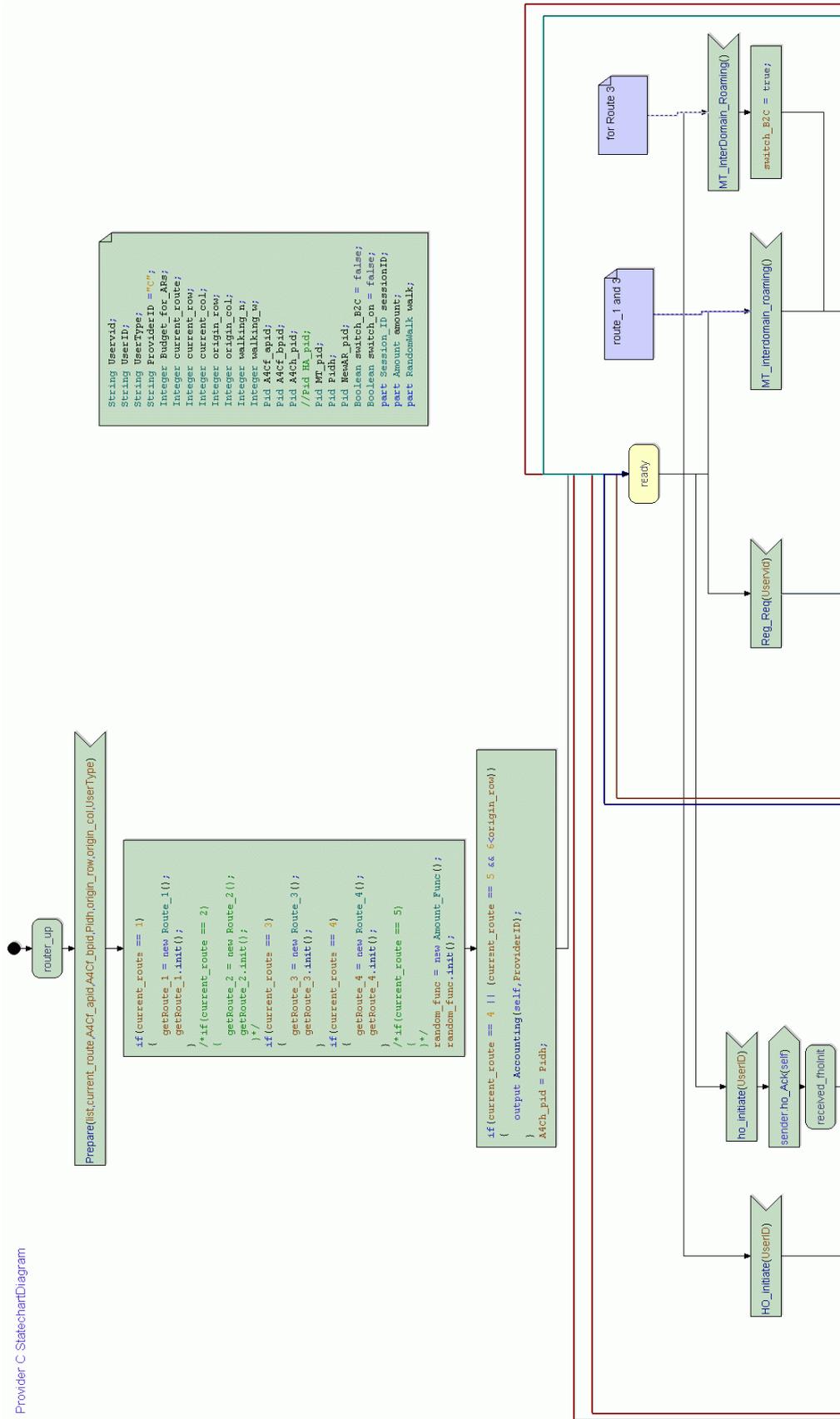


Abbildung 70: Zugangsrouten in der Heimatdomäne (1)

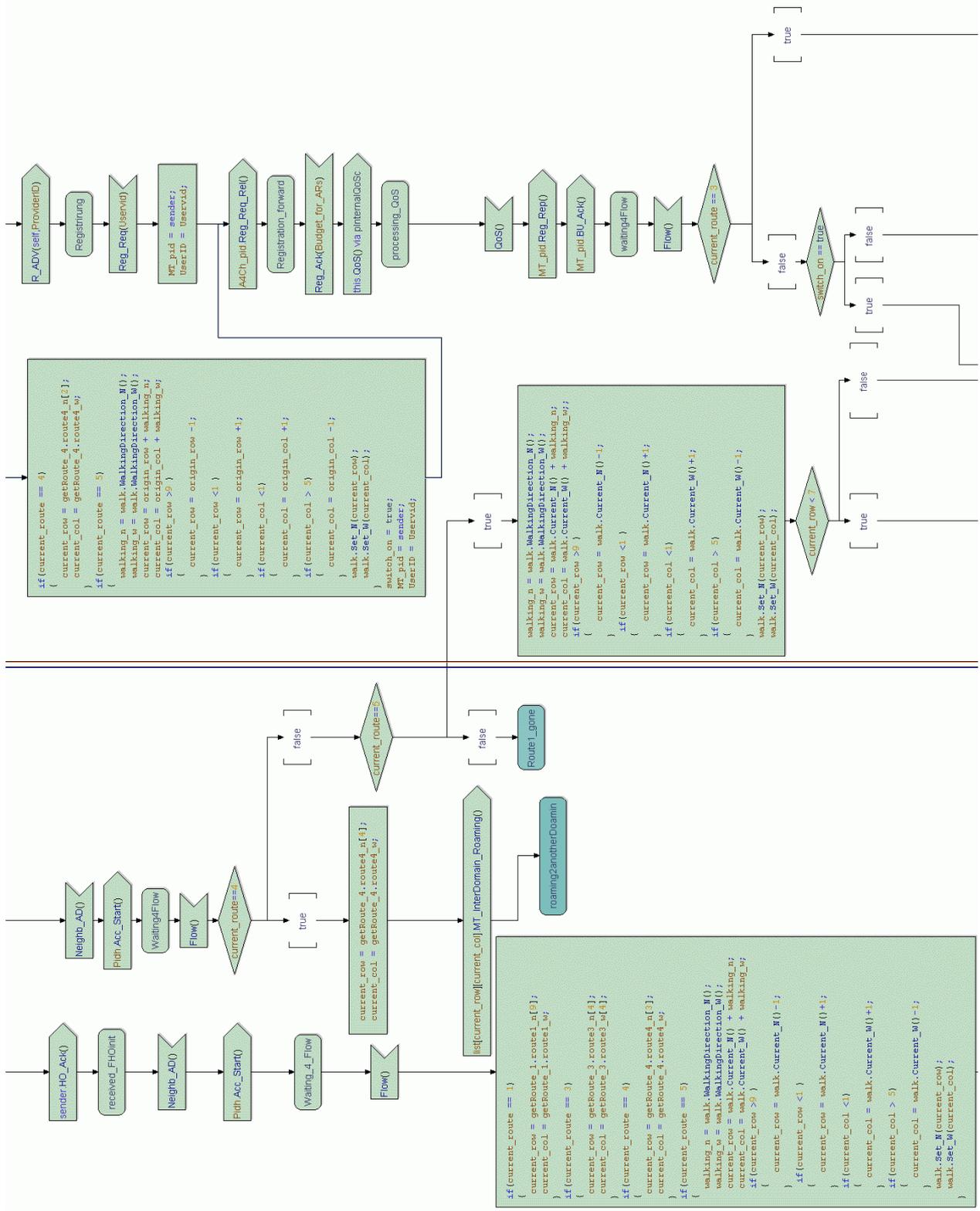


Abbildung 70: Zugangsrouter in der Heimatdomäne (2)

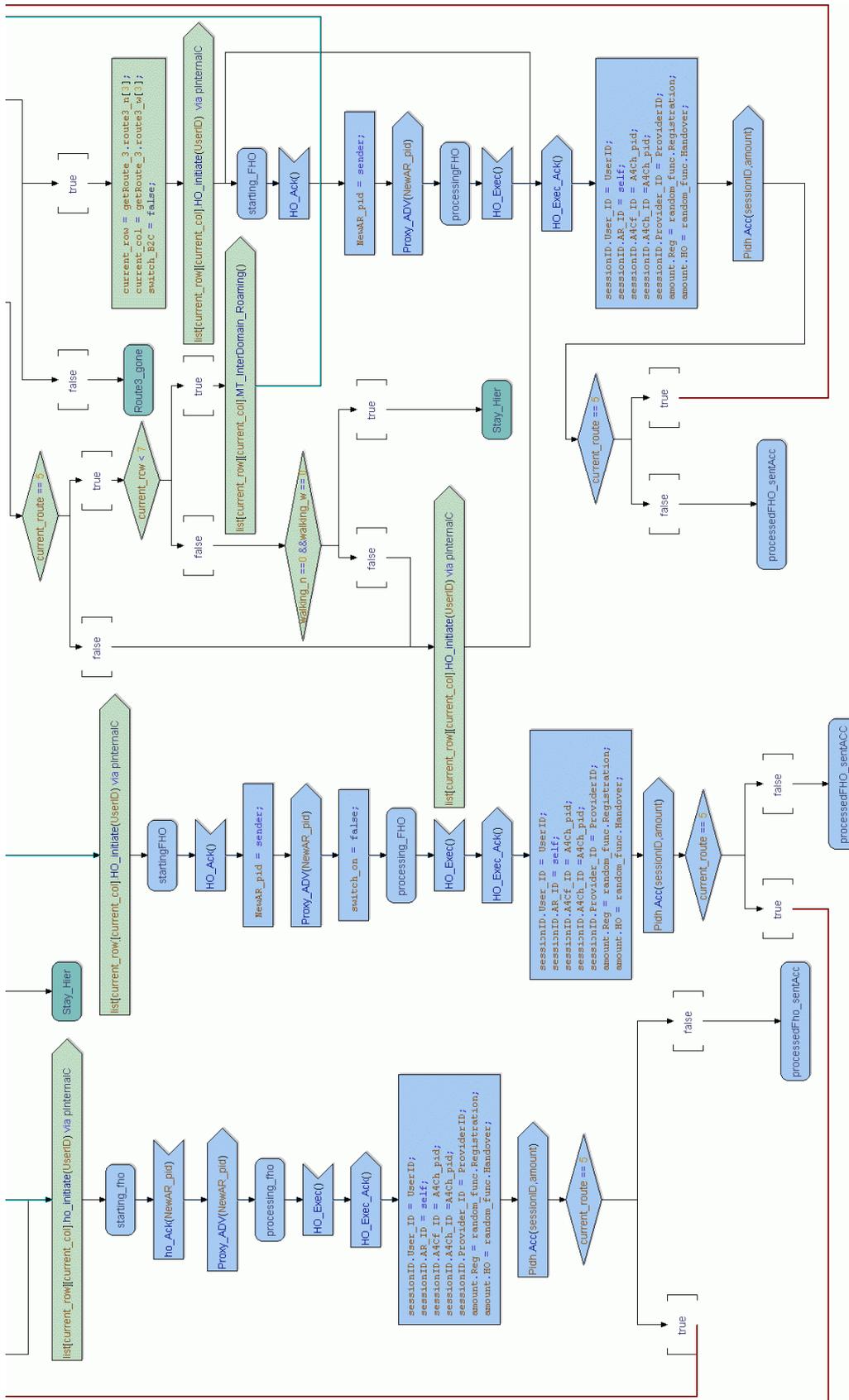


Abbildung 70: Zugangsrouten in der Heimatdomäne (4)