

# **Verteilt organisierte Mehrteilnehmer–Videokonferenzen für geschlossene Gruppen im Internet**

Von der Fakultät für Mathematik, Naturwissenschaften und Informatik  
der Brandenburgischen Technischen Universität Cottbus

zur Erlangung des akademischen Grades eines  
Doktors der Ingenieurwissenschaften (Dr.–Ing.)

genehmigte Dissertation

vorgelegt von

Diplom–Ingenieur

Mario Zühlke

geboren am 02.Mai 1974 in Spremberg

Gutachter: Prof. Dr.–Ing. habil. Hartmut König

Gutachter: Prof. Dr.–Ing. Klaus Fellbaum

Gutachter: Prof. Dr.–Ing. Wolfgang Effelsberg

Tag der mündlichen Prüfung: 27.Mai 2004



---

## Vorwort

---

Die vorliegende Arbeit ist während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Lehrstuhl für Rechnernetze und Kommunikationssysteme der Brandenburgischen Technischen Universität Cottbus vom September 1998 bis September 2003 entstanden. Ich habe dort im Rahmen des vom DAAD geförderten Forschungsprojekts „Software Support for Scalable Multiparty Interactive Multimedia Applications“ und im Rahmen des vom MWFK geförderten Projekts „Entwicklung eines Videokonferenzdienstes für Teleteaching-Anwendungen in den Hochschulen des Landes Brandenburg“ gearbeitet. Das Ziel der Projekte — Mehrteilnehmer-Videokonferenzen geschlossener Gruppen über das Internet — hat den Gegenstand meiner Forschungstätigkeit geprägt.

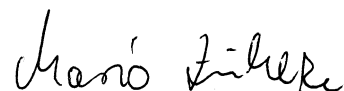
An dieser Stelle möchte ich mich bei Herrn Prof. Dr. König bedanken, der die Arbeit und die in ihrem Zusammenhang entstandenen Veröffentlichungen ermöglicht, unterstützt und mit konstruktiver Kritik mitgeformt hat.

Mein besonderer Dank gilt meiner Frau Bettina, die mir oft den nötigen Rückhalt gegeben hat, und meinem Sohn Magnus-Benedikt, der hin und wieder den notwendigen Abstand zur Arbeit vermittelt hat.

Bedanken möchte ich mich auch bei den Herren Prof. Dr. Fellbaum und Prof. Dr. Effelsberg für den Aufwand der Begutachtung der vorliegenden Arbeit.

Weiterhin danke ich meinen Kollegen und Kolleginnen sowie Diplomanden und Studienarbeitern für die gute Zusammenarbeit. Besonderer Dank gilt dabei Dr. Eduard Popovici, Ralf Mahlo und Fuwen Liu für die fruchtbare gemeinsame Forschungsarbeit im Bereich der Videoübertragung und -skalierung sowie Joachim Paschke für die gute technische Unterstützung.

Spremberg, im Juni 2004





---

# Zusammenfassung

---

Videokonferenzen gelten als eine populäre multimediale Anwendung in modernen Hochleistungsnetzen. Trotz langjähriger Forschungsarbeiten und Erfahrungen gehören Videokonferenzen noch nicht zu den alltäglich genutzten Telekommunikationsdiensten wie Email oder WWW. Videokonferenzdienste im Internet unterstützen vor allem offene Konferenzgruppen zur Übertragung von öffentlichen Veranstaltungen, wie Vorlesungen oder Fernsehprogrammen. Jedoch haben die Mehrzahl der alltäglichen Treffen und Beratungen einen eher geschlossenen Charakter. Systeme, die geschlossene Gruppen unterstützen, sind noch relativ teuer und basieren zumeist auf spezieller Studioteknik. Desktop-Systeme lassen nur ein begrenztes Leistungsangebot zu. Besonders beim Aufbau von Mehrteilnehmer-Videokonferenzen ist es nach wie vor schwierig, eine gute Dienstqualität bereitzustellen. Probleme bereiten dabei vor allem die Unterstützung heterogener Systemumgebungen, die Anpassungsfähigkeit (Skalierbarkeit) an sich ändernde Systembedingungen und die Überwindung großer Distanzen. Dies hat zur Folge, dass Videokonferenzen für geschlossene Gruppen bisher keine breite Anwendung gefunden haben.

Die vorliegende Arbeit untersucht Gestaltungsprinzipien für geschlossene Mehrteilnehmer-Videokonferenzsysteme in globalen öffentlichen Netzwerken wie dem Internet. Dabei werden zwei Ansätze verfolgt. Der erste Ansatz untersucht den Aufbau von globalen Konferenzen über den Verbund lokaler Videokonferenzsysteme. Damit können vorhandene lokale Systeme auch für die globale Kommunikation genutzt und Investitionen geschützt werden. In diesem Zusammenhang werden das Problem der Qualitätssicherung bei der Übertragung von kontinuierlichen Daten, wie zum Beispiel Video und Audio, in globalen öffentlichen Umgebungen untersucht. Der zweite Ansatz beschäftigt sich daraus abgeleitet mit der Gestaltung von Videokonferenzsystemen für geschlossene Gruppen im Internet. Der Fokus liegt dabei auf einem dezentral organisierten Peer-to-Peer-Konzept, durch den, im Vergleich zu existierenden Systemen, kostengünstigere und netzwerkunabhängige Lösungen ermöglicht werden sollen. Diese besitzen eine höhere Flexibilität und sollen insbesondere einen spontanen Aufbau von Konferenzen sowie die Mobilität der Anwender unterstützen. Die in der Arbeit untersuchten Ansätze wurden implementiert und in die Videokonferenzsysteme OCTOPUS und BRAVIS integriert. Die Lösungen werden aus verschiedenen Gesichtspunkten bewertet. Umfangreiche Messungen bestätigen die Anwendbarkeit der untersuchten Ansätze.



---

## Abstract

---

Video conferences are considered as an attractive application in high-performance networks. Nevertheless video conference services are not as broadly used as services like e-mail or the WWW. Most existing video conference systems support open group conferences for transmitting public events, like lectures, TV programs, music concerts, talks, and conferences. Most everyday meetings and consultations however, have a closed character. Systems supporting closed group communication are relatively expensive and are mostly based on special studio hardware. The performance of desktop video conference systems is limited. In particular, there are difficulties in establishing multi party video conferences with sufficient quality of service (QoS). The heterogeneity of system environments, the scalability depending on changing system conditions, and the delivery over large distances are still open issues. Therefore, multi party video conference tools for closed group meetings are rarely used.

This Ph.D. thesis investigates the design principles of closed multi party video conference systems for global public networks. Hereby two approaches are described. The first approach investigates a global video conferencing service that interconnects local conference systems over long distances where no assumptions on the architecture of the local systems are made. Thus the investment into a local video conference system can be saved. In this context the important problem discussed is how to ensure the QoS of the transport of real time traffic like audio and video in global public network environments. The second approach deals with the design of closed multiparty conferences in the Internet. It applies a distributed Peer-to-Peer approach which allows less expensive and network independent solutions compared with existing systems. Peer-to-Peer approaches are more flexible. They support a spontaneous set up of conferences and the mobility of the users as well. Both approaches have been implemented in the video conference systems OCTOPUS and BRAVIS. The applicability of the proposed concepts are discussed based upon numerous performance measurements.





---

# Inhaltsverzeichnis

---

<b>Abbildungsverzeichnis</b>	<b>V</b>
<b>Tabellenverzeichnis</b>	<b>IX</b>
<b>Glossary</b>	<b>XI</b>
<b>1. Einleitung und Problemstellung</b>	<b>1</b>
1.1. Motivation . . . . .	1
1.2. Problemstellung . . . . .	4
1.3. Zielsetzung und Aufbau der Arbeit . . . . .	8
<b>2. Gestaltungsprinzipien von Videokonferenzsystemen</b>	<b>13</b>
2.1. Anwendungsbereiche . . . . .	13
2.2. Arten von Videokonferenzsystemen . . . . .	15
2.3. Architektur von Videokonferenzsystemen . . . . .	20
2.4. Management eines Videokonferenzsystems . . . . .	21
2.4.1. Gruppenmanagement . . . . .	22
2.4.2. QoS-Management . . . . .	23
2.4.3. Floorkontrolle . . . . .	25
2.5. Medienverarbeitung . . . . .	26
2.5.1. Videokompression . . . . .	27
2.5.2. Audioverarbeitung . . . . .	33
2.6. Netzwerkunterstützung und Transport . . . . .	36
2.7. Überblick über vorhandene Videokonferenzlösungen . . . . .	44
2.8. Gestaltungsmöglichkeiten von Videokonferenzsystemen . . . . .	46
2.8.1. Einsatzgebiet . . . . .	46
2.8.2. Offene und geschlossene Videokonferenzsysteme . . . . .	50
2.8.3. Topologien von Videokonferenzsystemen . . . . .	53

2.9. Zusammenfassung . . . . .	59
<b>3. Gestaltung geschlossener Videokonferenzsysteme über das Internet</b>	<b>63</b>
3.1. Der OCTOPUS-Ansatz . . . . .	63
3.1.1. Systemarchitektur des OCTOPUS-Ansatzes . . . . .	63
3.1.2. Floorkontrolle . . . . .	66
3.1.3. Gruppenkoordinator . . . . .	67
3.1.4. Kommunikationsszenarien . . . . .	69
3.1.5. Bewertung des OCTOPUS-Ansatzes . . . . .	76
3.2. Der BRAVIS-Ansatz . . . . .	76
3.2.1. BRAVIS-Funktionalität . . . . .	77
3.2.2. Topologie . . . . .	79
3.2.3. Architektur . . . . .	81
3.2.4. Bewertung des BRAVIS-Ansatzes . . . . .	82
3.3. Zusammenfassung . . . . .	83
<b>4. Sicherung der Geschlossenheit der Gruppe in globalen Videokonferenzsystemen</b>	<b>87</b>
4.1. Anforderungen an einen dezentralen Managementansatz . . . . .	88
4.2. GCP <sub>ATM</sub> . . . . .	89
4.3. GCP <sub>IP</sub> . . . . .	94
4.3.1. Der GCP <sub>IP</sub> -Dienst . . . . .	95
4.3.2. Das GCP <sub>IP</sub> -Protokoll . . . . .	99
4.3.3. Diskussion der GCP <sub>IP</sub> Lösung . . . . .	113
4.4. Signalisierung in hierarchischen Topologien . . . . .	127
4.4.1. Signalisierungsprinzipien in OCTOPUS . . . . .	127
4.4.2. Abschätzung der Signalisierungsverzögerung . . . . .	129
4.5. Zusammenfassung . . . . .	131
<b>5. Übertragung multimedialer Daten in globalen Videokonferenzsystemen</b>	<b>133</b>
5.1. Medientransport in Videokonferenzsystemen . . . . .	133
5.1.1. Anforderungen . . . . .	133
5.1.2. Überlastkontrolle in existierenden Systemen . . . . .	138

5.1.3. Fehlerbehandlung . . . . .	139
5.1.4. Topologien . . . . .	141
5.1.5. Fazit . . . . .	143
5.2. Medienübertragung in OCTOPUS und BRAVIS . . . . .	143
5.3. Ein Ansatz zur QoS-Sicherung der Audiodaten . . . . .	145
5.3.1. Einflussgrößen auf die Sendeleistung von Audioströmen . . . . .	146
5.3.2. Anpassung der Sendeleistung eines Audiostroms . . . . .	147
5.4. Ein Ansatz zur QoS-Sicherung der Videodaten . . . . .	151
5.4.1. Einflussgrößen auf die Sendeleistung von Videoströmen . . . . .	151
5.4.2. Ein Ansatz zur Bandbreitenanpassung durch Abstimmung . . . . .	153
5.4.3. Bewertung des Ansatzes der Bandbreitenanpassung . . . . .	161
5.4.4. Bandbreitenanpassung in BRAVIS . . . . .	167
5.4.5. Skalierung von Videoströmen . . . . .	170
5.5. Zusammenfassung . . . . .	174
<b>6. Konzepte der Implementierung</b>	<b>177</b>
6.1. Auswahl der Systemplattform . . . . .	177
6.2. Gruppenkoordinator im OCTOPUS . . . . .	178
6.2.1. Signalisierung . . . . .	178
6.2.2. Medienübertragung . . . . .	181
6.3. Endsysteme im OCTOPUS . . . . .	183
6.4. Konzepte des BRAVIS . . . . .	184
6.5. Zusammenfassung und Einschätzung der Implementierung . . . . .	187
<b>7. Zusammenfassung und Ausblick</b>	<b>189</b>
<b>Anhang</b>	<b>195</b>
<b>A. Automatengraph des GCP<sub>IP</sub></b>	<b>195</b>
<b>B. Fehlerfälle des GCP<sub>IP</sub></b>	<b>197</b>
B.1. Konferenzinitialisierung . . . . .	197
B.2. Tokenübergabe . . . . .	198
B.3. Datenübertragung . . . . .	199
B.4. Verlassen der Konferenz . . . . .	201

<b>C. GCP<sub>B</sub></b>	<b>205</b>
<b>Literaturverzeichnis</b>	<b>207</b>

---

# Abbildungsverzeichnis

---

1.1. Anforderungen für die Gestaltung von Videokonferenzsystemen .	7
1.2. Überblick über die Arbeit . . . . .	10
2.1. Wachstumsprognosen auf dem Videokonferenzmarkt . . . . .	16
2.2. Beispiel für ein Studiosystem (entnommen [118]) . . . . .	17
2.3. Beispiel für ein Roll-About-System [153] . . . . .	17
2.4. Beispiel einer Set-Top-Box [153] . . . . .	18
2.5. Beispiel für ein Desktop-System (entnommen [166]) . . . . .	19
2.6. Allgemeine Struktur eines Videokonferenzsystems . . . . .	21
2.7. Allgemeiner Ablauf der Medienverarbeitung . . . . .	27
2.8. Funktionsweise der räumlichen Skalierung . . . . .	31
2.9. Möglichkeiten der temporalen Skalierung mit M-JPEG . . . . .	32
2.10. Möglichkeiten der temporalen Skalierung mit H.261 (Ebene 1: I-Frames, Ebene 2: P-Frames) . . . . .	33
2.11. Möglichkeiten der temporalen Skalierung mit MPEG (Ebene 1: I-Frames, Ebene 2: B-Frames, Ebene 3: P-Frames) . . . . .	33
2.12. Das Kodierungsprinzip des G.722 Standards . . . . .	35
2.13. Vereinfachter Aufbau eines XCast-Paketes . . . . .	40
2.14. Mesh First Ansatz für Overlay Multicast . . . . .	42
2.15. Tree First Ansatz für Overlay Multicast . . . . .	43
2.16. Das Prinzip des Layered Multicast in RENDEZVOUS . . . . .	49
2.17. Beispiel einer zentralisierten Architektur eines Videokonferenz- systems . . . . .	54
2.18. Beispiel einer dezentral organisierten Topologie eines Videokon- ferenzsystems . . . . .	56
2.19. Beispiel einer hierarchischen Gruppenkommunikation . . . . .	58

2.20. Die H.323-Struktur am Beispiel des DFN-Videokonferenzdienstes DFNVideoConference nach [97] . . . . .	59
3.1. Die Struktur des OCTOPUS-Systems . . . . .	64
3.2. Ebenenstruktur von OCTOPUS . . . . .	65
3.3. Architektur des Gruppenkoordinators . . . . .	68
3.4. Globaler Konferenzaufbau . . . . .	71
3.5. Beitritt eines Teilnehmers . . . . .	72
3.6. Änderung der Sprecherwarteschlange . . . . .	73
3.7. Verlassen einer Gruppe . . . . .	74
3.8. Ausfall einer Gruppe . . . . .	75
3.9. Die BRAVIS Topologie . . . . .	79
3.10. BRAVIS-Architektur . . . . .	81
4.1. Das Prinzip des GCP <sub>ATM</sub> . . . . .	92
4.2. Der GCP <sub>IP</sub> -Protokollstack . . . . .	94
4.3. Erfolgreiche Initialisierung einer Konferenz . . . . .	97
4.4. Erfolgreicher Beitritt eines Teilnehmers zur Gruppe . . . . .	97
4.5. Erfolgreiche Übertragung eines Datenpakets . . . . .	98
4.6. Austritt eines Teilnehmers aus der Gruppe . . . . .	98
4.7. Prinzip des Datenaustauschs im GCP <sub>IP</sub> . . . . .	100
4.8. Initialisierungsphase der Konferenz . . . . .	102
4.9. Zeitablaufdiagramm für den Beitritt eines Teilnehmers . . . . .	104
4.10. Optimierung der Paketlaufzeiten im GCP <sub>IP</sub> -Ring . . . . .	105
4.11. Rekonfiguration des logischen Rings bei einem Beitritt eines neuen Teilnehmers . . . . .	106
4.12. Tokenübergabe im GCP <sub>IP</sub> . . . . .	108
4.13. Zeitablaufdiagramm einer Datenübertragung im GCP <sub>IP</sub> . . . . .	109
4.14. Zeitablaufdiagramm für das Verlassen der Gruppe . . . . .	110
4.15. Fehlererkennung . . . . .	112
4.16. Messungen der Verlustrate zu verschiedenen Standorten . . . . .	114
4.17. Messung und Approximation der Tokenumlaufzeit $t_{\text{RTT-Token},1}$ . . . . .	117
4.18. Messung und Approximation der Datenübertragungszeiten $t_{\text{DAT},1}$ . . . . .	118

---

4.19. Approximation der maximalen Gesamtverzögerung und Messung der tatsächlichen, durchschnittlichen Verzögerung durch $GCP_{IP}$ . . . . .	119
4.20. Messung der Verzögerungen zu verschiedenen Standorten . . . . .	120
4.21. Maximaler Durchsatz von $GCP_{IP}$ . . . . .	123
4.22. Bandbreitenabschätzung für den Signalisierungsoverhead ohne Anpassung . . . . .	125
4.23. Bandbreitenabschätzung für den Signalisierungsoverhead mit Anpassung . . . . .	126
4.24. Anforderung des Rederechts im OCTOPUS-System . . . . .	128
4.25. Anordnung für die Abschätzung der Gesamtverzögerung der Signalisierung im OCTOPUS-System . . . . .	129
5.1. Prinzip der Forward Error Correction . . . . .	140
5.2. Das Konzept der Audiokommunikation . . . . .	149
5.3. Konzept der Audioübertragung in IVS . . . . .	150
5.4. Beispielszene einer Auswahl in einer OCTOPUS-Sitzung . . . . .	155
5.5. Der auswahlbasierte Bandbreitenzuweisungsalgorithmus . . . . .	158
5.6. Bandbreitenzuweisung basierend auf der Abstimmung . . . . .	160
5.7. Übertragungsverzögerungen für verschiedene Paketgrößen . . . . .	161
5.8. Verzögerung, Framerate und Bandbreite auf der Verbindung von Hong Kong nach Cottbus . . . . .	163
5.9. Verzögerung, Framerate und Bandbreite auf der Verbindung von Cottbus nach Hong Kong . . . . .	164
5.10. Bandbreitenanpassung ohne einen Sprecher in der Gruppe . . . . .	165
5.11. Bandbreitenanpassung mit einem Sprecher in der Gruppe . . . . .	166
5.12. Bandbreitenanpassung mit einer Änderung der Prioritäten zwischen den Teilnehmern $P1$ und $P2$ sowie einer Änderung des Sprechers von $S_p$ zu $P3$ . . . . .	167
5.13. Bildung von künstlichen Untergruppen in BRAVIS . . . . .	169
5.14. Skalierungsschemen für Videostreams . . . . .	172
6.1. Warteschlangenkonzept des $GCP_{IP}$ . . . . .	180
6.2. Softwarearchitektur des ATM-RTP-Gateway für den Gruppenkoordinator . . . . .	182

A.1. Automatengraph des $GCP_{IP}$ . . . . .	195
B.1. Verlust des JR . . . . .	197
B.2. Verlust des JC . . . . .	198
B.3. Verlust des CA (a) vom Initiator und (b) vom Eingeladenen . . .	199
B.4. Verlust des TF . . . . .	200
B.5. Verlust des TF_ACK . . . . .	200
B.6. Verlust des DT . . . . .	201
B.7. Verlust des LR . . . . .	202
B.8. Verlust des LC . . . . .	203
B.9. Verlust des LA . . . . .	203
C.1. Messung der tatsächlichen, durchschnittlichen Verzögerung durch GCP <sub>B</sub> . . . . .	206



---

## Tabellenverzeichnis

---

2.1. Analyse bekannter Videokonferenzsysteme hinsichtlich einiger wichtiger Merkmale . . . . .	60
4.1. Tokenumlaufzeiten des $GCP_{ATM}$ nach [10] . . . . .	130
5.1. Anzahl der Teilnehmer, die abhängig von der Bildschirmauflösung und der Videobildgröße auf dem Bildschirm Platz finden .	154
6.1. Schnittstellenbeschreibung der OCTOPUS-Signalisierung (gesendete Daten) . . . . .	181
6.2. Schnittstellenbeschreibung der OCTOPUS-Signalisierung (empfangene Daten) . . . . .	181



---

## Glossary

---

### A

**AAL** ATM Adaptation Layer.

**ADPCM** Adaptive Differential Pulse Code Modulation.

**AF** Assured Forwarding.

**AGB** Allgemeine Geschäftsbedingung.

**ATM** Asynchronous Transfer Mode.

**AVO** Audio Visual Objects.

### C

**CD** Compact Disc.

**CIF** Common Interchange Format.

**CODEC** Coder and Decoder.

**CS-ACELP** Conjugate Structure Algebraic Codebook Excitation Linear Prediction.

### D

**DBA** Dynamic Bandwidth Allocation.

**DCT** Discrete Cosine Transformation.

**DiffServ** Differentiated Services.

**DTV** Digital Television.

**DVD** Digital Versatile Disc.

## **E**

**EDF** Earliest—Deadline—First.

**EF** Expedited Forwarding.

**Email** Electronic Mail.

## **F**

**FEC** Forward Error Correction.

**FIFO** First—In—First—Out.

## **G**

**G-WiN** Deutsches Wissenschaftsnetzwerk des DFN.

**GCP** Group Communication Protocol.

**GCP<sub>ATM</sub>** Group Communication Protocol over ATM.

**GCP<sub>B</sub>** Group Communication Protocol for BRAVIS.

**GCP<sub>IP</sub>** Group Communication Protocol over IP.

## **I**

**IntServ** Integrated Services.

**ISDN** Integrated Services Digital Network.

**ISO** International Standardization Organisation.

**ISP** Internet Service Provider.

**ITU-T** International Telecommunication Union – Telecom Standardization.

**J**

**JPEG** Joint Picture Coding Expert Group.

**L**

**LD–CELP** Low–Delay Code Excited Linear Prediction.

**M**

**M–JPEG** Motion JPEG.

**Mbone** Multicast Backbone over the Internet.

**MC** Multipoint Controller.

**MCU** Multipoint Control Unit.

**MP** Multipoint Processor.

**MP–MLQ** MultiPulse–Maximum Likelihood Quantizer.

**MP3** MPEG Layer 3.

**MPEG** Motion Picture Expert Group.

**MPLS** Multi–Protocol–Label–Switching.

**N**

**NTP** Network Time Protocol.

**P**

**PC** Personal Computer.

**PCM** Pulse Code Modulation.

**PHB** Per Hop Behavior.

**Q**

**QCIF** Quater Common Interchange Format =  $\frac{1}{4}$  CIF.

**QoS** Quality of Service.

**QQCIF** Quarter Quater Common Interchange Format =  $\frac{1}{16}$  CIF.

**R**

**RSVP** Ressource Reservation Protocol.

**RTCP** Realtime Transport Control Protocol.

**RTP** Realtime Transport Protocol [137].

**S**

**SAP** Session Announcement Protocol [53].

**SDP** Session Description Protocol [54].

**SID** Session ID, Session Identifier.

**SIP** Session Initiation Protocol [55].

**SQCIF** Small Quater Common Interchange Format =  $\frac{1}{16}$  CIF.

**ST** Internet Stream Protocol.

**T**

**TCP** Transport Control Protocol.

**U**

**UDP** Universal Datagram Protocol.

**USB** Universal Serial Bus.

**V**

**VoIP** Voice over IP.

**W**

**WWW** World Wide Web.

**X**

**XCast** eXplicit Multicast.

*GLOSSARY*

---



---

# 1. Einleitung und Problemstellung

---

## 1.1. Motivation

„A man in New York makes a phone call. But instead of just holding a handset to his head, he watches a small screen on the phone. He and his wife, in Florida, see each other as they chat, thanks to tiny camera on each phone. They do not only talk. They interact by gesturing and expressing themselves just as they would in person. They have arrived in the future: 1964.“

*(International Harald Tribune, Friday, November 29, 2002)*

Seit der Erfindung des Telefons besteht der Traum der Menschen, nicht nur die Sprache sondern auch das Bild des Kommunikationspartners zu übertragen. Dieser Traum ist im Jahre 1936<sup>1</sup> Realität geworden [121, 169]. Trotzdem wird fast 70 Jahre nach der Entwicklung des ersten Bildtelefons die Videotelefonie kaum genutzt. Videokonferenzen sind ebenso noch nicht so ein gebräuchliches Kommunikationsmittel wie beispielsweise WWW oder Email. Mit der Entwicklung von Mikroprozessoren, Computern und schnellen weltweiten Kommunikationsnetzwerken sind die technischen Voraussetzungen für die Nutzung von Videokonferenzdiensten heutzutage zunehmend gegeben. Immer mehr Nutzer verfügen über multimediafähige Computer. Durch die wachsenden Übertragungsraten im Internet wird die erforderliche Infrastruktur bereitgestellt. Waren Videokonferenzsysteme bisher bevorzugt an speziell ausgestattete Studios und kostenintensive Technik gebunden, so besteht mit der Verfügbarkeit multimedialer PCs und Workstations immer mehr die Möglichkeit, unmittelbar vom Arbeitsplatz aus an Videokonferenzen teilzunehmen.

---

<sup>1</sup>Der erste *Fernsehsprechdienst* zwischen Berlin und Leipzig wurde auf der Leipziger Messe am 1.3.1936 vorgestellt. 1964 präsentierte AT&T sein erstes Videofon auf der Weltausstellung in New York.

Beratungen unter Nutzung von Videokonferenzsystemen bieten eine Vielzahl von Vorteilen gegenüber „vor-Ort“ Beratungen. Sie sind leichter zu organisieren, wodurch sich Personen unterschiedlicher Standorte schneller und öfter zu einer Beratung treffen können. Somit können Entscheidungen über Projekt- und Produktentwicklungen schneller getroffen werden, und es kann flexibler auf Ereignisse reagiert werden. Bei der Nutzung von Videokonferenzsystemen muss nicht aus Kostengründen auf die Teilnahme von Personen an Beratungen verzichtet werden. Vielmehr kann zusätzlich qualifiziertes Fachpersonal hinzugezogen werden.

Ein wesentlicher Vorteil von Videokonferenzsystemen ist die Steigerung der Produktivität, da wertvolle Zeit und Kosten für Reisen eingespart werden können. Insbesondere interkontinentale Geschäftsreisen sind für die Beteiligten oft eine körperliche Belastung, beispielsweise durch das Jetlag. Es werden oft Tage zur Erholung benötigt. Es wurde sogar beobachtet, dass eine Beratung mittels Videokonferenzsystem effizienter ist als eine persönliche Beratung, weil die Konzentration der Teilnehmer beim Verfolgen der Beratung höher ist [135]. Personal in Schlüsselpositionen kann ebenfalls effizienter arbeiten, da sich Treffen mit Vertretern von unterschiedlichen Kontinenten an einem Tag realisieren lassen. Zunehmend sind auch Sicherheitsaspekte von Bedeutung. Das Risiko, auf der Reise in einen Unfall verwickelt zu werden, steigt. Besonders nach dem Terroranschlag vom 11. September 2001 ist die Nachfrage nach einem Videokonferenzsystem gestiegen [30].

Aber nicht nur in der Industrie sondern auch in der Medizin werden zunehmend Videokonferenzsysteme eingesetzt. Regelmäßige Beratungen von Spezialisten unterschiedlicher Krankenhäuser und Kliniken gehören zum Alltag, um Probleme und spezielle Erkrankungen zu diskutieren sowie um Erfahrungen auszutauschen. Durch ein Videokonferenzsystem ist es nicht nur möglich, statische Bilder, sondern auch Bildsequenzen zu übertragen, was für eine Ferndiagnose sehr vorteilhaft sein kann. Aber auch schwierige Operationen, bei dem ein Arzt durch einen über ein Videokonferenzsystem verbundenen Spezialisten unterstützt wird, sind keine Seltenheit mehr. Auch hier werden Reisekosten — vor allem aber die in der Medizin lebenswichtige Zeit — gespart.

Email-Services, News-Services und Telefonhotlines werden immer häufiger für verschiedenste Produkte angeboten, um bei auftretenden Problemen be-

hilffich zu sein. Jedoch fällt es den Betroffenen aufgrund der gestiegenen Komplexität von technischen Geräten zunehmend schwerer, das konkrete Problem detailliert in Worte zu fassen. Somit warten sie oft tagelang oder sogar vergebens auf eine Antwort, die das Problem auch nicht löst. Telefonhotlines sind zum Teil recht teuer und die verbale Beschreibung des Problems kann auch zu Missverständnissen führen. Mit einem Videokonferenzsystem können Bildsequenzen übertragen werden, die ein Problem schnell und detailliert beschreiben können. Somit können Probleme effizienter gelöst werden.

Die Aus- und Weiterbildung in Wirtschaft, Lehre, Forschung und im öffentlichen Dienst sind in der heutigen Zeit wichtiger de je, um mit neuen und sich verändernden Technologien umgehen zu können. Oft sind aber Bildungsangebote in der näheren Umgebung unzureichend. Angebote von entfernten Standorten sind jedoch nicht zeit- und kostengünstig zu erreichen. Somit bleibt bis heute nur der Weg über eine Fernausbildung, bei der Aufgaben und Probleme über den Postweg diskutiert werden. Videokonferenzsysteme bieten jedoch eine Möglichkeit, spontan Diskussionen mit dem Ausbilder als auch mit Studienkollegen zu führen, die durch Bilder und Bildsequenzen unterstützt werden. Des Weiteren können Lernende an Lehrveranstaltungen per Videokonferenz teilnehmen, um so ein besseres Verständnis für das Ausbildungsgebiet zu bekommen.

Im Rahmen der Globalisierung und der Entwicklung auf dem Arbeitsmarkt sind Familien zunehmend gezwungen, an verschiedenen Standorten zu leben und zu arbeiten. Aufgrund hoher Reisekosten ist es hier auch oft nicht möglich, sich wöchentlich oder monatlich zu besuchen. Zusätzlich zu einem Telefongespräch können mit einem Videokonferenzsystem Mimik und Gestik übertragen werden, was in einem privaten Gespräch oft eine große Rolle spielt.

Videokonferenzen können jedoch kaum psychologische Faktoren übertragen, die beim Kennenlernen von bisher unbekanntem Gesprächspartnern eine wesentliche Rolle spielen. So wird der erste Eindruck nicht durch das persönliche Auftreten, dem Geruch oder einem festen Händedruck bestimmt, sondern beschränkt sich auf die Stimme und den Bildausschnitt der Gesprächspartner. Es ist jedoch erwiesen, dass oft der erste persönliche Eindruck eines Menschen über Sympathie oder Antipathie beziehungsweise über Erfolg oder Misserfolg einer geschäftlichen Beratung. Deshalb sollte vor dem Einsatz eines Videokon-

ferenzsystems immer ein persönliches Zusammentreffen der Verhandlungspartner erfolgen.

### 1.2. Problemstellung

In der Vergangenheit basierte die Mehrzahl der Videokonferenzsysteme auf spezieller Hardware, die in dafür vorgesehenen Studios installiert wurde. Verschiedene Anbieter vermieten solche Studios, die in ihrer Anschaffung und ihrem Betrieb sehr kostenintensiv sind [60, 118]. Die hohe Nachfrage und die Reduzierung der Baugröße elektronischer Geräte führte zu kompakten Lösungen, die an kein Studio mehr gebunden sind, sondern in einen gewöhnlichen Fernsehschrank oder auf ein TV-Gerät passen [58, 122, 143, 153]. Diese Systeme, die zumeist den H.320 und den H.323 Standard unterstützen, eignen sich aber weniger für spontane Beratungen, sind zudem sehr preisintensiv und können kaum aktuellen Tendenzen und Neuerungen in der Technologie folgen.

Flexibler, kostengünstiger und spontan von jedem Arbeitsplatz ausführbar sind dagegen Desktop-Videokonferenzsysteme. Durch die ständige Weiterentwicklung der Rechentechnik, der Rechengeschwindigkeit und der Softwarealgorithmen sind moderne Computer in der Lage, multimediale Konferenzen durchzuführen. Deshalb geht der aktuelle Trend hin zu softwaregestützten Desktop-Videokonferenzsystemen.

Durch die Globalisierung werden Videokonferenzsysteme zunehmend für die weltweite Kommunikation von Geschäftspartnern, Forschungsgruppen, Lehrenden und Lernenden sowie Privatpersonen eingesetzt. Eine ausreichend gute Qualität der Video- und Audioübertragung kann jedoch nur durch Telefon-, ISDN- oder ATM-Netzwerke garantiert werden, woran wiederum sehr hohe Kosten gebunden sind. Die preisgünstige Alternative, das Internet, lässt aber noch keine garantierte Übertragungsqualität zu. Stattdessen teilen sich tausende Nutzer die Ressourcen des Netzwerks. Ansätze wie *Integrated Services (IntServ)* [21, 180] oder *Differentiated Services (DiffServ)* [15] versuchen Dienstgüte (*Quality of Service, QoS*) über das Internet zu sichern, konnten sich aber bis jetzt noch nicht weltweit durchsetzen. Videokonferenzen sind jedoch

Anwendungen, die hohe Bandbreiten<sup>2</sup> benötigen und somit das Netzwerk stark belasten oder sogar blockieren können. Populäre Videokonferenzsysteme wie die Mbone-Tools VAT [113], VIC [112], RENDEZVOUS [131] sowie Standardanwendungen wie Microsoft Netmeeting [102], SUNForum [149] und Whitepine CUSeeMe [174] überlassen es dem Nutzer, die Überlastung des Netzwerks zu verhindern. Dies erfordert jedoch vom Nutzer viele Erfahrungen im Bereich der Netzwerktechnik. Somit wird die Bedienbarkeit dieser Systeme erschwert. Selten hingegen sind automatische Anpassungsmechanismen zur Vermeidung von Überlastsituationen.

Ein weiteres Problem für die Gruppenkommunikation mit Videokonferenzsystemen über öffentliche, globale Netze ist die Heterogenität des Netzwerks. Einerseits kann die Netzwerkanbindung einzelner Endsysteme recht verschieden sein, beispielsweise von einem Anschluss mit analogem Modem (64 kbit/s) über DSL-1500 (1536 kbit/s) bis zur Standleitung (mehrere Mbit/s) reichen. Auf der anderen Seite können die Endsysteme von unterschiedlicher Leistungsfähigkeit sein, wie Workstations, Desktop-PCs, Notebooks oder Handheld-PCs. Gängige Videokonferenzsysteme unterstützen kaum heterogene Umgebungen, sondern überlassen es wiederum dem Nutzer, die optimale Qualität für alle Teilnehmer einzustellen. Automatische Mechanismen einiger Systeme, wie zum Beispiel VCON [166, 33] oder POLYCOM [122, 116], unterstützen nur Punkt-zu-Punkt-Verbindungen. Weiterhin existieren Systeme wie beispielsweise USMINT [47] und Ivs [163], die Ansätze zur automatischen Beeinflussung der Dienstgüte in Mehrteilnehmer-Videokonferenzen enthalten. Diese Lösungen richten sich jedoch nach dem leistungsschwächsten Teilnehmer der Gruppe. Dadurch bekommen die verbleibenden Teilnehmer eine geringere Qualität geboten, als sie verarbeiten können. Somit sinkt die Akzeptanz dieser Lösungen. Einzig RENDEZVOUS [131] bietet eine Lösung für heterogene Systemumgebungen, die jedoch individuell durch den Nutzer angepasst werden muss. Somit bleibt es bis heute hin schwierig, globale Videokonferenzen mit einer zufrieden stellenden Qualität für die Nutzer aufzusetzen.

---

<sup>2</sup>Der Begriff der Bandbreite wird in dieser Arbeit nicht im kommunikationstechnischen Sinne als Frequenzbereich, gemessen in Hertz (Hz), zwischen einer Minimal- und einer Maximalfrequenz benutzt. Stattdessen wird die Bandbreite als in der Literatur durchaus übliches Synonym für die Bitrate, gemessen in Bit/s, verwendet.

Die meisten Videokonferenzsysteme unterstützen einen offenen Charakter der Gruppenkommunikation, das heißt, die Anzahl der Teilnehmer ist nicht genau bestimmbar. Ein Sender weiß nicht genau, wer seine Daten empfängt. Die Teilnehmer können der Gruppe ohne Autorisierung beitreten. Solche Konferenzen werden besonders für öffentliche Veranstaltungen, öffentliche Vorlesungen, TV-Übertragungen oder Übertragungen von Space Shuttle Starts der NASA genutzt. Die Mbone-basierten Videokonferenzsysteme VAT [113], VIC [112], RENDEZVOUS [131] und USMINT [47] unterstützen offene Gruppen.

Die Mehrzahl der tagtäglichen Kommunikation wie Beratungen in Firmen, medizinische Konsultationen, Gerichtsverhandlungen, private Unterhaltungen und Hotline-Dienste haben jedoch einen geschlossenen Charakter. Geschlossene Mehrteilnehmer-Konferenzen besitzen einen begrenzten Teilnehmerkreis und dienen der Diskussion vertraulicher Themen, der gemeinsamen Lösung von Problemen oder der kooperativen Bearbeitung von Objekten. Die Teilnahme an der Konferenz basiert zumeist auf einer expliziten Einladung. Außen stehende Personen haben keinen direkten Zugriff auf die Kommunikation. Geschlossene Kommunikation ist mit dem Telefondienst vergleichbar. Ein weiteres Merkmal geschlossener Systeme ist, dass jeder Teilnehmer die multimedialen Daten der anderen Teilnehmer empfängt und verarbeitet. Eine geschlossene Mehrteilnehmer-Konferenz wird durch ATM-basierte Systeme, wie COVIS [12] oder HKUST [25], durch das proprietäre IP-Videokonferenzsystem DAVIKO [34] und durch Systeme nach H.320 für ISDN, H.321 für ATM, H.323 für IP und H.324 für das analoge Telefon [76, 74, 78, 79] unterstützt.

H.32x Systeme realisieren die geschlossene Gruppe durch einen zentralistischen Ansatz für die Verteilung der Medienströme und für das Gruppenmanagement. Zentrale Ansätze sind nicht frei skalierbar, da die zentralen Komponenten einen Flaschenhals der Kommunikation bilden und zu einem so genannten *Single-Point-of-Failure* werden können. Des Weiteren besteht immer eine Abhängigkeit von der zentralen Komponente und von dessen Hersteller beziehungsweise Provider. In globalen Umgebungen entstehen Schwierigkeiten, insbesondere wenn sich die zentrale Komponente an einem weit entfernten Standort befindet.

Dezentrale Ansätze, wie beispielsweise die Mbone-Tools, USMINT, RENDEZVOUS, DAVIKO und COVIS, verteilen Aufgaben des Managements und

der Medienverarbeitung auf die Endsysteme. Sie vermeiden somit einen Flaschenhals der Kommunikation und sind weniger fehleranfällig. Des Weiteren sind dezentrale Ansätze nicht von Providern, Netzwerkstrukturen oder Netzwerkkomponenten abhängig. Ein weiterer Vorteil der dezentralen Ansätze liegt in der Spontaneität eine Videokonferenz aufzubauen, da keine übergeordnete Instanz den Zugriff auf Ressourcen, wie einen zentralen Server, regeln muss. Dezentrale Ansätze erfordern jedoch einen Mehraufwand im Management der Gruppe, da jeder Teilnehmer den aktuellen Status der Gruppe kennen muss.

Aus den Unzulänglichkeiten der existierenden Lösungen folgt die Frage, die sich diese Arbeit stellt, wie geschlossene Videokonferenzen über öffentliche Netzwerke realisiert werden können. Um dieses Problem lösen zu können,

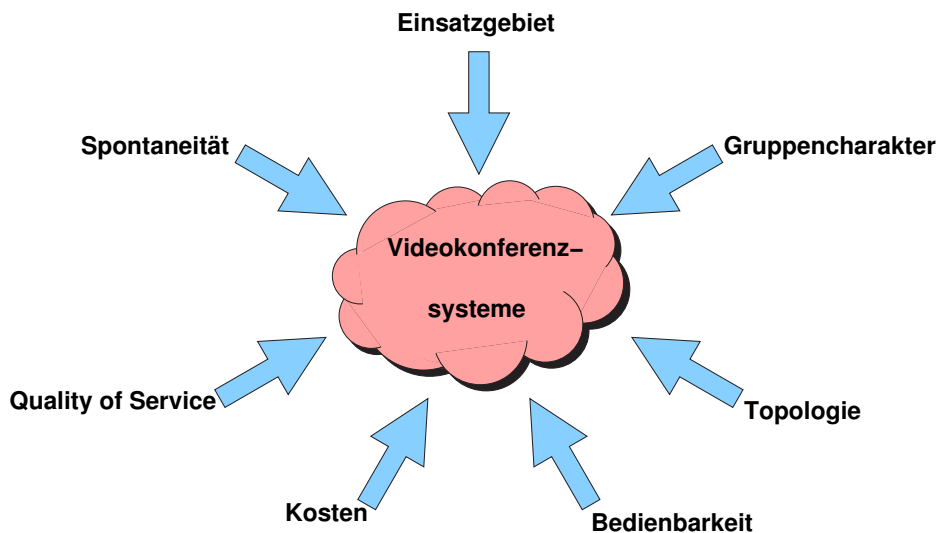


Abbildung 1.1.: Anforderungen für die Gestaltung von Videokonferenzsystemen

müssen die in Abbildung 1.1 gezeigten Anforderungen für die Gestaltung von Videokonferenzsystemen berücksichtigt werden.

**Einsatzgebiet:** Das Videokonferenzsystem sollte für globale Konferenzen über öffentliche Netzwerke wie dem Internet geeignet sein. Dies erfordert Mechanismen für die Anpassung an Leistungsengpässe und Unterstützung heterogener Systemumgebungen.

**Spontaneität:** Im Moment liegt die Vorbereitungszeit für eine Videokonferenzsitzung im Bereich von Stunden bis Tagen. Spontane Sitzungen sind

kaum möglich. Dies ist jedoch ein wichtiges Kriterium, damit die Nutzung von Videokonferenzsystemen eine breitere Akzeptanz findet.

**Gruppencharakter:** Da die Mehrzahl der menschlichen Kommunikation einen geschlossenen, vertraulichen Charakter hat, muss eine neue Lösung geschlossene Gruppen unterstützen, ähnlich dem Telefondienst.

**Quality of Service:** Die Dienstgüte muss weitestgehend stabil gehalten werden. Leider steht dies im Widerspruch zu einer globalen und kostengünstigen Kommunikation. Alternativ sollten daher Anpassungsmechanismen zur Vermeidung von Leistungsengpässen integriert werden.

**Topologie:** Die Topologie eines Videokonferenzsystem stellt ein besonderes Problem dar, da zum einen Engpässe der Kommunikation und Single-point-of-failures, aber zum anderen zu komplexe Steuerungsmechanismen der Konferenz vermieden werden müssen.

**Kosten:** Die Kosten für ein Videokonferenzsystem müssen sowohl in der Anschaffung als auch im Betrieb gering sein. Ansonsten wiegen sich die reduzierten Reisekosten, die erhöhte Effizienz der Arbeit und die erhöhte Sicherheit nicht auf.

**Bedienbarkeit:** Videokonferenzsysteme müssen einfach zu bedienen sein. Es sollten dem Nutzer automatische Mechanismen für die Regulierung der Dienstqualität zur Verfügung stehen, die allen Teilnehmern automatisch die bestmögliche Qualität während der gesamten Konferenz bereitstellt.

### 1.3. Zielsetzung und Aufbau der Arbeit

In der vorliegenden Arbeit werden Prinzipien für die Gestaltung von geschlossenen Videokonferenzsystemen für globale öffentliche Netzwerke untersucht. Beratungen, Meetings und Konsultationen, wie sie im täglichen Leben vorkommen, haben einen geschlossenen Charakter. Globale Videokonferenzen werden durch die zunehmende Globalisierung des Weltmarktes notwendig. Die Arbeit verfolgt hierbei zwei Ansätze.

Der erste Ansatz untersucht den Aufbau von globalen Konferenzen über den Verbund lokaler Videokonferenzsysteme, um vorhandene lokale Systeme auch



für globale Konferenzen nutzen zu können. Damit werden einerseits Investitionen erhalten und andererseits entfällt die Einarbeitung der Nutzer in ein anderes neues System. Ein wesentliches Ziel ist die Untersuchung der Möglichkeiten für die Erhaltung und Sicherung der Dienstqualität in globalen öffentlichen Umgebungen, um die in den lokalen Systemen gewohnte Dienstgüte zu gewährleisten.

Der zweite Ansatz beschäftigt sich daraus abgeleitet mit der Gestaltung von Videokonferenzsystemen für geschlossene Gruppen im Internet. Der Fokus liegt dabei auf einem dezentralen Peer-to-Peer-Konzept, durch den, im Vergleich zu existierenden Systemen, kostengünstigere Lösungen ermöglicht werden sollen. Diese besitzen eine höhere Flexibilität und sollen insbesondere einen spontanen Aufbau von Konferenzen sowie die Mobilität der Anwender unterstützen.

Um beide Zielstellungen erreichen zu können, muss die Frage nach dem Konferenzmanagement für geschlossene Gruppen, nach der optimalen Topologie für die Kommunikation und nach den Möglichkeiten der Dienstgütesicherung der Echtzeitdaten diskutiert werden. Dazu ist eine Analyse vorhandener Konzepte für die Verwaltung einer Konferenz, für die Echtzeitverarbeitung der multimedialen Daten und für die Übertragung der Management- und Echtzeitdaten über öffentliche Netzwerke notwendig.

Die Arbeit ist wie folgt gegliedert (siehe Abbildung 1.2): Kapitel 2 gibt zunächst einen Überblick über die Grundlagen der Gestaltung von Videokonferenzsystemen. Ausgehend von einer allgemeinen Architektur von Videokonferenzsystemen werden die wichtigsten Komponenten das Konferenz-Management und die Medienverarbeitung erläutert. Prinzipien der Echtzeitübertragung von Daten und vorhandene Netzinfrastrukturen werden ebenfalls diskutiert. Darauf aufbauend wird eine Diskussion über die Kriterien für die Gestaltung von Videokonferenzsystemen geführt. Das Kapitel schließt eine zusammenfassende Betrachtung und Bewertung von bisherigen Lösungen in Bezug auf die Unterstützung globaler, geschlossener Videokonferenzen ab.

Aufbauend auf dieser Diskussion werden im 3. Kapitel die zwei entwickelten Ansätze, OCTOPUS und BRAVIS, vorgestellt, die geschlossene, globale Mehrteilnehmerkonferenzen unterstützen. OCTOPUS stellt ein Verbundsystem dar, das lokale Videokonferenzsysteme verschiedener Architektur zusammenschalten gestattet. BRAVIS vereint entwickelte Konzepte des OCTOPUS-Ansatzes in

## 1. Einleitung und Problemstellung

---

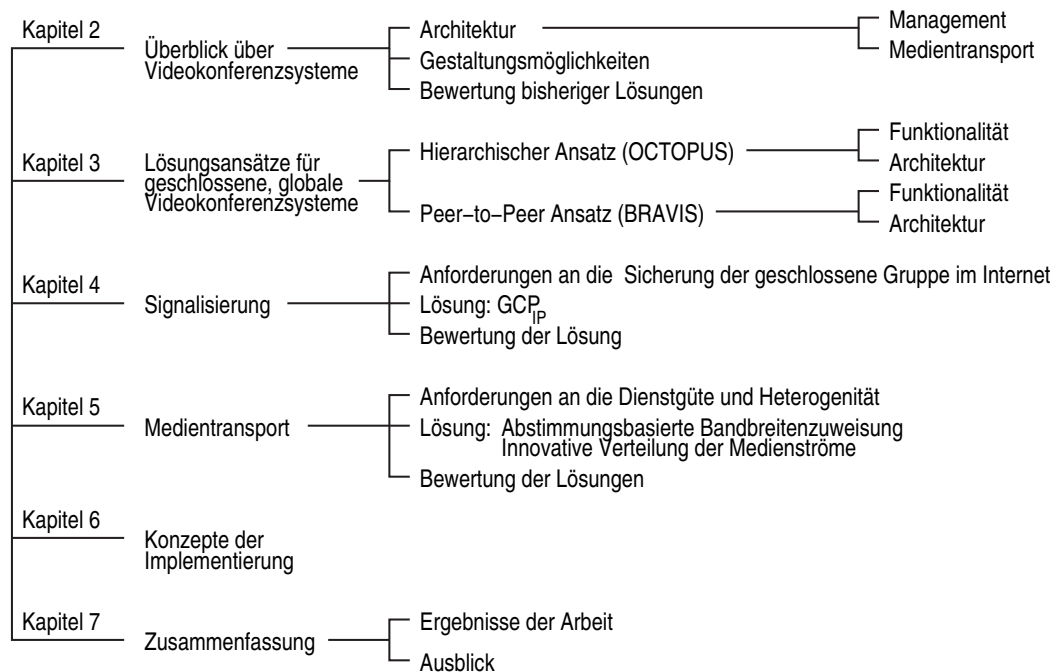


Abbildung 1.2.: Überblick über die Arbeit

einem eigenständigen Videokonferenzsystem auf der Basis von IP-Netzwerken. In den folgenden Kapiteln werden grundlegende Überlegungen zur Gestaltung von Videokonferenzsystemen für geschlossene Konferenzen anhand dieser Systeme diskutiert.

Kapitel 4 beschäftigt sich mit der Signalisierung von Videokonferenzsystemen in einem globalen Einsatzgebiet, wobei dessen Aufgaben und die Anforderungen an diese diskutiert werden. Dabei werden vorhandene Ansätze auf ihre Tauglichkeit zur Erfüllung der Signalisierungsaufgaben untersucht. Diese Untersuchung führt zu der Notwendigkeit der Herausarbeitung eines eigenständigen Gruppenkommunikationsprotokolls für die Realisierung einer Signalisierung in geschlossenen, globalen Videokonferenzsystemen. Das vorgeschlagene Protokoll wird spezifiziert und anschließend bewertet. Der Abschluss dieses Kapitels bildet die Diskussion um die Einsatzfähigkeit des neuartigen Protokolls, sowohl in OCTOPUS als auch in BRAVIS.

Das 5. Kapitel ist dem Kernzweck eines Videokonferenzsystems, der Medienübertragung, gewidmet. Hierbei wird auf die Besonderheiten der Video- und Audioübertragung am Beispiel von OCTOPUS und BRAVIS eingegangen,

und es werden Ansätze zur Unterstützung der Dienstgüte und Heterogenität in globalen Netzwerken beschrieben und diskutiert.

In Kapitel 6 folgt eine Beschreibung der Implementierung beider Systeme. Dabei werden spezielle Entwurfsentscheidungen diskutiert.

Das Kapitel 7 enthält eine Zusammenfassung der in dieser Arbeit entwickelten Ansätze und Strategien. Ein Ausblick auf noch zu lösende Probleme und Erweiterungen der beschriebenen Konzepte wird abschließend gegeben.

## *1. Einleitung und Problemstellung*

---

---

## 2. Gestaltungsprinzipien von Videokonferenzsystemen

---

Videokonferenzen gelten als eine populäre multimediale Anwendung in modernen Hochleistungsnetzen. Trotz langjähriger Forschungs- und Entwicklungsarbeiten werden Videokonferenzen nicht so häufig im Alltag genutzt wie Email oder das World Wide Web. Die technischen Voraussetzungen für die Nutzung von Videokonferenzdiensten sind heutzutage zunehmend gegeben. Immer mehr Nutzer verfügen über multimediafähige Computer. Durch die wachsenden Übertragungsraten im Internet wird die erforderliche Infrastruktur bereitgestellt. Waren Videokonferenzsysteme bisher bevorzugt an speziell ausgestattete Studios und an eine spezielle, kostenintensive Technik gebunden, so besteht mit der Verfügbarkeit multimedialer PCs und Workstations immer mehr die Möglichkeit, unmittelbar vom Arbeitsplatz aus an Videokonferenzen teilnehmen.

Dieses Kapitel gibt einen umfassenden Überblick über den Stand der Technik im Bereich der Videokonferenzen. Vorhandene Lösungen werden dabei auf verschiedene Eigenschaften hin untersucht. Den Abschluss dieses Kapitels bildet die Herausarbeitung der Zielstellung dieser Arbeit.

### 2.1. Anwendungsbereiche

In der Vergangenheit kamen Videokonferenzsysteme vorrangig für das Entertainment zum Einsatz, wobei Space Shuttle Starts der NASA, öffentliche Veranstaltungen oder TV-Kanäle übertragen wurden. Heute fällt diese Anwendung in den Bereich des *Media-Streaming* und zählt nur noch als Randgebiet der eigentlichen Videokonferenzanwendungen. Eine Echtzeitkommunikation ist

Entertainment
---------------

## 2. Gestaltungsprinzipien von Videokonferenzsystemen

---

hierbei nicht unbedingt erforderlich, da die Empfänger kein Feedback an den Sender geben müssen.

*geschäftliche Beratungen*

Um Informationen, Arbeitsergebnisse, neue Aufgaben und Ziele auszutauschen, kommen asynchrone Kommunikationsmittel wie Fax, Email, FTP (File Transfer Protocol) oder Briefpost zum Einsatz. Für einen synchronen Gedankenaustausch eignet sich das Telefon, wobei jedoch audiovisuelle und textuelle Informationen kaum ausgetauscht werden können. Deshalb führen Unternehmen der Wirtschaft und des öffentlichen Dienstes regelmäßig Beratungen und Projekttreffen durch. In großen und mittleren Unternehmen sind die Mitarbeiter oft auf mehrere Standorte verteilt. Kooperierende Unternehmen und Universitäten sind zumeist ebenfalls räumlich voneinander getrennt. Somit entstehen hohe Reisekosten und Ausfallzeiten für die Reise selbst. Durch den Einsatz eines Videokonferenzsystems können audiovisuelle, textuelle und graphische Informationen synchron ausgetauscht werden, ohne dass Reisekosten und Ausfallzeiten entstehen.

*medizinische Konsultationen*

Immer häufiger nutzen auch Ärzte ein Videokonferenzsystem, um Informationen auszutauschen, um Ferndiagnosen zu stellen oder gemeinsam ein medizinisches Problem zu lösen. Aufgrund der regionalen Entfernung zwischen den Spezialisten konnten bisher informative Diskussionen nur per Telefon, Briefpost, Fax, Email oder auf Tagungen geführt werden. Eine spontane Diskussion war nicht möglich und eine Tagung verursachte hohe Kosten sowie Ausfallzeiten. Aus diesen Gründen haben Videokonferenzsysteme verstärkt Einzug in Kliniken und Krankenhäuser gehalten.

*Telesupport*

Im täglichen Leben werden immer häufiger Telefon-Hotlines eingerichtet, um Unterstützung für technische, medizinische oder sonstige Probleme zu bieten. Aber oft reicht die Kommunikation mittels der Sprache nicht aus, um das Problem eindeutig zu beschreiben. Somit bleiben Konsultationen vor Ort oder das Einschicken eines technischen Gerätes nicht aus. Durch Hinzunahme des Mediums Video können jedoch Probleme visuell veranschaulicht werden, defekte Stellen eines technischen Gerätes gezeigt werden, das Krankheitsbild einem Arzt übertragen werden und so weiter. Videokonferenzsysteme können somit auch in diesem Bereich Fahrtkosten, Reparaturkosten und lange Wartezeiten vermeiden. Zudem können mehrere unabhängige Institutionen an mehreren Standorten gleichzeitig zu einem gestellten Problem Stellung nehmen.

Professoren und Dozenten werden oft gebeten, Gastvorlesungen an anderen Universitäten zu halten, um das Wissen der Universitäten untereinander auszutauschen. Neben der Einsparung der Reisekosten, bietet sich der Vorteil, dass eine Lehrveranstaltung gleichzeitig an mehreren Standorten verfolgt werden kann. Auf der anderen Seite bieten Videokonferenzsysteme die Möglichkeit, dass Lernende das Angebot an Lehrveranstaltungen auch von anderen Standorten für sich nutzen können. Videokonferenzen können aber auch für andere Arten von Lehrveranstaltungen, wie zum Beispiel Übungen und Seminare, eingesetzt werden. Hierbei bietet sich die Möglichkeit der interaktiven und Campus übergreifenden Diskussion.

*Teleteaching*

Durch die zunehmende Globalisierung des Arbeitsmarktes sind Familien gezwungen, auch weit entfernt voneinander zu arbeiten und zu leben. Das Telefon bietet zwar die Möglichkeit des Gedankenaustausches, aber oft ist auch ein visueller Eindruck der Gestik und Mimik der Gesprächspartner gewünscht. Videokonferenzsysteme werden eines Tages auch in diesem Bereich zunehmend genutzt werden.

*private Unterhaltung*

Der Anstieg des Umsatzes für Videokonferenzsysteme wird mit einem gesteigerten Sicherheitsbedürfnis nach dem Terroranschlag vom 11. September 2001 begründet. [30] und [170] berichten, dass der Anteil von persönlichen Geschäftstreffen von 54% auf 45% gesunken ist, aber dafür Web-, Video- und Audiokonferenzen um 62%, 25% und 11% gestiegen sind. Auch das Bewusstsein der Menschen über die neuen Kommunikationsmittel und -medien ist nach dem 11. September gestiegen. Der Umsatzwachstum für Videokonferenzsysteme wird nach Einschätzung mehrerer unabhängiger Marktforschungsinstitute [31, 98, 170, 172, 177] in den nächsten Jahren weiter ansteigen (siehe Abbildung 2.1).

*aktueller Trend*

## 2.2. Arten von Videokonferenzsystemen

Videokonferenzsysteme können ihrem Aufbau nach in vier Kategorien unterschieden werden: Studiosysteme, die fest in einen Raum beziehungsweise einem Studio installiert sind, Roll-About-Systeme, die einem rollbaren Studio-System entsprechen, Set-Top-Boxen, als tragbare Variante eines Studiosys-

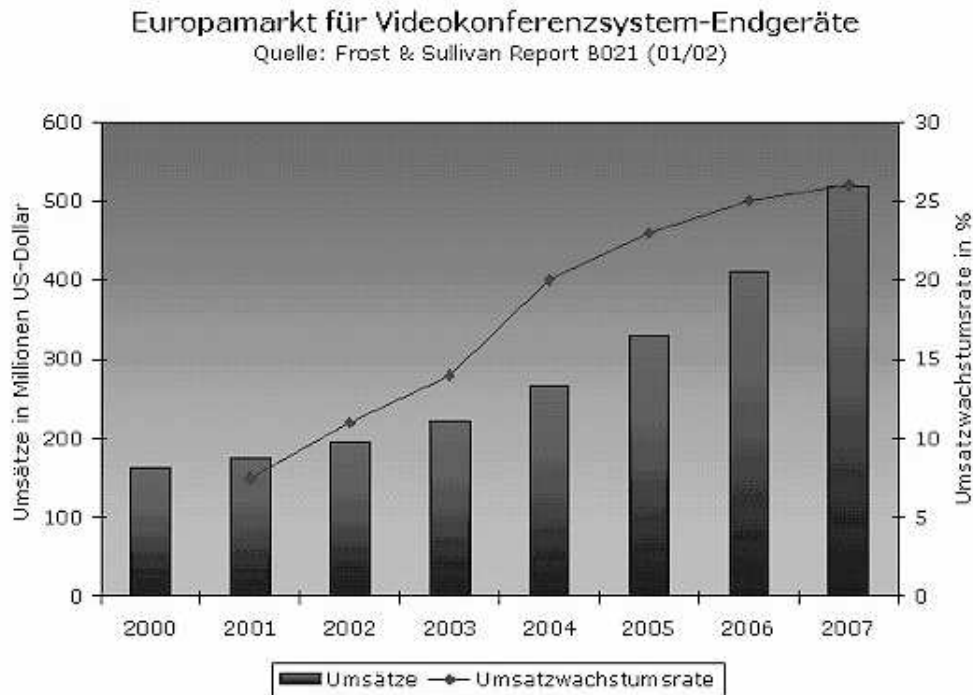


Abbildung 2.1.: Wachstumsprognosen auf dem Videokonferenzmarkt

tems, und Desktopsysteme, die als eine Anwendung von jedem Personal Computer (PC) gestartet werden können.

### Studiosysteme

Studiosysteme sind fest installierte Hardwaresysteme, die mit Richtmikrofonen, Kameras, Dokumentenkameras sowie teilweise mit einem PC ausgestattet sind (siehe Abbildung 2.2). Sie bieten den Vorteil, dass die ausgewählten Komponenten optimal aufeinander abgestimmt und an die räumliche Umgebung angepasst werden können. So kann der Raum durch Scheinwerfer mit spezieller Farbe ähnlich einem Fotostudio optimal ausgeleuchtet werden. Dem gegenüber stehen entscheidende Nachteile. Die Anschaffungskosten und die laufenden Kosten eines solchen Systems sind enorm. Zudem ist die Ausnutzung eines Studiosystems zumeist gering. Um den Grad der Ausnutzung eines Studiosystems zu erhöhen, existieren Mietmodelle am Markt [60, 118]. So stellen Anbieter in diversen größeren Städten ein oder mehrere Videokonferenzstudios zur Verfügung, die von ortsansässigen Firmen genutzt werden können. Dies ist für Firmen preisgünstiger, als ein selbst eingerichtetes Studio





Abbildung 2.2.: Beispiel für ein Studiosystem (entnommen [118])

zu unterhalten. Studiosysteme sind in ihrer Flexibilität recht eingeschränkt, da sie räumlich gebunden sind und die Einsatzgebiete sich auf Gruppenbesprechungen beschränken. Einzelgespräche, wie sie in der Telefonie alltäglich vorkommen, sind zwar möglich, jedoch sehr verwaltungsaufwendig. Spontane Gespräche und Beratungen sind mit einem Studiosystem nicht möglich.



Abbildung 2.3.: Beispiel für ein Roll-About-System [153]

Etwas flexibler sind die kleineren beweglichen Roll-About-Systeme. Die in einen Rollschrank eingebauten Systeme bestehen zumeist aus einem oder zwei handelsüblichen Fernsehgeräten, einer Kamera sowie einem Steuergerät. Das

*Roll-About-  
Systeme*

## 2. Gestaltungsprinzipien von Videokonferenzsystemen

---

Steuergerät verfügt über diverse Schnittstellen zu Kameras, TV-Geräten, Audio-Geräten, PCs und Netzwerken (IP und ISDN). Es übernimmt die Steuerung der Konferenz und die Verarbeitung der Video- und Audiosignale. Die Hardwarekomponenten eines Roll-About-Systems sind wie bei einem Studio-system optimal aufeinander abgestimmt, jedoch ist eine Feinjustierung zur Anpassung an die Licht- und Klangverhältnisse eines Raumes kaum möglich und auch nicht unbedingt vorgesehen. Roll-About-Systeme sind ebenfalls nicht geeignet, spontan eine Konferenz aufzusetzen, da sie nicht an jedem Arbeitsplatz installiert werden können. Tandberg [153], Intel [58], Polycom [122], VCON [166] und SONY [143] bieten beispielsweise Roll-About-Systeme an.

### Set-Top-Boxen

Mit sehr kompakten und tragbaren Set-Top-Boxen (siehe Abbildung 2.4), die an ein handelsübliches TV-Gerät oder an einen Monitor angeschlossen werden, können ebenfalls Videokonferenzen gehalten werden. Sie beinhalten bereits eine Kamera und ein Mikrofon. Ansonsten verfügen sie über die Eigenschaften der Roll-About-Systeme. Um spontan eine Videokonferenz mit Set-Top-Boxen aufzusetzen, ist ein solches System an jedem Arbeitsplatz notwendig. Dies ist sehr teuer, und es wird eventuell zusätzlicher Platz für einen zweiten Monitor benötigt.



Abbildung 2.4.: Beispiel einer Set-Top-Box [153]

Studio- sowie Roll-About-Systeme haben einen gemeinsamen Nachteil. Sie sind, wie alle Hardwarelösungen, nicht flexibel in Bezug auf neue Technologien. So können nicht ohne Weiteres optimierte Kodierungstechniken, leistungsfähigere Signalisierungsprotokolle oder neuartige, Dienstgüte sichernde Netzwerkdienste ausgetauscht oder nachgerüstet werden. Oft ist eine preisintensive Neuanschaffung notwendig.

### Desktopsysteme

Desktopsysteme hingegen werden auf einen handelsüblichen PC installiert

und können somit vom Arbeitsplatz aus genutzt werden. Beispiele für Desktopsysteme sind die MBone-Tools [88, 112, 113, 157, 158], USMINT [47], RENDEZVOUS [95] und COVIS [11]. Aber auch Netmeeting [102] und CUSeeMe [174] als Vertreter des H.323-Standards sind Desktopsysteme. Ein Beispiel eines solchen Systems zeigt Abbildung 2.5. Der große Vorteil von Desktop-



Abbildung 2.5.: Beispiel für ein Desktop-System (entnommen [166])

Videokonferenzsystemen sind die geringen Anschaffungskosten, da eine Minimalconfiguration eines PCs mit einem USB-Anschluss (Universal Serial Bus) und einer angeschlossenen USB-Kamera sowie einer integrierten Soundkarte ausreicht, um per Bild und Ton an einer Konferenz teilnehmen zu können. Die Videokonferenzsoftware auf dem PC ist in der Lage, die Konferenz zu steuern, Audio- und Videodaten aufzubereiten und über diverse Netzwerkanschlüsse zu verteilen. Des Weiteren können Desktopsysteme flexibel auf neue Technologien reagieren und diese integrieren. Parallel dazu können die PCs für andere Applikationen, wie beispielsweise Büroanwendungen, genutzt werden. Somit erzielen Desktopsysteme einen höheren Nutzungsgrad als Studiosysteme. Durch den technologischen Fortschritt auch auf dem mobilen PC-Markt, sind Notebooks heutzutage ebenfalls in der Lage, an Videokonferenzen teilzunehmen. Die örtliche Beschränkung von Studiosystemen wird damit durch Desktopsysteme aufgehoben. Mit Desktopsystemen lassen sich spontan Videokonferenzen aufsetzen, ohne zuvor spezielle Hardware anzufordern. Es ist keine Reservierung eines Studios notwendig, und das Videokonferenzsystem kann von jedem Mitarbeiter genutzt werden.

Neben den genannten reinen Softwaresystemen existieren Desktop-Lösungen, die durch Hardware-Einsteckkarten unterstützt werden. Ein Beispiel dafür sind Systeme der Firma VCON [166]. Hardwarekarten sind jedoch vergleichsweise teuer, unterstützen oft nur ein Betriebssystem und sind oft auf einen Standard für die Video- und Audiokodierung beschränkt. Sie können somit kaum dem Stand der Technik folgen. Des Weiteren sind sie meist nur für eine Punkt-zu-Punkt-Kommunikation geeignet. Mehrteilnehmerkonferenzen können nur durch externe Komponenten aufgebaut werden, wodurch die Spontaneität des Konferenzaufbaus verloren geht.

Desktop-Videokonferenzsysteme haben sich bis heute noch nicht durchsetzen können, da die Leistungsfähigkeit der Endsysteme, der Netzwerke und der Netzwerkanbindungen erst in den letzten Jahren so weit gestiegen ist, dass multimediale Konferenzen möglich wurden. Aufgrund der geringen Anschaffungskosten, der Flexibilität, der Spontaneität und der hohen Verfügbarkeit geht der Trend jetzt zu Desktop-Systemen hin, die nur aus Softwarekomponenten bestehen.

### 2.3. Architektur von Videokonferenzsystemen

Videokonferenzsysteme bestehen im Kern aus den selben vier Grundelementen: der Schnittstelle zum Nutzer, dem Konferenzmanagement, den Komponenten für die Medienverarbeitung und der Schnittstelle zum Netzwerk für die Übertragung der Management- und Mediendaten. Abbildung 2.6 zeigt eine typische Anordnung dieser Komponenten.

*Nutzerschnitt-  
stelle*

Die Nutzerschnittstelle dient zur Interaktion zwischen dem System und dem menschlichen Nutzer mittels akustischer und visueller Signale. Akustische Schnittstellen in einem Videokonferenzsystem sind das Mikrofon zur Spracheingabe und die Lautsprecher zur Ausgabe von Sprache beziehungsweise von Warnmeldungen. Die visuelle Schnittstelle kann zur Ausgabe ein Bildschirm, ein Fernsehgerät oder ein Drucker und zur Eingabe eine Kamera oder ein Scanner haben. Hierbei können sowohl bewegte Bilder als auch Einzelbilder ausgetauscht sowie graphische Steuerelemente angezeigt werden. Eine weitere Schnittstelle zur Eingabe sind Tastatur und Maus, um die Konferenz zu steuern.

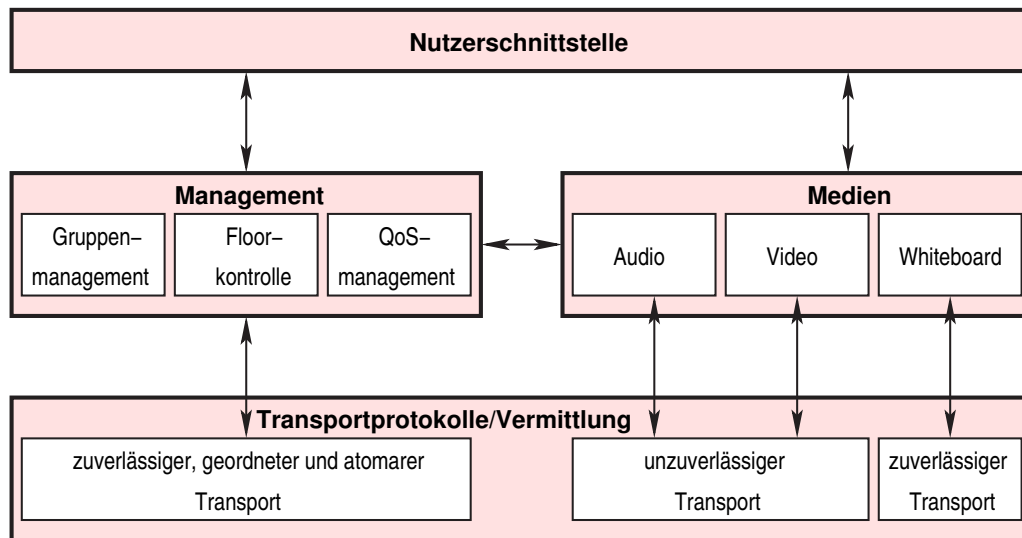


Abbildung 2.6.: Allgemeine Struktur eines Videokonferenzsystems

Die Gestaltung der visuellen Schnittstelle zum Benutzer wird in der Regel nach ergonomischen oder funktionellen Gesichtspunkten konzipiert, die in dieser Arbeit nicht betrachtet werden.

Auf die Komponenten für das Konferenzmanagement, für die Medienverarbeitung und auf die Netzwerkschnittstelle soll in den folgenden Abschnitten ausführlicher eingegangen werden.

## 2.4. Management eines Videokonferenzsystems

Die Aufgaben der Signalisierung können recht unterschiedlich sein. In dieser Arbeit werden die drei wichtigsten Komponenten des Konferenzmanagements:

- das Gruppenmanagement,
- das QoS-Management und
- die Floorkontrolle

diskutiert. Das Schlüsselmanagement für eine verschlüsselte Konferenzsitzung ist eine weitere Aufgabe der Signalisierung, die in dem Rahmen dieser Arbeit nicht betrachtet wird.

### 2.4.1. Gruppenmanagement

Die Aufgabe des Gruppenmanagements ist die Initialisierung, Verwaltung und Beendigung einer Konferenz. Dazu gehören unter Anderem folgende Teilaufgaben:

- Ankündigung der Konferenz,
- Registrieren der Teilnehmer beim Beitritt, Verlassen und Einladen,
- Überwachung der Verfügbarkeit der Teilnehmer.

Eine Konferenz kann im Allgemeinen auf zwei Arten initialisiert und vergrößert werden: durch Ankündigung oder durch Einladung. Beide Möglichkeiten sind für den Aufbau von *dynamischen* Gruppen und *statischen* Gruppen geeignet. Dynamische Gruppen können die Teilnehmerzahl während der Konferenz ändern, statische Gruppen nicht.

#### Ankündigung

Eine Videokonferenzsitzung kann über Email oder *Session Directories* angekündigt und aufgebaut werden. Das wohl populärste Session Directory ist das SDR [158] der Mbone-Tools. Die Ankündigung kann dabei ein vordefiniertes Format besitzen, um die Informationen „wo“, „wann“ und „wie“ die Konferenz stattfindet, darzustellen. Dazu wird in der Regel das *Session Description Protocol (SDP)* [54] genutzt. Diese Informationen werden per Email oder einem Ankündigungsdienst verbreitet. SDR benutzt das *Session Announcement Protocol (SAP)* [53] für die Verbreitung der SDP-Daten. Aufgrund einer empfangenen Ankündigung kann der Nutzer entscheiden, ob er der Konferenz mit den im SDP-Paket beschriebenen Parametern beitreten möchte oder nicht. Eine Ankündigung zum Aufbau einer Konferenz ist sehr effektiv, um große Gruppen in einer geringen Zeit aufzubauen. Der Nachteil ist, dass die Gruppenzusammensetzung nicht gesteuert werden kann. Jeder Nutzer, der eine Einladung auf irgendeinem Weg bekommen hat, kann der Konferenz beitreten.

#### Einladung

Eine singuläre Einladung, die explizit bestätigt werden muss, ist eine zweite, jedoch ineffizientere Möglichkeit, eine Konferenz aufzubauen. Dazu schickt der Initiator einer Konferenz oder ein bereits involviertes Mitglied einer Sitzung eine Einladung an einen gewünschten Teilnehmer. Mit der Annahme der Einladung wird der neue Teilnehmer in die Konferenz involviert. Auf diese Weise

kann dann dynamisch eine Konferenzgruppe aufgebaut werden. Explizite Einladungen für Videokonferenzen sind nur für kleine Gruppen empfehlenswert, da die Zeit für den Aufbau der Gruppe zu groß wird. Der Vorteil liegt darin, dass der Nutzerkreis beschränkt und bekannt ist. Das *Session Initiation Protocol (SIP)* [55] ist der bekannteste Vertreter eines Einladungsprotokolls im Internet.

Während der Konferenz überwacht das Gruppenmanagement die Verfügbarkeit der involvierten Teilnehmer. Es muss ständig geprüft werden, ob jeder Teilnehmer noch erreichbar ist und aus Sicht der Konferenz fehlerfrei arbeitet. Bei einem Fehlverhalten muss das Gruppenmanagement entsprechend reagieren und versuchen, den Fehler zu beseitigen oder die Gruppe wieder in einen stabilen Zustand zu führen.

Verfügbarkeit

Das Verlassen und Beenden einer Konferenz gehört ebenfalls zu den Aufgaben des Gruppenmanagements. Hierbei gibt es zwei grundlegende Ansätze. Im ersten Fall verlässt ein Teilnehmer nach dem anderen die Konferenz. Dabei muss das Gruppenmanagement sicherstellen, dass beim Verlassen des Teilnehmers nicht die gesamte Konferenz beendet wird. Dies darf erst beim Verlassen des vorletzten Teilnehmers passieren. Im zweiten Fall veranlasst ein ausgewählter Teilnehmer der Konferenz, beispielsweise ein Moderator, dass die Sitzung geschlossen wird. Hier werden alle Verbindungen zeitgleich abgebaut.

Verlassen und  
Beenden

### 2.4.2. QoS-Management

Das QoS-Management ist für eine optimale Dienstgüte während der Gruppenkommunikation verantwortlich. Der Begriff der Dienstgüte kann dabei aus verschiedenen Gesichtspunkten interpretiert werden [62, 67, 72, 136, 181]. Eine häufig verwendete Klassifizierung in netzwerkabhängige QoS, kommunikationsspezifische QoS und anwendungsorientierte QoS wird in [148] gegeben.

Die netzwerkabhängige QoS bezieht sich auf die unteren beiden Schichten im TCP/IP-Protokollstapel, inklusive dem Medienzugriff. Übertragungsrate, Verlustrate, Jitter und Übertragungsverzögerung bilden die Grundlage zur Bewertung der Qualität einer Datenübertragung.

netzwerkabhän-  
gige QoS

Die kommunikationsspezifische QoS beschreibt die Qualität, die durch die Vermittlungs- und Transportschicht erbracht wird. Hier spielen Durchsatz, Verlustrate, Jitter und Übertragungsverzögerung ebenfalls eine große Rolle.

kommunika-  
tionsspezifische  
QoS

Zusätzlich zählen Priorität und Reihenfolge eines Datenstroms zur kommunikationsspezifischen QoS-Beschreibung.

anwendungs-  
orientierte  
QoS

Die anwendungsorientierten QoS definieren die QoS-Anforderungen aus der Sicht der Anwendungen. Es werden bei der Bewertung der Qualität die Stoßartigkeit (Burstiness), die maximale Größe der Dateneinheiten, der Durchsatz, die Bitfehlerrate, der Jitter, die Priorität, die Ende-zu-Ende-Verzögerung und die Inter-/Intrastream-Synchronisation zugrunde gelegt.

nutzerspezifische  
QoS

In [147] und [10] werden zusätzlich nutzerspezifische QoS definiert, die die Wahrnehmung der Dienstqualität durch den Nutzer beschreibt. Entsprechende QoS-Parameter sind beispielsweise die Verzögerung beim Verbindungsaufbau, die Qualität der Repräsentation, die Bedienbarkeit des Systems oder die Kosten.

Für die Videoübertragung in einem Videokonferenzsystem sind die Dienstmerkmale aus Nutzersicht die Bildwiederholungsrate (Framerate), die Farbtiefe und die Bildgröße (Pixelauflösung). Die Abtastrate und die Abtasttiefe sind die entscheidenden Qualitätsmerkmale eines Audiostroms. Um diese Dienstmerkmale bestmöglich über das Netzwerk zu übertragen, müssen die anderen drei QoS-Klassen von der Videokonferenzanwendung verwaltet und entsprechend konfiguriert werden. In den meisten existierenden Videokonferenzsystemen [78, 131, 175] wird kaum Wert auf die Einstellung der QoS-Merkmale gelegt. Stattdessen muss der Teilnehmer die nutzerspezifische QoS einstellen und darauf vertrauen, dass alle Teilnehmer die gewählte Qualität empfangen können. USMINT [47] enthält einen Ansatz zur Reservierung von Netzwerkressourcen, kann aber nicht garantieren, dass dafür Mechanismen im Netzwerk bereit stehen. Ob die Qualität vom Endsystem verarbeitet werden kann, wird von USMINT ebenfalls nicht betrachtet. Die Videokonferenzsysteme von VCON [166] bieten eine automatische und dynamische Anpassung der Sendeleistung an die zur Verfügung stehende Bandbreite des Transportwegs. Eine automatische und dynamische Anpassung der Dienstgüte an die verfügbare Leistung des Endsystems enthält einzig COVIS [12].

Ein Videokonferenzsystem, das in sehr heterogenen Umgebungen arbeitet, sollte jedoch die QoS sowohl an die Endsystemleistung als auch an die Netzwerksituation anpassen können. Dies ist bisher noch in keinem Videokonferenzsystem verwirklicht worden.



### 2.4.3. Floorkontrolle

Der Zugriff der Teilnehmer auf gemeinsam genutzte Ressourcen, zum Beispiel Audiokanal oder Whiteboard, sollte kontrolliert erfolgen. Das Management konkurrierender Zugriffe auf gemeinsame Ressourcen und Daten ist als Floorkontrolle bekannt. Die meisten Floorkontrollmechanismen versuchen, menschliche Verhaltensweisen und gesellschaftliche Protokolle nachzuahmen, um somit vertrautes Verhalten auf Konferenzen und Treffen zu reproduzieren. Andere Aspekte der Floorkontrolle sind Fairness oder Stabilität. In [23] und [40] werden die verschiedenen Mechanismen und Strategien der Floorvergabe zusammengefasst:

**Agenda:** Die Floorvergabe wird vor der Sitzung statisch festgelegt und ist für dynamische Diskussionen nicht geeignet.

**Zufall:** Der neue Floor-Inhaber wird nach einem Zufallsmechanismus ausgewählt. Dies ist aber in Videokonferenzsystemen kaum sinnvoll einsetzbar, da soziale Protokolle unberücksichtigt bleiben.

**Zeitbasiert:** Der Floor wird automatisch für eine bestimmte Zeit einem Teilnehmer gegeben. Teilnehmer, die den Floor nicht benötigen, bekommen ihn trotzdem. Dadurch kann wertvolle Zeit der Diskussion verschenkt werden.

**Definierte Ordnung:** Der Floor wird in einer vorgegebenen Reihenfolge weitergeleitet. Hier entstehen die gleichen Probleme wie bei der zeitbasierten Floorvergabe.

**Abstimmung:** Der neue Inhaber des Floors wird durch eine Abstimmung aller Teilnehmer gewählt. Besonders in großen Gruppen stellt dies einen erheblichen zeitlichen Mehraufwand dar.

**Warteschlange:** Eine Anfrage an den Floor wird in eine Warteschlange abgelegt und nach Zeit- oder Prioritätskriterien bedient. Beispiele sind *First-In-First-Out (FIFO)* oder *Earliest—Deadline—First (EDF)*. Das Warteschlangenprinzip stellt eine der fairsten Strategien dar und eignet sich für Beratungen.

**Moderator:** Ein bestimmter Teilnehmer, meist der Initiator der Konferenz, bestimmt über die Vergabe des Floors. Dieser entscheidet, wer wie lange sprechen darf. Dieses Schema ist gut geeignet für Diskussionen und Beratungen.

Praxistauglich sind nur die Floorkontrollstrategien, die nach dem Warteschlangen- und Moderationsprinzip arbeiten. Sie sind fair, unterstützen soziale Protokolle und können dynamisch an den inhaltlichen Diskussionsverlauf angepasst werden.

Obwohl die Floorkontrolle bei modernen Videokonferenzsystemen notwendig ist, ist die Anzahl der bekannten Systeme mit einer Floorkontrolle erstaunlich gering. Lediglich die Videokonferenzsysteme USMINT [47, 142] und COVIS [12, 13] haben bisher eine Floorkontrolle integriert. Die populären Mbone-Tools [88, 112, 113, 157, 158] und RENDEZVOUS [131] besitzen keine. In Microsoft Netmeeting [101, 102] und anderen H.323 Systeme, die mit dem T.120 Standard [75] konform sind, ist es möglich, den Zugriff auf gemeinsam genutzte lokale Anwendungen zu sperren. Eine explizite Floorkontrolle stellt dies jedoch nicht dar.

## 2.5. Medienverarbeitung

Der eigentliche Zweck eines Videokonferenzsystems liegt in der Kommunikation der Konferenzteilnehmer. Dazu müssen Video- und Audioinformationen übertragen werden.

Die Module der Medienverarbeitung auf der Senderseite haben die Aufgabe, die kontinuierlichen, analogen Audio- und Videosignale in kontinuierliche, digitale Signale umzuwandeln, zu komprimieren und dem Transportdienst zu übergeben. Die Empfangsmodule übernehmen die Audio- und Videodaten von der Transportschicht, dekomprimieren diese und wandeln sie wieder in analoge Signale um. Diskrete, digitale Signale, die beispielsweise ein Whiteboard oder ein Chatmodul erzeugen, können vor dem Transport ebenfalls komprimiert werden. Der konzeptionelle Aufbau dieser Komponenten ist in Abbildung 2.7 gezeigt.

Die Filterung dient zur Beschränkung des Wertebereiches eines analogen Eingangssignals und zur Entfernung von Rauschen im Ausgangssignal. Der

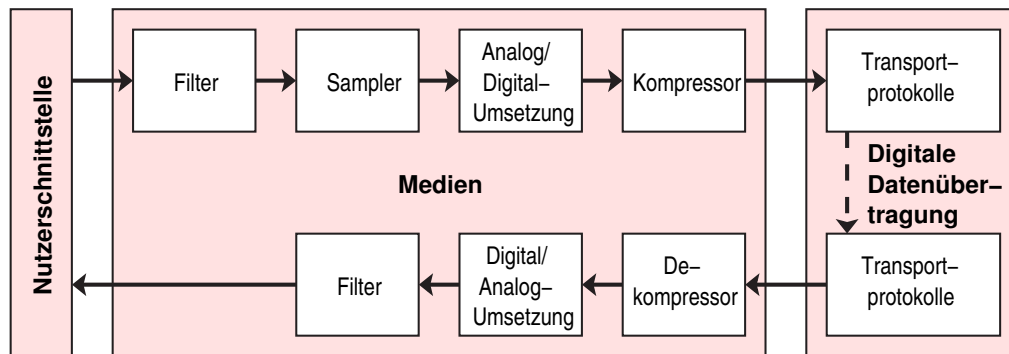


Abbildung 2.7.: Allgemeiner Ablauf der Medienverarbeitung

Sampler tastet zu festgelegten Zeitpunkten das Eingangssignal ab. Der abgetastete Wert wird anschließend digitalisiert. Die Digitalisierung liefert eine sehr hohe Datenrate, die von einem normalen Netzwerk nicht verkraftet wird. Beispielsweise liefert ein Video, das mit einer Bildgröße von CIF (Common Interchange Format, 352x288 Pixel), einer Farbtiefe von 8 bit und 25 Bildern pro Sekunde abgetastet wird, eine Datenrate von etwa 21 Mbit/s. Deshalb wird dieser digitale Strom vor der Übertragung durch das Netzwerk komprimiert. Über die Möglichkeiten der Video- und Audiokompression informieren die folgenden Abschnitte.

### 2.5.1. Videokompression

#### Grundlagen der Videokompression

Videosignale verursachen ein sehr hohes Datenaufkommen, welches von der Bildgröße, der Bildwiederholungsrate, der Farbtiefe aber auch vom Inhalt des Bildes selbst abhängt. Um diese Datenflut auf Computern mit begrenzten Speicherressourcen oder über Netzwerke mit begrenzten Bandbreitenressourcen zu übertragen, müssen Videodaten komprimiert werden.

Grundlage aller Videokompressionsverfahren ist die Kompression eines Einzelbildes. Hierbei kommen Ansätze wie die Entropiekodierung (zum Beispiel Huffman-Kodierung), die Querkodierung (Diskrete Cosinus Transformation, DCT) oder die Hybride Kodierung (zum Beispiel JPEG) zum Einsatz [146]. Videodaten haben die Eigenschaft, dass aufeinander folgende Bilder sehr ähnlich zueinander sind. Deshalb wird oft eine Bewegungserkennung eingesetzt,

um nur die Differenzen zwischen den Bildern übertragen zu müssen. Diese Differenzbilder werden auch *Interframes (P-Frames)* genannt. In verlustbehafteten Umgebung kann jedoch das Problem auftreten, dass ein Interframe verloren geht und somit der nachfolgende Strom nicht mehr dekomprimiert werden kann. Um solche Probleme zu umgehen, werden bei vielen Kompressionsmethoden so genannte *Intraframes (I-Frames)* in den Strom eingefügt, die wieder die komplette Information eines Bildes enthalten. Diese werden also als Einzelbild komprimiert. Somit ergibt sich also ein komprimierter Videostrom aus einem Wechsel von I-Frames und mehreren P-Frames.

### M-JPEG

Durch die *Joint Photographic Expert Group* [82] wurde das populäre Kompressionsformat JPEG für Standbilder eingeführt. JPEG unterstützt sowohl Graustufen- als auch Farbbilder. Aufgrund verschiedener Anwendungsfelder sind verlustlose und verlustbehaftete Kompressionsmodi definiert worden. Für Entertainment-Anwendungen ist eine verlustbehaftete Kompression ausreichend. Hingegen sind in einem medizinischen Umfeld Verluste unerwünscht.

M-JPEG steht für Motion-JPEG und ist eigentlich kein Videokompressionsverfahren. Bei M-JPEG folgen einzelne JPEG-Bilder zeitlich hintereinander, wodurch der Eindruck eines bewegten Bildes entsteht. M-JPEG ist ein einfacher Codec, der keine Bewegungserkennung und keine zeitliche Kompression durchführt, also nur aus I-Frames besteht. Aus diesem Grund ist die Ausgangsdatenrate eines M-JPEG-Kodierers recht hoch.

Für eine Videokonferenzanwendung eignet sich M-JPEG aufgrund der hohen Datenrate in Umgebungen mit hohen Bandbreitenressourcen. M-JPEG ist sehr robust gegenüber Übertragungsverlusten, da keine Abhängigkeiten zwischen den Bildern bestehen. Somit ist der Verlust eines Bildes für den Nutzer kaum spürbar. Dieser Vorteil kann zur dynamischen Skalierung von M-JPEG-Strömen genutzt werden, um die Datenrate an das Aufnahmevermögen des Teilnehmersystems anzupassen. So lässt sich zum Beispiel durch einen einfachen *Frame-Dropping-Filter* die Datenrate des Videostroms reduzieren, ohne eine aufwendige Dekodierung mit anschließender Neu-Kodierung durchzuführen.

## Entwicklungen der ISO

Die *Motion Picture Expert Group (MPEG)* [105] der *International Standardization Organisation (ISO)* hat das Ziel, Video- und Audiosequenzen mit möglichst geringer Speicherkapazität trotz hoher Qualität auf elektronischen Medien, wie einer CD oder DVD, zu speichern. Durch diese Gruppe wurden bis heute fünf Standards definiert: MPEG-1, MPEG-2, MPEG-4, MPEG-7 und MPEG-21, die jeweils in drei Ebenen unterteilt werden. Die erste Ebene enthält Systeminformationen, die zweite und dritte Ebene beschreiben die Video- beziehungsweise Audiokodierung.

Der *MPEG-1* [106] Videostandard unterstützt die Interframekodierung mittels Bewegungskompensation und bidirektionaler Interpolation. Die unabhängigen Intra-Frames sind JPEG kodiert. MPEG Kodierer sind in der Lage, eine konstante Bitrate bis zu 1,5 Mbit/s zu unterstützen.

MPEG-1

Dem *MPEG-2* [107] wurden gegenüber dem MPEG-1 neue Charakteristiken hinzugefügt. Herauszustellen ist ein zweistufiges Skalierungsschema, das es erlaubt, leistungsschwachen Endsystemen eine geringere Qualität zur Verfügung zu stellen (siehe auch *Layered Coding*). Weiterhin werden höhere Bandbreite von 4 bis 20 Mbit/s unterstützt.

MPEG-2

*MPEG-4* [110] wurde für die Echtzeitkommunikation zwischen zwei Partnern sowie für die Überwachung und das Infotainment, beispielsweise für DVD, Videobroadcast-Anwendungen (*Digital Television (DTV)*) und Video Streaming entwickelt. Der Standard spezifiziert *Audio Visual Objects (AVOs)* mit auditiven, visuellen oder audiovisuellen Inhalten, die im gemeinsamen Zusammenwirken eine audiovisuelle Szene beschreiben können. Der Vorteil von MPEG-4 besteht in der Möglichkeit, beliebige Konturen zu komprimieren. Es werden Bitraten von 5 kbit/s bis 10 Mbit/s unterstützt.

MPEG-4

Der *MPEG-7* Standard [109] wird auch *Multimedia Content Description Interface* genannt. Dieser beschreibt die Katalogisierung von AVOs, um anhand von Stichwörtern ein gewünschte audiovisuelle Szenen zu suchen. Der Einsatz in Videokonferenzsystemen ist unwahrscheinlich.

MPEG-7

Der *MPEG-21* [108] beschreibt ein Multimedia-Framework, um einen einheitlichen Zugriff auf Text, Grafik, Audio und Video zu ermöglichen. Für Videokonferenzsystemen ist MPEG-21 ebenfalls kaum relevant.

MPEG-21

Die Kodierer der MPEG stellen hohe Anforderungen an das Endsystem. Selbst Computer mittlerer Leistung sind oft nicht in der Lage, einen kontinuierlichen Datenstrom mit 25 fps und einer CIF-Auflösung zu komprimieren. Die Folge sind starke Verzögerungen sowie Reduzierung der Sendequalität. Um trotzdem MPEG-Ströme auf leistungsschwachen Systemen zu erzeugen, werden in der Regel preisintensive<sup>1</sup> Hardware-Kompressionskarten eingesetzt. Ein weiteres Problem ergibt sich, wenn ein MPEG-Strom an mehrere Empfänger über heterogene Netzwerke übertragen werden soll. Ein Empfänger mit einer schlechten Netzwerkanbindung erhält aufgrund von Paketverlusten nur noch Fragmente der ursprünglichen Daten, die zumeist keine zufrieden stellende Qualität liefern. Wird andererseits die Sendeleistung an den Empfänger mit der schlechten Netzwerkverbindung angepasst, kann sich die Qualität des MPEG-Stroms auch für Empfänger mit einer guten Netzwerkanbindung drastisch reduzieren. Aus diesem Grund lässt sich die MPEG-Komprimierung nur bedingt in heterogenen Netzwerkumgebungen, wie dem Internet, einsetzen.

### Entwicklungen der ITU-T

Parallel zu den Bemühungen der ISO entwickelte die ITU-T ebenfalls eine Reihe von Kompressionsverfahren und -standards speziell für Videokonferenzanwendungen.

H.261

Der *H.261*-Standard [68] wurde ursprünglich für ISDN basierte Videokonferenzsysteme mit einer Bandbreite von einem Vielfachen eines  $S_0$ -Anschlusses von 64 kbit/s entworfen, wodurch dieser auch als  $p \times 64$  Standard bezeichnet wird. Der Standard bietet konstante Bitraten an, indem die Quantisierungstiefe, das Zeitintervall zwischen zwei Intraframes und die Framerate variiert werden.

H.263

Der *H.263*-Standard [73] unterstützt sowohl hohe als auch geringere Bitraten, um auch mit einem einfachen  $S_0$ -Anschluss oder mit einem analogen Telefonanschluss mit einer Bitrate von maximal 56 kbit/s an einer Videokonferenz teilnehmen zu können. H.263 definiert dazu zusätzliche Bildformate, die von *Small QCIF (SQCIF)* mit  $128 \times 96$  Pixeln bis *16CIF* mit  $1408 \times 1152$  Pixeln reichen. Zur Aufzeichnung von Videosequenzen können in H.263 *bidirektiona-*

---

<sup>1</sup>Zur Zeit bewegen sich die Preise zwischen 200 bis 1500 €.

le Frames (*B-Frames*), die vom vorhergehenden und vom nachfolgenden Bild abhängig sind, optional genutzt werden.

H.26x Kodierer finden aufgrund der geringen Bitraten des komprimierten Videostroms häufigen Einsatz in Videokonferenzsystemen. Die Verwendung einer Bewegungserkennung benötigt jedoch eine hohe Rechenkapazität, so dass auch hier auf hardwarenahe Lösungen zurückgegriffen werden muss. H.26x ist wie MPEG nicht für heterogene Netzwerkkumgebungen geeignet.

### Layered Coding

*Layered Coding* Verfahren sind geeignet, heterogene Netzwerke zu unterstützen. Die Grundidee ist, dass durch einen Kodierer mehrere Qualitätsebenen des Bewegtbildes erzeugt werden, die sukzessiv aufeinander aufbauen. Dabei gibt es die Möglichkeiten der *räumlichen*, der *temporalen* und der *hybriden* Kompression.

Bei der räumlichen Kodierung wird ein Einzelbild eines Videostroms in mehrere Bilder mit verschiedenen Auflösungen, Frequenzanteilen oder Quantisierungsstufen zerlegt [6]. Beispielsweise enthält die unterste Ebene (Layer 4) nur ein Bild mit einer Auflösung von 98x72 Pixeln, wobei bis zur obersten Ebene (Layer 1) schrittweise Informationen hinzugefügt werden, um eine Auflösung von 784x576 Pixeln zu erreichen (siehe Abbildung 2.8). In ähnlicher Weise werden zusätzliche Frequenzanteile und Quantisierungsstufen in höheren Ebenen kodiert. Ein Beispiel stellt der modifizierte H.261 Codec LH.261 (*Layered H.261*) [87] dar. Dieses Art der Kompression erfordert wie MPEG und H.26x selbst eine hohe Leistung der kodierenden Endsysteme. In [99] wurde von einer maximalen Bildrate von 11 fps berichtet. Eine Hardwarelösung existiert in Form eines MPEG-2-Kodierers für HDTV-Systeme. MPEG-2 unterstützt jedoch eine Kompression von nur zwei Ebenen mit unterschiedlicher Auflösung oder unterschiedlicher Darstellungsqualität [140, 154].

räumliche Kompression

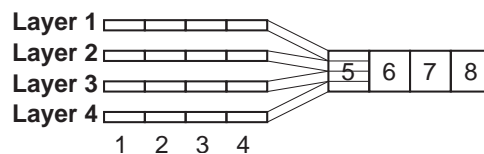


Abbildung 2.8.: Funktionsweise der räumlichen Skalierung

## 2. Gestaltungsprinzipien von Videokonferenzsystemen

temporale Kom-  
pression

Bei der temporalen Kompression werden vollständige, aufeinander folgende Bilder des Videostroms auf mehrere Ebenen verteilt, so dass das flüssige Abspielen eines Videostroms von der Anzahl der empfangenen Ebenen abhängt. So kann zum Beispiel ein Teilnehmer beim Empfang der untersten Ebene eine Framerate von 1 fps abspielen. Durch den Empfang von höheren Ebenen kann die Framerate bis zur maximale Bildwiederholungsrate des Senders erhöht werden. Kompressionsverfahren ohne Bewegungserkennung wie M-JPEG eignen sich sehr gut für die temporale Kompression, da jedes komprimierte Bild auf eine separate Ebene gelegt werden kann und somit eine feine Abstufung der Qualitätsebenen möglich ist (siehe Abbildung 2.9). Bei Kompressionsverfahren mit Bewegungserkennung lassen sich jedoch nur wenige Ebenen definieren, wodurch keine feine Abstufung möglich ist. Bei H.261 lassen sich beispielsweise zwei Ebenen und bei MPEG drei Ebenen unterteilen (siehe Abbildung 2.10 und 2.11). In MPEG-2 ist weiterhin eine Komponente zur temporalen Kompression von Stereo-Bildsequenzen standardisiert, die für Videokonferenzen jedoch nicht relevant ist.

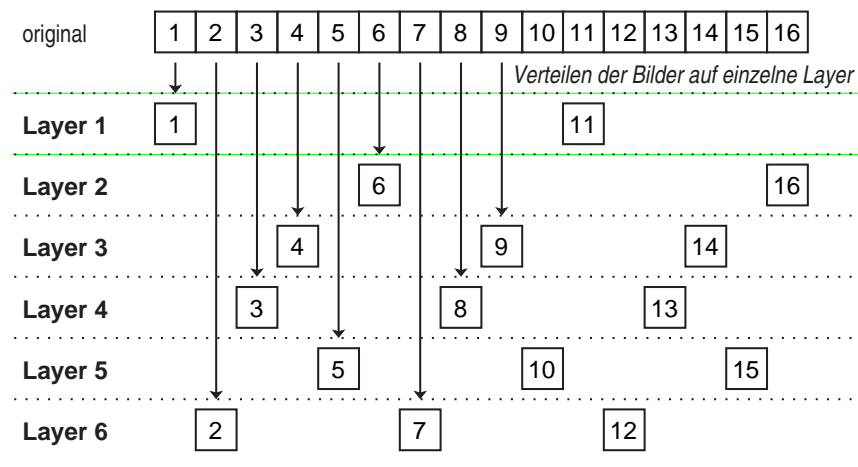


Abbildung 2.9.: Möglichkeiten der temporalen Skalierung mit M-JPEG

hybride Techni-  
ken

Neben den einzelnen Verfahren existieren hybride Techniken, die die temporale und räumliche Kompression hintereinander durchführen [87]. Die Reihenfolge und die Wichtung der Einzelkompression hängt dabei in einem starken Maße von der Anwendung ab.



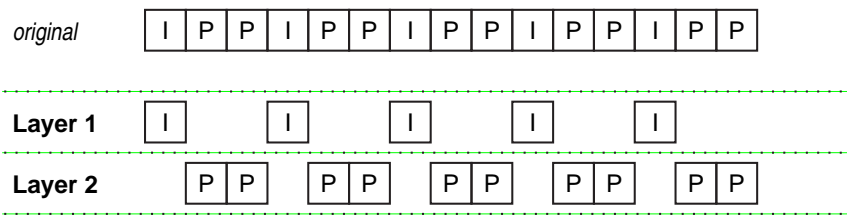


Abbildung 2.10.: Möglichkeiten der temporalen Skalierung mit H.261 (Ebene 1: I-Frames, Ebene 2: P-Frames)

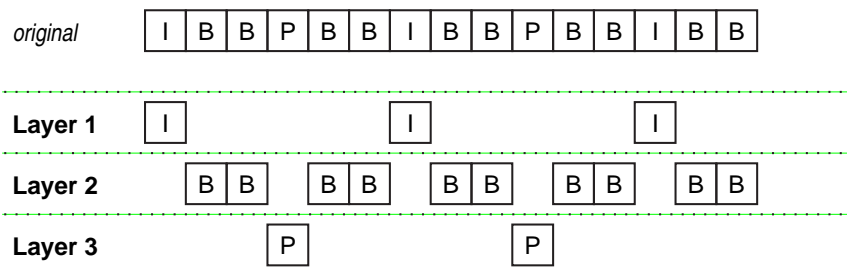


Abbildung 2.11.: Möglichkeiten der temporalen Skalierung mit MPEG (Ebene 1: I-Frames, Ebene 2: B-Frames, Ebene 3: P-Frames)

### Weitere Kompressionsmöglichkeiten

Die *Wavelet*-Kompression bildet eine andere Möglichkeit der Bildkomprimierung. Grundidee ist eine Transformation des originalen Bildes in einen Frequenzbereich unter Nutzung von linearen Funktionen innerhalb von bestimmten Intervallen. Zur Zeit wird die auch sehr aufwendige Waveletkompression nur für Standbilder eingesetzt, beispielsweise in JPEG2000 [81]. Hierbei ist man auch bestrebt, heterogene Systemumgebungen durch Layered-Coding-Techniken zu unterstützen [87].

Wavelet-  
Kompression

*Fractal Coding* basiert auf die Verwendung von affinen Gebieten, iterativen Systemfunktionen und dem Collage-Theorem. Der Kompressionsaufwand ist ebenfalls sehr hoch, so dass eine Echtzeitkompression für bewegte Bilder in Videokonferenzsystemen kaum möglich ist.

Fractal Coding

### 2.5.2. Audioverarbeitung

Die Sprache ist in einem Videokonferenzsystem das wichtigste Medium überhaupt, da in ihr die meisten Informationen enthalten sind. Somit muss eine ho-

he Qualität der Sprachübertragung auch in Situationen gewährleistet werden, in denen die Netzwerkkapazitäten knapp werden. Im Gegensatz dazu kann auf die hochwertige Übertragung der Videostreams bei schlechten Übertragungskapazitäten notfalls verzichtet werden, da das Bild meist nur das typische Kopf-Schulter-Portrait der Teilnehmer enthält [114].

Für die Audiokompression wurde eine Vielzahl von Ansätzen entwickelt, die unterschiedliche Datenraten und Qualitäten unterstützen, die maßgeblich von der ISO und der ITU-T geprägt wurden.

### Entwicklungen der ISO

Die ISO hat mit dem *MPEG Layer 3 (MP3)* ein populäres Kompressionsverfahren für Audiodaten, insbesondere Musikdateien, entwickelt. Weitere Einsatzgebiete sind im Bereich des Home-Entertainments und im Zusammenhang mit den MPEG-Videostandards zu finden. MP3 entfernt dabei Frequenzbereiche aus dem originalen Audiostrom, die das menschliche Gehör kaum wahrnimmt. Dabei werden Bitraten von 2.5 bis 192 kbit/s unterstützt. Eine Echtzeitkompression mit dem sehr komplexen Codec ist aber nur mit sehr leistungsstarken Endsystemen realisierbar, wodurch ein Einsatz für Videokonferenzsysteme nur bedingt möglich ist.

### Entwicklungen der ITU-T

Die ITU-T hat eine weitaus größere Anzahl von Verfahren für die Audiokompression entwickelt und standardisiert als die ISO.

Ein analoges Audiosignal, das bereits in elektrischer Form vorliegt, wird mittels *Pulse Code Modulation (PCM)* in ein binäres, digitales Signal umgewandelt. Die Parameter der PCM, die Bandbreite von ca. 3 kHz (300 – 3400 Hz), die Abtastrate von 8 kHz und die Bittiefe von 8 bit, wurden von der ITU-T im *G.711*-Standard standardisiert [63]. Dies entspricht einer normalen Telefonqualität. Daraus resultiert eine Bitrate von 64 kbit/s ( $= 8000 \frac{1}{s} \times 8 \text{ bit}$ ). Eine Kompression ist im *G.711* Standard nicht vorgesehen. Somit ist er wenig komplex, robust gegen Bitverfälschungen und hat eine gute Sprachqualität. Es wird jedoch eine relativ hohe Bandbreite benötigt.

G.711

G.721

Eine Audiokompression wird durch den *G.721* Standard [64] realisiert, wobei das durch den *G.711* Standard erzeugte digitale Abbild als Eingangssignal

für den Komprimierer genutzt wird. Als Basis wird die *Adaptive Differential Pulse Code Modulation (ADPCM)* genutzt, das die Differenz zwischen zwei aufeinander folgenden Abtastwerten in einem Wert von 4 bit Tiefe abbildet und somit eine Kompression von 50% erreicht. Die Bitrate reduziert sich dadurch auf nur 32 kHz. Die Sprachqualität unterscheidet sich kaum von der des G.711 Standards.

Eine weitaus höhere Audioqualität bietet der *G.722* Standard [65], die aus einer größeren analogen Audiobandbreite von 7 kHz resultiert. Dadurch werden insbesondere Reibelaute wie „f“ und „s“ deutlicher verstanden. G.722 unterstützt dabei verschiedene Ausgangsbitraten von 48, 56 und 64 kbit/s. Als Basis dient ein *Subband Coder* mit je einem ADPCM-Modul pro Subband. Das Prinzip ist in Abbildung 2.12 gezeigt. Das ADPCM-Modul des oberen Bandes gibt einen 2 bit Wert und das des unteren Bandes einen 4, 5 oder 6 bit Wert zurück. Durch das Multiplexen der beiden Werte resultieren daraus Bitraten von 48, 56 und 64 kbit/s.

G.722

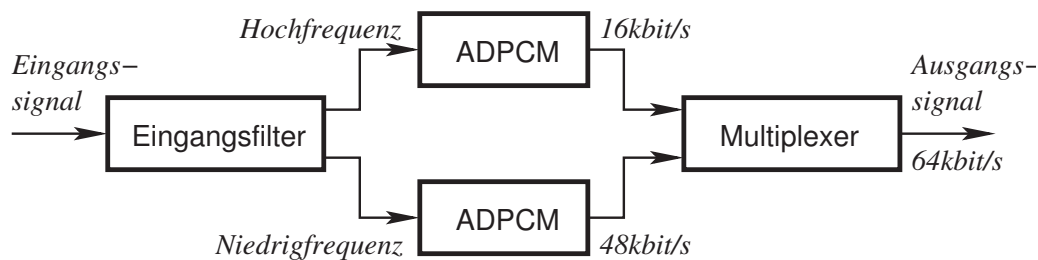


Abbildung 2.12.: Das Kodierungsprinzip des G.722 Standards

Der ITU-T Standard *G.728* definiert einen *CODEC (Coder and Decoder)*, der mit einer Bitrate von 16 kbit/s auskommt [66]. Die Komprimierung beruht auf dem *Low-Delay Code Excited Linear Prediction (LD-CELP)* Prinzip, das sich von den vorher erläuterten Methoden stark unterscheidet. G.728 nutzt die Charakteristika der menschlichen Sprache und des menschlichen Hörempfindens weitaus mehr aus, als dies G.711 und G.721 tun. Als Eingangswerte werden wiederum PCM-codierte Daten benötigt, wobei aus fünf aufeinander folgenden Werten ein so genannter *Quantisierungsvektor* gebildet und dann übertragen wird. Die Verzögerung, die durch den Kodierungs- und Dekodierungsprozess auftreten, beläuft sich auf etwa 7 bis 10 ms. Der CODEC ist sehr komplex, aber trotzdem relativ robust gegenüber zufälligen Bitverfälschungen.

G.728

G.729

Die G.729 Empfehlung der ITU-T [71] reduziert die benötigte Bandbreite auf 8 kbit/s. Die Basis für die Komprimierung bildet das *Conjugate Structure Algebraic Codebook Excitation Linear Prediction (CS-ACELP)*. Dieser CODEC ist besonders für kabellose Übertragungen mit hohen Fehlerraten geeignet. Die Verzögerung vom Sender zum Empfänger belaufen sich bei diesem recht komplexen Verfahren auf circa 35 ms.

G.723.1

Der *de facto* Standard für *Voice over IP (VoIP)* ist der G.723.1 Standard [70], der mit einer Bitrate von 6,3 oder 5,3 kbit/s auskommt. Das ACELP-Verfahren wird dabei für die niedrige Bitrate genutzt. Die höhere Bitrate wird durch den Sprachkodierer *MultiPulse-Maximum Likelihood Quantizer (MP-MLQ)* erreicht. Die starke Kompression hat allerdings ihren Preis in einer starken Verzögerung von mindestens 100 ms. Die Qualität der Sprache ist mit der eines herkömmlichen Telefons vergleichbar.

## 2.6. Netzwerkunterstützung und Transport

Einen entscheidenden Einfluss auf die Qualität eines Videokonferenzsystems hat die Netzwerkunterstützung und die Leistungsfähigkeit des Netzwerks.

Videokonferenzen können sowohl über leitungsvermittelnde als auch über paketvermittelnde Netzwerke realisiert werden. Traditionell wurden leitungsvermittelnde Netzwerke, zum Beispiel ISDN, bevorzugt für Videokonferenzsysteme eingesetzt, weil aufgrund der physikalisch geschalteten Leitung die QoS garantiert werden kann. Die Bitrate liegt stabil bei einem Vielfachen von 64 kbit/s, und die Verzögerungen sind gering und konstant. Von Nachteil sind der hohe Verwaltungsaufwand, die fehlende Flexibilität bei der Nutzung, insbesondere bei einem Ausfall einzelner Leitungen, und die schwierige Umsetzung der Mehrteilnehmerkommunikation. Weiterhin wachsen die Betriebskosten stark an, insbesondere wenn die Teilnehmer über eine große räumliche Entfernung kommunizieren.

paketvermittelnde Netze

In paketvermittelnden Netzwerken ist die Absicherung der QoS noch problematisch, da keine direkte physikalische Verbindung zwischen den Kommunikationspartnern besteht. Stattdessen werden auf einer physikalischen Leitung mehrere logische Verbindungen verschiedener oder derselben Kommunikationspartner aufgebaut. Dies bewirkt eine bessere und flexiblere Nutzung der Lei-

tung. Aus diesem Grund geht zur Zeit der Trend zu paketvermittelnden Netzen wie dem Internet. Selbst der klassische Telefondienst wird über IP-Netzwerke abgebildet. Für die Sicherung der QoS in IP-Netzwerken existieren zur Zeit drei unterschiedlichen Ansätze: die Integrated Services, die Differentiated Services und das Multi-Protocol-Label-Switching. Parallel dazu existiert seit den späten siebziger Jahren der Ansatz *Internet Stream Protocol*.

Die *Integrated Services* reservieren Ressourcen in den Netzwerkknoten unter Nutzung des *Resource Reservation Protocol (RSVP)* [21, 180]. Hierbei werden vom Sender zum Empfänger *PATH*-Nachrichten geschickt, die die Weginformation vom Sender zum Empfänger beinhalten. Die *RESV*-Nachrichten werden auf dem *PATH* vom Empfänger zum Sender geschickt und enthalten die Reservierungsanforderungen. Die *PATH*- und *RESV*-Nachrichten werden periodisch gesendet, um die Reservierung in den Routern aufrecht zu erhalten, da diese nach der *Soft-State*-Idee Statusinformationen bei Ablauf von Timern löschen. Fällt ein Teilnehmer aus, so wird auf diese Weise seine Reservierung nach einiger Zeit automatisch gelöscht. Jeder Router verwaltet für jede reservierte Verbindung *Flow* einen *Soft State*, wodurch RSVP nicht sehr gut skalierbar in Bezug auf die gleichzeitige Nutzung durch mehrere Nutzer ist. Des Weiteren wird RSVP nicht durch öffentliche Netzbetreiber unterstützt, da entsprechende Abrechnungsmechanismen fehlen. Aus diesem Grund können Integrated Services für Videokonferenzsysteme nur in lokalen, selbst administrierten Netzwerken eingesetzt werden.

*Integrated Services*

Im Gegensatz zu RSVP werden bei den *Differentiated Services* [15] keine separaten Signalisierungen entlang der einzelne Datenwege durchgeführt, sondern es werden allgemein gültige Verhaltensregeln in den Routern definiert, die so genannten *Per Hop Behaviors (PHBs)*. Im *Type-of-Service*-Feld des IP-Headers werden *DiffServ-Codepoints* gesetzt, die zur Unterscheidung von Datenströmen mit unterschiedlichen PHBs dienen. Zur Zeit existieren zwei PHBs, das *Assured Forwarding (AF)*[56] und das *Expedited Forwarding (EF)*[80].

*Differentiated Services*

Unter *Assured Forwarding (AF)* [56] werden verschiedene Klassen (AF1 bis AF4) mit einer so genannten *Drop Precedence* zur Verfügung gestellt, wobei Pakete unterschiedlicher Klassen vom Router differenziert behandelt werden. Pakete gleicher Klassenzuteilung haben die gleiche Wahrscheinlichkeit, verworfen zu werden. Mit diesem Verfahren lassen sich zum Beispiel Audioströme als

*Assured Forwarding*

wichtige kontinuierliche Daten in Klassen mit niedriger *Drop Precedence* einordnen, um die Verlustrate gering zu halten. Eine konstante geringe Verzögerung lässt sich aber mit diesem Ansatz nicht erreichen [160]. Es können lediglich Verlustwahrscheinlichkeiten von Paketen beeinflusst werden.

*Expedited  
Forwarding*

Mit dem *Expedited Forwarding (EF)* Ansatz [80] wird ein Dienst ähnlich einer leitungsvermittelten Strecke realisiert. Ankommende EF-Pakete werden ohne Verzögerung sofort weitergeleitet. EF setzt jedoch voraus, dass die Menge des ankommenden EF-Verkehrs kleiner oder gleich den Ressourcen auf den weiterführenden Verbindungen ist. Dadurch wird gewährleistet, dass sich maximal *ein* Paket pro eingehender Verbindung in der EF-Warteschlange befindet. Somit ist mit diesem Ansatz eine Minimierung der Verzögerung möglich [160].

*MPLS*

Der dritte Ansatz, das *Multi-Protocol-Label-Switching (MPLS)*, leitete sich vom *IP over ATM* ab [132]. Das Verfahren sucht für jedes Datenpaket den kürzesten Weg durch ein Netz, wodurch das zeitaufwendige Weiterleiten der Datenpakete im Hop-by-Hop-Routing umgangen wird. Jedes MPLS-Paket kann dabei auch über ein Label mit Informationen über dessen Priorität verfügen. Mit MPLS lässt sich damit jeder Anwendung eine differenzierte Dienstgüte zuordnen. MPLS ist jedoch sehr komplex und erfordert einen hohen Administrationsaufwand [133].

Zur Zeit lassen sich aber MPLS, Differentiated Services sowie die Integrated Services nicht in öffentlichen, globalen Umgebungen nutzen, sondern werden eher in Intranets eingesetzt, in denen Echtzeitverkehr wie VoIP übertragen wird. Projekte wie QUASAR [133, 126], GÉANT/TFNGN [155], SEQUIN [32] oder INTERNET2 [59] können nicht auf öffentliche Netze übertragen werden, da zum einen noch keine Verfahren entwickelt worden sind, um die höherwertigen Dienste für die Echtzeitübertragung abzurechnen. Zum anderen müsste jeder Router im öffentlichen Internet um funktionelle Einheiten erweitert werden, die diese Dienste erbringen. Dies ist ein ähnliches Problem, wie die Migration des IP von Version 4 auf Version 6. Somit ist in naher Zukunft kaum mit einer weltweiten Einführung von DiffServ oder MPLS zu rechnen.

*Internet Stream  
Protocol*

Das *Internet Stream Protocol (ST)* [44], das inzwischen weiterentwickelt worden ist und als ST-2 und ST-2+ vorliegt [161, 36], beschreibt eine Kommunikationsarchitektur für den Transport von multimedialen Echtzeitdaten wie Audio und Video. Mittels ST lassen sich Netzwerkverbindungen mit ei-

ner garantierten Dienstgüte aufbauen. Es wird das Addressierungsschema des IP-Protokolls benutzt, wodurch ST neben IP existieren kann. Weiterhin wird vom ST-Sender für jeden gesendeten Datenstrom eine *Session ID (SID)* generiert, die im gesamten Netzwerk eindeutig ist. Jeder Netzwerkknoten, über den dieser Datenstrom übertragen wird, muss diese SID und die damit verknüpften Dienstgüteanforderungen solange verwalten, bis die Verbindung abgebaut wird. Dieser Ansatz ist mit dem von ATM vergleichbar. ST konnte sich bisher jedoch nicht durchsetzen, da der Ansatz sehr komplex ist, und sämtliche Internet-Router und -Endsysteme um ST-Funktionalitäten erweitert werden müssten. Dies ist mit denselben Schwierigkeiten verbunden, wie die Einführung von IPv6 oder DiffServ. Deshalb ist ST immer noch in einem experimentellen Zustand.

Das Problem der fehlenden QoS-Garantien in öffentlichen paketvermittelnden IP-Netzen bleibt also immer noch zu lösen. Deshalb müssen zur Zeit andere Ansätze genutzt werden, um eine Echtzeitübertragung in globalen paketvermittelnden Netzen mit einer maximal möglichen Dienstgüte zu realisieren. In Kapitel 5 wird dazu ein Vorschlag unterbreitet.

Ein Vorteil der Paketvermittlung gegenüber der Leitungsvermittlung liegt darin, dass eine Mehrteilnehmer-Kommunikation recht leicht über Multicastdienste realisiert werden kann, bei denen Netzwerkknoten die Erzeugung und Weiterleitung von Duplikaten übernehmen. Zur Zeit ist der empfängerbasierte IP-Multicastdienst standardisiert. Empfängerbasiert bedeutet, dass jede Station, die ein Empfänger einer Multicastgruppe sein möchte, dieser ohne Einschränkungen beitreten kann. Dadurch können sehr große Gruppen aufgebaut werden. Das weltweit größte Multicast Backbone (Mbone) wurde Anfang der neunziger Jahre durch universitäre Einrichtungen ins Leben gerufen. Hierbei verbinden sich einzelne IP-Multicastinseln über IP-Tunnel und bilden so ein gemeinsames Multicast-Backbone-Netzwerk.

IP-Multicast

Der entscheidende Nachteil des empfängerbasierten Multicast-Ansatzes ist, dass die Gruppenzugehörigkeit nicht beschränkt werden kann und somit ein Sender einer Multicastgruppe nicht genau weiß, wer die Daten empfängt. Des Weiteren ist IP-Multicast aufgrund seines beschränkten Adressraums nicht beliebig in Bezug auf die Anzahl der gleichzeitig stattfindenden Konferenzen skalierbar. IP-Multicast konnte sich außerdem nicht weltweit durchsetzen, da

die Kosten für die Routertechnik recht hoch sind und die Bandbreite für den Multicastdienst in den Routern auf typischerweise 1024 kbit/s beschränkt wird. Dadurch wird die Anzahl der gleichzeitig über einen Router betriebenen Videokonferenzen ebenfalls beschränkt. Des Weiteren haben private Anwender kaum eine Möglichkeit zur Nutzung von IP-Multicast, da die *Internet Service Provider (ISP)* oft keinen Multicastdienst zur Verfügung stellen oder sogar durch ihre Allgemeinen Geschäftsbedingungen (AGB) ausschließen. Somit bleibt nur die Möglichkeit, selbst einen IP-Multicasttunnel durch das ISP-Netzwerk aufzubauen, wobei sich dann die Frage stellt, wer einen Endpunkt des Tunnels zur Verfügung stellt.

ST

Im Konzept von ST wurde bereits ein senderbasierter Multicastdienst integriert. Senderbasiert bedeutet, dass der Sender festlegt, wer die Empfänger seiner Daten sind. Ein ST-Sender baut zu jedem Empfänger eine explizite Verbindung mit derselben *Session ID (SID)* auf. Die Netzwerkknoten entscheiden beim Verbindungsaufbau aufgrund der Routinginformationen, ob sie den Verbindungspfad aufspalten müssen. ST konnte sich jedoch, wie schon erwähnt, bisher nicht durchsetzen und spielt deshalb keine Rolle mehr für ein senderbasiertes Multicast.

XCast

In der Standardisierungsphase befindet sich zur Zeit ein senderbasiertes Multicastprotokoll XCast [16] für IP. XCast definiert einen eigenen Paketkopf (siehe Abbildung 2.13), der zwischen IP und UDP platziert wurde und die Zieladressen aller Empfänger (*Zieladresse\_1, ..., Zieladresse\_n*) des Datenpakets enthält. Die Empfängerliste eines XCast-Datenpakets wird nur von XCast-fähigen Routern analysiert und gegebenenfalls das Datenpaket dupliziert. Der letzte XCast-Router vor dem Ziel des Pakets entfernt den XCast-Paketkopf und schickt ein normales UDP-Paket an den oder die Empfänger.



Abbildung 2.13.: Vereinfachter Aufbau eines XCast-Paketes

Die XCast-Router gehören einer definierten IP-Multicastgruppe an, über die die XCast-Datenpakete verteilt werden. Die IP-Zieladresse eines XCast-



Pakets muss deshalb eine für die XCast-Router festgelegte Multicastadresse enthalten. Erst der letzte XCast-Router ersetzt die Zieladresse des IP-Kopfes mit der IP-Adresse des Empfängers.

XCast ist zum gegenwärtigen Zeitpunkt noch nicht ausgereift. Es sind noch eine Reihe konzeptioneller Probleme zu lösen, unter anderem das Entstehen von Schleifen in Ethernet-Umgebungen [90]. Durch die Nutzung des empfängerbasierten IP-Multicastdienstes für die Kommunikation zwischen den XCast-Routern ist trotz des senderbasierten XCast-Ansatzes keine Aussage über die genaue Anzahl der aktuellen Teilnehmer möglich, da einer IP-Multicastgruppe unbegrenzt Teilnehmer beitreten dürfen. Außerdem gibt es bisher keine Messungen, die die Leistungsfähigkeit des XCast für Gruppen mit mehr als 10 Teilnehmern bewerten.

Aufgrund der schwierigen Einführung von Multicastdiensten im Internet wurden in den letzten zwei Jahren im Kontext der Peer-to-Peer Technologie verstärkt Ansätze untersucht, um Multicastverteilstrukturen in die Anwendungsebene zu verlagern. Diese Technik wird *Application Layer Multicast*, *End System Multicast* oder *Overlay Multicast* genannt. Overlay Multicast-Ansätze haben den Vorteil, dass die Multicastdienste unabhängig von Netzwerktechnologien, und -topologien sind. Jedoch werden Ressourcen des Netzwerks und der Endsysteme stärker belastet. Es haben sich drei wesentliche Ansätze entwickelt: der *Mesh First* Ansatz, der *Tree First* Ansatz und der *Implizite* Ansatz [8].

<i>Overlay Multicast</i>
------------------------------

Der *Mesh First* Ansatz, der dem NARADA Protokoll [28, 29] zugrunde liegt, geht von einer vorhandenen vermaschten Topologie aus (siehe Abbildung 2.14). Aus dieser vermaschten Topologie wird mittels bekannter Routingprotokolle, wie zum Beispiel DVMRP [35], ein Baum pro Teilnehmer für die Datenverteilung aufgebaut. Es werden periodisch Nachrichten über den Zustand des logischen Netzes ausgetauscht, um dynamisch die optimalen Pfade zu bestimmen sowie Fehler von Endsystemen zu erkennen. Dies erfordert jedoch insbesondere für große Gruppen einen hohen Mehraufwand.

Das *Your Own Internet Distribution (Yoid)* [45] und *Host Multicast (HMTP)* [179] basieren auf dem *Tree First* Ansatz. Hierbei wird ein Baum für die Verteilung der Daten sukzessiv aufgebaut, der von allen Teilnehmern genutzt wird. Dies führt unter Umständen dazu, dass die Wurzel stark belastet wird und hohe

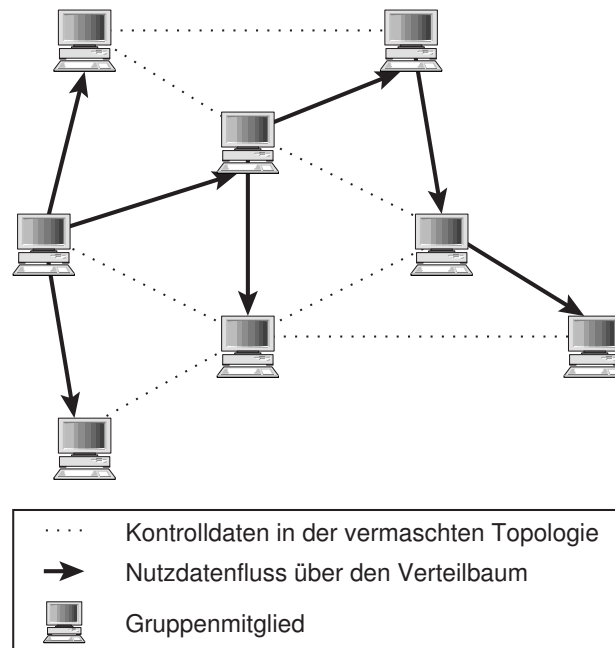


Abbildung 2.14.: Mesh First Ansatz für Overlay Multicast

Verzögerungen auftreten. Der Baum wird wie folgt aufgebaut. Ein so genannter *Rendezvous Point*, der nicht unbedingt Teilnehmer der Gruppe sein muss, verwaltet eine Liste mit allen Teilnehmern einer Gruppe. Diese Liste wird von neuen Teilnehmern vor dem Beitritt zur Gruppe angefordert (siehe Abbildung 2.15 Schritt (1) und (2)). Der neue Teilnehmer kontaktiert eine Untermenge der Gruppenmitglieder aus der angeforderten Liste, um daraus denjenigen zu bestimmen, mit welchem er eine Verbindung für die Datenübertragung aufbaut (Schritte (3) bis (5)). Tree First Ansätze unterstützen einen empfängerbasierten Multicastdienst und sind somit stark an das IP-Multicast angelehnt. Des Weiteren kennen die Teilnehmer der Gruppe nicht den gesamten Baum, sondern nur ihre direkten Nachbarn.

Nachdem ein Baum erstellt wurde, können zusätzliche Verbindungen zwischen den Teilnehmern aufgebaut werden, um diesen dynamisch zu optimieren. Dies führt somit zu einer vermaschten Topologie.

*Implizite Ansätze* wie SCRIBE [24], NICE [9] und *Content-Addressable Network (CAN)-Multicast* [130] erstellen den Verteilbaum und die Vermaschung quasi gleichzeitig. Sie unterstützen den Aufbau von großen Multicastgruppen.

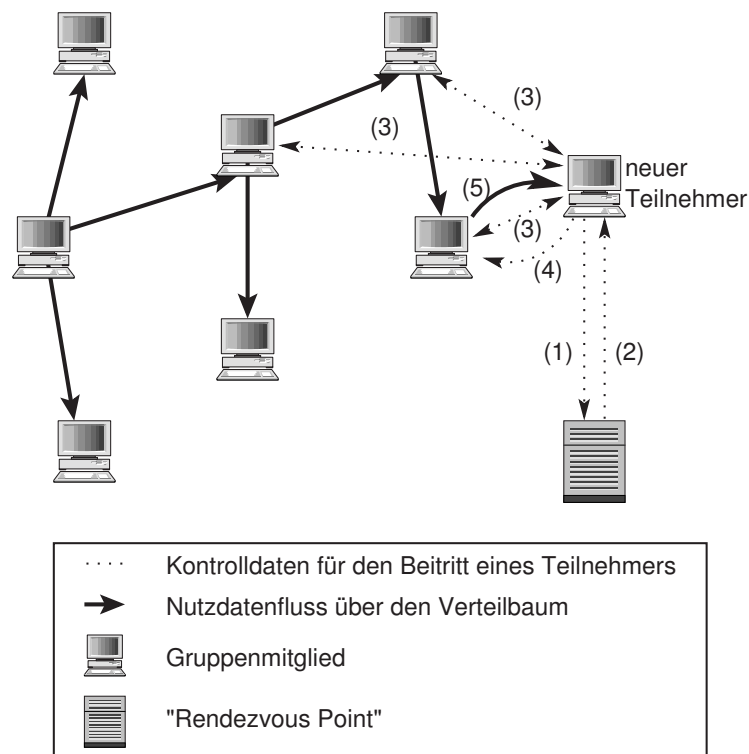


Abbildung 2.15.: Tree First Ansatz für Overlay Multicast

NICE basiert auf einer hierarchischen Topologie, bei der die Teilnehmer in einzelne Cluster unterteilt werden. Dadurch wird eine gleichmäßige Ausnutzung der Verbindungen zwischen den Teilnehmern gewährleistet. *CAN-Multicast* liegt einem  $n$ -dimensionalen kartesischen Raum zugrunde, wobei jeder Teilnehmer ein Teil dieses Raumes ist. Dieser Ansatz ist sehr komplex und ist nicht in der Lage, einen optimalen Multicastbaum zu erstellen [8]. SCRIBE basiert auf einer Verteilung der Daten über ein Peer-to-Peer-Netzwerk wie beispielsweise PASTRY [134]. Jedoch werden Funktionen für die Weiterleitung von Paketen auch von Teilnehmern des Peer-to-Peer-Netzes übernommen, die nicht unbedingt Teilnehmer der Multicastgruppe sind. Die Kopplung von Overlay Multicast auf Basis eines Peer-to-Peer-Netzes erhöht die Komplexität des SCRIBE Ansatzes.

Peer-to-Peer Ansätze sind grundsätzlich für die Übertragung kontinuierlicher Daten geeignet und können somit in Videokonferenzsystemen eingesetzt werden. Die Gruppenzusammensetzung kann jedoch nicht eingeschränkt wer-

den, da sich die bisherigen Ansätze sehr stark an IP–Multicast und dessen Eigenschaften anlehnen. Nachdem sich ein neuer Teilnehmer bei einem *Rendezvous Point* über die Teilnehmer einer Gruppe informiert hat, kann er ohne Weiteres der Gruppe beitreten. Um dies zu verhindern, müsste eine explizite Einladungsprozedur oder ein Sicherheitskonzept integriert werden.

Transportschicht

Über jedem Netzwerkprotokoll, ob nun IP, IP–Multicast, XCast oder Overlay Multicast, wird für die multimediale Echtzeit–Datenübertragung das *Universal Datagram Protocol (UDP)* genutzt<sup>2</sup>. UDP ist jedoch nicht in der Lage, Pakete zu ordnen, unterschiedliche Quellen zu unterscheiden oder Nutzlasttypen zu definieren. Dazu wird sich des bewährten *Realtime Transport Protocol (RTP)* in Verbindung mit dem *Realtime Transport Control Protocol (RTCP)* [137] bedient. RTP definiert einen Header für multimediale Datenpakete, der unter Anderem einen Nutzdatentyp, eine Sequenznummer, einen Quellidentifikator und einen Zeitstempel enthält. Das Protokoll enthält jedoch keine Protokollfunktionen, wie zum Beispiel die Wiederherstellung der Paketreihenfolge. Dies bleibt der Anwendung überlassen. RTCP wird hauptsächlich genutzt, um Feedback–Informationen an den Sender eines RTP–Stromes zu geben. Es werden aber auch teilnehmerspezifische Daten, wie Name, Telefonnummer, Email–Adresse und so weiter, mittels RTCP ausgetauscht.

## 2.7. Überblick über vorhandene Videokonferenzlösungen

Es gibt zur Zeit ein breites Spektrum von Videokonferenzsystemen, die zum Teil recht unterschiedliche Konzepte verfolgen. Dieser Abschnitt soll zunächst einen Überblick über derzeitige Videokonferenzsysteme geben, die in der Forschung und Praxis eingesetzt werden. Dabei werden nur die Ursprünge der Entwicklung und deren Hintergründe erläutert, bevor abschließend die wesentlichen Merkmale von Videokonferenzsystemen in Abschnitt 2.8 diskutiert werden.

H.32x

Von der ITU–T wurde eine Serie von Standards entwickelt, denen die meisten kommerziellen Videokonferenzsysteme zugrunde liegen. Die bedeutendsten

---

<sup>2</sup>In einem ATM–Netzwerk stehen statt dem UDP verschiedene *ATM Adaptation Layers (AAL)* zur Verfügung, die echtzeitfähig sein können, zum Beispiel AAL–2.

ten Standards sind dabei H.320 für Videokonferenzen über ISDN [76] und H.323 für Videokonferenzen über IP [78], der später auch für VoIP interessant wurde. Des Weiteren existieren Standards für das analoge Telefon und ATM [79, 74]. H.32x Systeme bestehen aus Endgeräten (Terminals), Gatekeepern, Gateways und Multipoint Control Units (MCUs). Endgeräte können sowohl Studio-, Roll-About- als auch Desktop-Systeme sein. Gatekeeper werden zur Verbindungsverwaltung, der Abrechnung, Autorisierung und gegebenenfalls zur Adressumwandlung benötigt. Gateways ermöglichen die Kopplung unterschiedlicher Standardsysteme wie H.323 und H.320. MCUs bestehen aus einem *Multipoint Controller (MC)* und einem *Multipoint Processor (MP)*. Sie realisieren Mehrteilnehmerkonferenzen, indem die Medienströme mehrerer Quellen im MC verwaltet und im MP gemixt werden. Bedeutende Studiosysteme sind Polycoms ViewStation und iPower [122], die Tandberg-Systeme 800 und 6000 [153], VCONS MC und Falcon Systeme [166], Intel ProShare [58] sowie das SONY Contact [143]. Desktop-Systeme mit Unterstützung durch Hardwarekarten werden durch die Firma VCON (Escort, ViGo Pro, Cruiser) und Polycom (ViaVideo) [122] angeboten. Reine Softwarelösungen sind Microsofts Netmeeting [102], CUSeeMe [174] und SUNs SUNForum [149]. Ebenfalls auf Software basieren so genannte Web-Konferenzen, beispielsweise von FVC [42], die einen Web-Browser als grafische Schnittstelle zum Nutzer anbieten.

Viele durch Universitäten und Forschungseinrichtungen entwickelte Videokonferenzsysteme sind Desktop-Systeme und basieren auf dem Multicast-Backbone [175]. Diese können kostenfrei genutzt werden. Die bedeutendste Entwicklung sind die Mbone-Tools VIC zur Videoübertragung [112], VAT [113], RAT [157] zur Audioübertragung und das Whiteboard WB [88]. Weiterhin existieren die Mbone-basierten Systeme USMINT [47] mit den Komponenten NEVIT zur Videoübertragung und NEVOT zur Audioübertragung sowie das Videokonferenzsystem RENDEZVOUS [131], die Weiterentwicklung des IVS-Systems [163]. Die Mbone-basierten Systeme sind jeweils selbständige Applikationen, die unabhängig von einander genutzt werden können. Außer einer vorhandenen Multicaststruktur sind keine Voraussetzungen für den Betrieb dieser Videokonferenzsysteme zu erfüllen. Mbone-basierte Systeme sind ausschließlich Desktop-Systeme, die ohne spezielle Hardware auskommen.

*Mbone-basierte  
Systeme*

Das Videokonferenzsystem der Firma DAVIKO GmbH [34] ist ein IP-basier-

*DaViKo*

tes serverloses Desktop-Videokonferenzsystem, das sich insbesondere durch die Verwendung von hochwertigen Kompressionsmethoden für die Audio- und Videodaten auszeichnet. Des Weiteren werden Application Sharing und die Aufzeichnung von Konferenzsitzungen unterstützt.

ATM-Entwicklungen

Mitte der neunziger Jahre wurden einige Videokonferenzsysteme entwickelt, die auf dem *Asynchronous Transfer Mode (ATM)* basieren. Diese Systeme wurden ebenfalls an Universitäten als Desktop-Systeme entwickelt. Beispiele sind ISABEL [37, 38], COVIS [12] und das System der Hong Kong University of Science and Technology HKUST [25]. Das an der BTU Cottbus entwickelte Videokonferenzsystem COVIS (*COttbuser VIDEOconference Service*) ist aus dem Ansatz GCSVA (*Group Communication and Scalability in Videoconferencing over ATM*) hervorgegangen. Aufgrund der rückläufigen Entwicklung des ATM-Marktes besitzen diese Systeme keine Einsatzperspektive mehr.

## 2.8. Gestaltungsmöglichkeiten von Videokonferenzsystemen

Die Vielzahl der Videokonferenzsysteme lässt auf den ersten Blick die Unzulänglichkeiten einzelner kaum erkennen. Aus diesem Grund werden die Merkmale

- Einsatzgebiet (lokal/global),
- Charakter (offen/geschlossen) und
- Topologie (zentral/dezentral/hierarchisch)

in Anlehnung an die in Abbildung 1.1 aufgestellten Anforderungen an ein Videokonferenzsystems diskutiert.

### 2.8.1. Einsatzgebiet

Videokonferenzsysteme können einerseits in lokalen Bereichen, wie beispielsweise einem Firmen- oder Campusnetzwerk, eingesetzt werden. Andererseits müssen Videokonferenzsysteme durch die Globalisierung in der Wirtschaft

auch über lokale Netzwerkgrenzen hinaus eine zufrieden stellende Qualität unterstützen. Sie müssen Strecken in globalen öffentlichen Netzen überwinden. Inwieweit Videokonferenzsysteme für globale Konferenzen eingesetzt werden können, hängt maßgeblich von drei Kriterien ab:

1. Die gegenwärtigen Videokonferenzsysteme nutzen ein breites Spektrum an Netzwerktechnologien, vom analogen Telefonnetz über ISDN, ATM bis hin zu IP. Hierbei stellen sich Fragen nach der Verfügbarkeit der genutzten Technologie und nach den Kosten für eine globale Konferenz.
2. In einem globalen Netzwerk können sich die Übertragungsbedingungen ständig ändern. Für Videokonferenzsysteme, die in einem globalen Umfeld eingesetzt werden sollen, ist es deshalb notwendig, sich an diese Änderungen anzupassen. Somit muss ein globales Videokonferenzsystem eine automatische Regelung der QoS-Parameter enthalten.
3. Globale Netzwerke können sehr heterogen sein. Zum einen kann die Bandbreite zwischen einzelnen Netzwerkknoten variieren, und zum anderen können die angeschlossenen Endsysteme sehr unterschiedliche Leistungsparameter haben. Deshalb ist es für globale Videokonferenzsysteme wichtig, heterogene Systemumgebungen zu unterstützen.

ISDN und das analoge Telefonnetz sind weltweit vorhanden, und es wird eine gesicherte Qualität der Verbindungen geboten. Jedoch verursachen global geschaltete Verbindungen hohe Kosten, wodurch der Betrieb von H.320 und H.324 Videokonferenzsystemen teuer werden kann. Da beide Netze nicht multicastfähig sind, entstehen weitere Kosten für den Betrieb einer MCU, um Mehrteilnehmerkonferenzen aufzusetzen.

ATM-Netzwerke spielen aufgrund der hohen Anschaffungskosten in der weltweiten Vernetzung von Computersystemen kaum eine Rolle. Zwar setzen vereinzelt lokale Netzwerke auf ATM auf, aber auch das ist eher eine Ausnahme. Somit können Videokonferenzsysteme wie COVIS, HKUST, ISABEL und Systeme nach dem H.324 Standard nicht für den Aufbau einer globalen Videokonferenz genutzt werden.

Die Internettechnologie hat sich in den letzten Jahren sehr rasant entwickelt und ist heute quasi weltweit verfügbar. Die Einfachheit der Best-Effort Paketvermittlung und die niedrigen Kosten haben zu diesem Erfolg geführt.

Netzwerktechnologie
---------------------

## 2. Gestaltungsprinzipien von Videokonferenzsystemen

---

Grundsätzlich wären also H.323, DAVIKO und die Mbone-Tools geeignet, um globale Videokonferenzen aufzusetzen.

Die Grundlage für die Videokonferenzsysteme RENDEZVOUS, USMINT und die Mbone-Tools bildet das Mbone auf Basis von IP-Multicast. Aufgrund der bereits erwähnten Einschränkungen dieser Technik können diese Systeme zukünftig kaum noch für globale Konferenzen eingesetzt werden.

### Adaptivität

Die Einfachheit des IP-Protokolls beschränkt es in der Nutzung für Echtzeitanwendungen wie Videokonferenzen. Das globale öffentliche Internet wird gleichzeitig von Millionen Nutzern genutzt, wodurch kaum ausreichend Ressourcen für Echtzeitanwendungen zur Verfügung gestellt werden können. Aus diesem Grund sind zusätzliche Mechanismen notwendig, die die Qualitätssicherung der Datenübertragung übernehmen. Wie die Diskussion über Integrated Services und Differentiated Services im Abschnitt 2.6 gezeigt hat, werden diese Ansätze heutzutage und auch in naher Zukunft nicht in einem breiten Maße eingesetzt. Demnach sind Nutzer eines IP-basierten Videokonferenzsystems auf den Best-Effort-Dienst des Internets angewiesen und müssen die Dienstqualität an die vorhandenen Ressourcen anpassen.

Bei den H.323-Systemen, den Mbone-Tools, DAVIKO, ISABEL und RENDEZVOUS existieren keine Mechanismen, um die Sendeleistung an die verfügbaren Ressourcen automatisch anzupassen. Stattdessen muss der Nutzer die Sendeleistung von Hand einstellen, weshalb die Bedienung dieser Werkzeuge einschlägige Erfahrungen vom Nutzer erfordert. Eine dynamische Anpassung an die zur Verfügung stehenden Bandbreite zwischen den Teilnehmern wird nur durch einige H.323-Systeme, beispielsweise der Firma VCON mit der Erweiterung *PacketAssist<sup>TM</sup>* [33] und Polycom mit der Erweiterung *Dynamic Bandwidth Allocation (DBA)* [116], sowie durch IVS [19] und USMINT mit dem in [141] vorgestellten Ansatz unterstützt.

### Heterogenität

Globale Netze, wie beispielsweise das Internet, können unterschiedlich leistungsstarke Komponenten enthalten. Die Bandbreite zwischen einzelnen Netzen variiert sehr stark, wodurch sich unterschiedliche Übertragungskapazitäten zwischen einzelnen Endsystemen ergeben. Des Weiteren kann die Leistung der Endsysteme sehr unterschiedlich sein. Diese Heterogenitäten muss ein Videokonferenzsystem berücksichtigen.



Die automatischen Bandbreiten-Anpassungsmechanismen *PacketAssist<sup>TM</sup>* und DBA arbeiten nur auf Basis von Punkt-zu-Punkt-Verbindungen. Sie sind nicht geeignet Mehrpunktverbindungen zu unterstützen. In USMINT und IVS wird der Teilnehmer einer Konferenz mit den geringsten Netzwerkressourcen berücksichtigt. Dadurch lassen sich Situationen nicht vermeiden, in denen Teilnehmer mit ausreichenden Ressourcen nur eine nicht zufrieden stellende Qualität geboten bekommen. Dies ist in den meisten Fällen nicht akzeptabel. Selbst mit der manuellen Steuerung in H.323-Systemen oder in den Mbone-Tools ist es nur möglich, die Sendeleistung an das leistungsschwächste Mitglied anzupassen. Ansonsten wären Stausituationen die Folge, die selbst leistungsstarke Verbindungen beeinträchtigen können.

RENDEZVOUS hat ein Layered-Coding-Kompressionsverfahren für die Videoübertragung integriert. Einzelne Qualitätsebenen werden an unterschiedliche Multicastgruppen gesendet, denen die Empfänger entsprechend ihrer Leistungsfähigkeit oder der Leistungsfähigkeit der Verbindung zum Sender beitreten können. Dieses Prinzip wird auch als *Layered Multicast* bezeichnet [100] (siehe Abbildung 2.16). Somit kann die Qualität der Videoübertragung sehr gut in heterogenen Netzwerken angepasst werden.

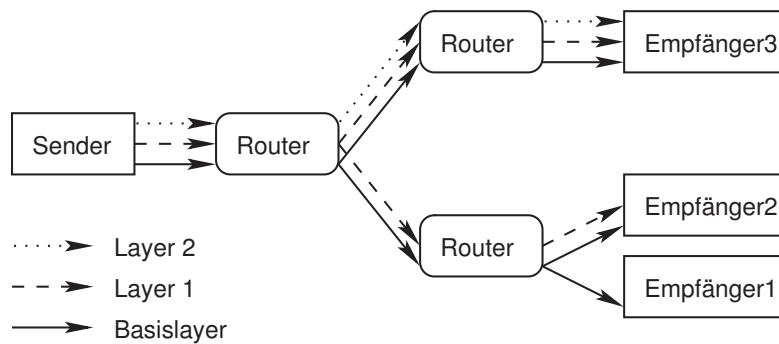


Abbildung 2.16.: Das Prinzip des Layered Multicast in RENDEZVOUS

Im angegebenen Beispiel sendet der Sender drei verschiedene Qualitätsebenen  $L_1$  bis  $L_3$  an drei Multicastgruppen. Der Empfänger 1 hat die höchste Leistung und tritt allen drei Multicastgruppen bei, die der Sender mit Daten beliefert. Es kann so die maximal mögliche Qualität empfangen. Die Empfänger 2 und 3 sind nicht so leistungsstark und treten deshalb nur den Multicastgruppen mit den zwei Ebenen beziehungsweise der einen Grundebene bei. Die Qua-

lität reduziert sich dabei bei dem Empfänger 3 auf eine minimal akzeptable Qualität.

COVIS und HKUST enthalten dynamische QoS-Mechanismen, die heterogene Systemumgebungen unterstützen. Sie berücksichtigen aufgrund der vorhandenen Reservierungsmechanismen in ATM jedoch nicht das Netzwerk, sondern nur das Endsystem. COVIS richtet sich im Gegensatz zu HKUST nach dem leistungsstärksten Teilnehmer in der Gruppe. Schwächere Teilnehmer müssen den empfangenen Medienstrom so skalieren, dass sie ihn problemlos verarbeiten können [13]. Da diese Ansätze zur QoS-Adaptierung nicht das Netzwerk betrachten, sind sie für globale Videokonferenzen in IP-Netzen kaum einsetzbar.

### 2.8.2. Offene und geschlossene Videokonferenzsysteme

Videokonferenzsysteme können die Kommunikation von offenen und geschlossenen Gruppen unterstützen. Der Charakter der Gruppe entscheidet darüber, für welche Anwendung das System geeignet ist.

Offene Konferenzen

Offene Videokonferenzsysteme, auch *lose gekoppelte (lightweight / informal / loose coupled) Systeme* genannt [83], finden Einsatz für öffentliche Übertragungen wie Space-Shuttle-Starts, Lehrveranstaltungen und TV-Kanäle. Dabei ist eine  $n : m$ -Kommunikationsbeziehung ( $n \ll m$  und  $n \neq m$ ) typisch. Im Extremfall ist  $n = 1$  und  $m \rightarrow \infty$ <sup>3</sup> Die Mehrzahl der heute existierenden Videokonferenzsysteme unterstützen offene Konferenzen, wie beispielsweise die Mbone-Tools, USMINT und RENDEZVOUS.

*Offen* kann bei den meisten existierenden Systemen damit gleichgesetzt werden, dass das Management der Konferenz nur rudimentär oder gar nicht vorhanden ist. Das fehlende Gruppenmanagement der Mbone-Tools und des RENDEZVOUS-Systems führt dazu, dass das System für jeden beliebigen Teilnehmer ohne Autorisierung einer anderen Partei zugänglich ist. Ein beliebiger Nutzer kann einer Gruppe beitreten und die übertragenen multimedialen Daten mitverfolgen, ohne dass die aktuellen Teilnehmer der Gruppe benachrichtigt werden. In USMINT wird zumindest eine explizite Einladung angeboten,

---

<sup>3</sup>Einen Spezialfall bildet eine 1:1-Kommunikation über Unicastverbindungen. In diesem Fall handelt es sich um eine geschlossene Gruppe, da sich hierbei jeder Kommunikationspartner kennt.

wodurch teilweise eine geschlossene Sitzung möglich ist. Jedoch ist bei Kenntnis der IP–Multicast–Adresse auch der unerlaubte Beitritt möglich.

Der Zugriff auf gemeinsam genutzte Ressourcen und die Einstellung der optimalen QoS–Parameter lassen sich in offenen Systemen kaum sinnvoll steuern, da die genaue Anzahl der beteiligten Endsysteme und deren Leistungsfähigkeit nicht bekannt ist. Einzig USMINT enthält eine Floorkontrolle *iFloor*[142], die moderierte und unmoderierte Sitzungen unterstützt. Da keine geschlossene Gruppensicht existiert, sendet jeder Teilnehmer, der das Rederecht erlangen möchte, periodisch ein diesbezügliches Signal an die Gruppe. Das Signal enthält einen Zeitstempel, mit dem jeder Teilnehmer eine zeitlich geordnete Sprecherwarteschlange zusammenstellen kann. Der Teilnehmer, der am Anfang dieser Liste steht, ist der aktuelle Sprecher. Zusätzlich wird diese Liste durch einen so genannten *Announcer* periodisch an die Gruppe verteilt, was allerdings zu einer weiteren Belastung des Netzwerkes führt. Die zeitliche Ordnung kann jedoch nur eindeutig realisiert werden, wenn die Uhren aller Teilnehmer synchronisiert sind. Mit dem *Network Time Protocol (NTP)* [104] sind zwar Genauigkeiten bis in den Nanosekundenbereich möglich, jedoch hängt dieser Wert stark von dem verwendeten Betriebssystem und der Qualität der Netzwerkverbindungen ab. Der Verlust einer Floorkontrollnachricht kann außerdem zeitweilig zu Inkonsistenzen der Sprecherwarteschlange führen. Somit ist eine zuverlässig und fair arbeitende Floorkontrolle nicht voll gewährleistet.

In einem *geschlossenen*, auch *eng gekoppelten (formal / tightly coupled)*, Videokonferenzsystem wird von einer bekannten und in der Regel kleinen Anzahl von Teilnehmern ausgegangen [83]. Die Teilnehmer einer Sitzung kennen sich untereinander, und jeder Beitritt eines neuen Teilnehmers wird allen Partnern angezeigt. Dadurch wird ein gewisser Grad an Vertraulichkeit erreicht, der dem des Telefondienstes entspricht. Geschlossene Videokonferenzsysteme finden vorwiegend Anwendung bei internen Firmengesprächen, privaten Treffen oder Konsultationen, zum Beispiel in der Medizin, im Rechtswesen oder in der Finanzberatung. Ein Einsatz in Teleteaching– und Telelearning–Umgebungen ist ebenfalls vorstellbar, zum Beispiel um Seminare und Diskussionsrunden durchzuführen. Längst haben Videokonferenzsysteme schon in Gerichtssälen Einzug gehalten, so dass auch im Bereich des *e–Government* Anwendungen gefunden werden können. Geschlossene Videokonferenzsysteme haben in der

<i>geschlossene Konferenzen</i>
-------------------------------------

Praxis eine größere Bedeutung als offene Konferenzen, da sie letztlich Treffen und Beratungen des täglichen Lebens abbilden, die eine weitaus größere Häufigkeit haben als öffentliche Veranstaltungen.

Jeder Teilnehmer in einem geschlossenen Videokonferenzsystem ist, ähnlich dem Peer-to-Peer Konzept, gleichzeitig Sender und Empfänger von Audio- und Videodaten, wodurch eine  $n:n$ -Kommunikationsbeziehung existiert. Die Systeme enthalten ein explizites Gruppen- und QoS-Management sowie zu meist eine Floorkontrolle. Das QoS-Management ermöglicht in Zusammenarbeit mit der Gruppenverwaltung eine individuelle Anpassung der Übertragungsqualität an jeden Teilnehmer der Gruppe.

H.32x-Systeme, DAVIKO, ISABEL, HKUST und COVIS unterstützen eine geschlossene Gruppenkommunikation<sup>4</sup>. Das Management in H.32x, ISABEL und HKUST ist zentral organisiert (siehe Abschnitt 2.8.3), wodurch auch der geschlossene Charakter der Konferenz gesichert wird. COVIS hingegen ist dezentral organisiert und basiert auf einem entsprechenden Transportdienst, der geschlossene Konferenzen ermöglicht. Des Weiteren enthält COVIS eine Floorkontrolle, die moderierte und nicht moderierte Sitzungen erlaubt sowie ein automatisches QoS-Management, das sich auf die Endsysteme bezieht. H.32x, DAVIKO, ISABEL und HKUST verfügen weder über eine Floorkontrolle noch über ein komfortables QoS-Management.

*Geschlossenheit  
vs. Sicherheit*

Die Geschlossenheit der Gruppe wird oft mit einer sicheren, vertraulichen Kommunikation gleichgesetzt. Geschlossene Gruppen sind a priori nicht vollständig vertraulich. Sie bieten eine mit dem Telefonnetz vergleichbare Vertraulichkeit. Durch gezielte Angriffe innerhalb des Kernnetzes können deshalb vertrauliche Inhalte in falsche Hände gelangen. Eine entsprechend starke Verschlüsselung der Daten kann dies verhindern. Jedoch werden dafür zusätzliche Ressourcen benötigt, um das Datenaufkommen einer Videokonferenz bewältigen zu können. Zudem entsteht ein Mehraufwand für die Verteilung der Schlüssel. Auf der anderen Seite sind vertrauliche und verschlüsselte Konferenzen nicht gleich geschlossene Konferenzen. Werden verschlüsselte Daten über einen empfangerbasierten Multicastdienst, wie dem IP-Multicast, verschickt, weiß der Sender trotzdem nicht genau, wer die Daten in welcher Qualität empfängt beziehungsweise für eine mögliche Entschlüsselung aufzeichnet.

---

<sup>4</sup>H.323 und DAVIKO können das IP-Multicast nutzen, wodurch auch offene Konferenzen möglich werden.

Die Gewährleistung der Sicherheit in Videokonferenzsystemen bezüglich Authentifizierung, Authentizität und Integrität ist ein eigenständiges Forschungsgebiet, das insbesondere unter dem Aspekt des verteilten Konferenzmanagements noch eine Reihe ungelöster Probleme enthält. Die Lösung dieser Probleme geht über die Zielstellung dieser Arbeit hinaus. Sie werden in einem anderen Promotionsvorhaben am Lehrstuhl erarbeitet. Deshalb wird die Diskussion hinsichtlich dieser Problematik nicht weiter vertieft.

### 2.8.3. Topologien von Videokonferenzsystemen

Bei der Gestaltung von Videokonferenzsystemen spielt die Wahl der Topologie eine entscheidende Rolle, wobei zentralisierte und verteilte Ansätze vorrangig genutzt werden. Daneben gibt es hierarchische Videokonferenzsysteme sowie hybride Ansätze, die unterschiedliche Topologien für die Signalisierung und die Medienübertragung benutzen.

Zentralisierte Systeme, wie beispielsweise die H.32x Systeme, lehnen sich stark an das Client–Server–Modell an. Clients sind teilnehmende Hosts der Konferenz, die sich mit dem Server in Verbindung setzen (siehe Abbildung 2.17). Der Server kontrolliert eine oder mehrere voneinander unabhängige Sitzungen und regelt die Verteilung der Medienströme an die Clients. Dazu gehört zumeist auch, dass der Server die Medienströme aller Teilnehmer einer Sitzung empfängt und anschließend mixt. Der gemixte Strom wird den Teilnehmern zugestellt. Somit benötigen die Clients nur eine geringe Funktionalität, um die Medien aufzunehmen und abzuspielen.

Zentrale Systeme

Zentralisierte Systeme sind recht einfach zu realisieren und zu handhaben, da alle Managementfunktionen im Server durchgeführt werden. Im H.32x–Standard werden dazu ein *Gatekeeper* und eine *Multipoint Control Unit (MCU)* definiert. Der Gatekeeper steuert den Verbindungsauf– und –abbau. Die MCU verteilt die Audio– und Videoströme an die einzelnen Teilnehmer.

Zentralisierte Ansätze haben den Nachteil, dass der gesamte Datenverkehr der Konferenz den Server durchläuft. Dieser kann dadurch zum Flaschenhals der Kommunikation werden. Aus diesem Grund werden MCUs als Hardwarelösungen mit zum Teil noch sehr hohen Anschaffungskosten angeboten. Eine MCU von einem Provider zu mieten, senkt zwar die Anschaffungskosten, aber dafür sind die Betriebskosten höher, und der Nutzer ist von der Verfügbarkeit

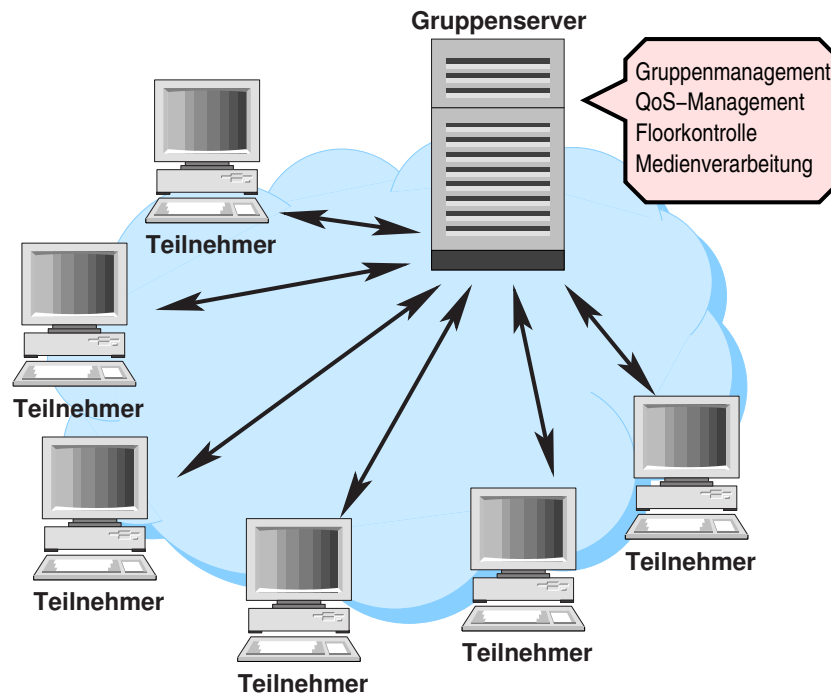


Abbildung 2.17.: Beispiel einer zentralisierten Architektur eines Videokonferenzsystems

des Providers abhängig. Ein weiterer Nachteil einer Hardwarelösung ist die geringe Flexibilität in Bezug auf neue Technologien und die schlechte Skalierbarkeit in Bezug auf die Gruppengröße sowie parallel laufenden Konferenzsitzungen. Eine handelsübliche MCU ist für 4, 8 oder 16 Teilnehmer ausgelegt. Ist die MCU ausgelastet, können keine weiteren Teilnehmer der Konferenz beitreten. Erweiterungen einer Einheit sind kaum möglich. Des Weiteren ist die Steuerung parallel laufender Konferenzen durch eine MCU beschränkt. Einfache MCUs unterstützen nur eine Sitzung. Preisintensive MCUs, wie sie beispielsweise von CISCO im DFN-Videokonferenzdienst (DFNVC) [127, 168] eingesetzt werden, unterstützen maximal 100 Teilnehmer, das heißt 33 Dreier-, 25 Vierer-, 20 Fünfer-Konferenzen et cetera. Größere Gruppen oder eine höhere Anzahl paralleler Sitzungen ist nur unter Anschaffung weiterer MCUs möglich. Um diese Problem zu lösen, wurden Software-MCUs entwickelt, wie beispielsweise der *Microsoft Exchange Video Server* [101]. Damit ist es möglich, unterschiedliche Sitzungen aufzusetzen und parallel abzuhalten. Die Leistungsfähigkeit des Videokonferenz-Servers muss jedoch recht hoch sein, um nicht der Flaschen-

hals der Kommunikation zu werden. Bei der Überlastung des Server können keine weiteren Teilnehmer einer laufenden Konferenz hinzugefügt werden und auch keine neuen voneinander unabhängige Konferenzen aufgebaut werden. Um eine MCU oder einen Server vor Überlast zu schützen, werden die Nutzungsrechte und -zeiten der jeweiligen Komponente von einer übergeordneten Instanz verwaltet. Allerdings unterstützt dieses Konzept spontane Sitzungen kaum.

Der Server selbst stellt darüber hinaus einen Single-Point-of-Failure dar. Fällt beispielsweise die MCU aus, bricht die gesamte Videokonferenz zusammen. H.323 lässt es in diesem Zusammenhang zu, dass die MCU-Fähigkeiten in einem Endsystem integriert werden können, wie zum Beispiel bei Microsoft Netmeeting [102] oder ISABEL. Der Initiator einer Konferenz startet seine Endsystem-Komponenten und die MCU-Komponente für die Konferenz. Bei Verlassen oder bei Fehlverhalten des Initiators kann jedoch die Konferenz beendet werden, obwohl die verbleibenden Teilnehmer noch weiter kommunizieren wollen. Dasselbe würde bei Ausfall der Verbindung zu der MCU passieren.

In globalen Konferenzen ist ein zentraler Ansatz auch deshalb ungeeignet, weil sich die Verzögerungen durch die globalen Entfernungen negativ bemerkbar machen können. Um dies zu verdeutlichen, soll folgendes Szenarium als Beispiel dienen. Der Konferenzserver einer überregional arbeitenden Firma steht an ihrem Hauptsitz. Zwei oder mehr Teilnehmer, die sich in einer entfernten Zweigstelle befinden, müssen dann auf dem Umweg über den Server am Hauptsitz kommunizieren. Die Verzögerung der Medienströme wäre enorm.

In dezentralen Lösungen übernimmt jedes Endsystem die Aufgaben der Verwaltung, Signalisierung und Medienübertragung. Sie sind vergleichbar mit dem Peer-to-Peer-Konzept. Dezentrale Lösungen unterstützen das spontane Aufsetzen von Mehrteilnehmerkonferenzen, ohne dabei von Providern abhängig zu sein. Sie sind flexibel einsetzbar und vermeiden Leistungsengpässe. Im Gegensatz zu zentralen Ansätzen sind sie weniger fehleranfällig und preiswerter. Des Weiteren beschränkt sich die Administration auf die Endsysteme. Server-, spezielle Netzwerk- und MCU-Konfigurationen sind nicht notwendig. Dezentrale Systeme sind besser skalierbar. Obwohl das Endsystem eines Teilnehmers ausgelastet sein kann, können trotzdem weitere Teilnehmer der Gruppe beitreten. In Bezug auf das gleichzeitige Aufbauen von mehreren voneinander unabhä-

<i>Dezentrale Systeme</i>
---------------------------

## 2. Gestaltungsprinzipien von Videokonferenzsystemen

gen Konferenzen sind dezentrale Ansätze ebenfalls vorteilhafter als zentrale Lösungen. Der Aufwand für die Verwaltung einer dezentral organisierten Konferenz ist jedoch im Allgemeinen höher, da alle Verwaltungsinformationen auf jedem teilnehmenden Endsystem gespeichert und bearbeitet werden müssen (siehe Abbildung 2.18). Hinzu kommt ein Mehraufwand im Netzwerk für die Verteilung der Managementdaten.

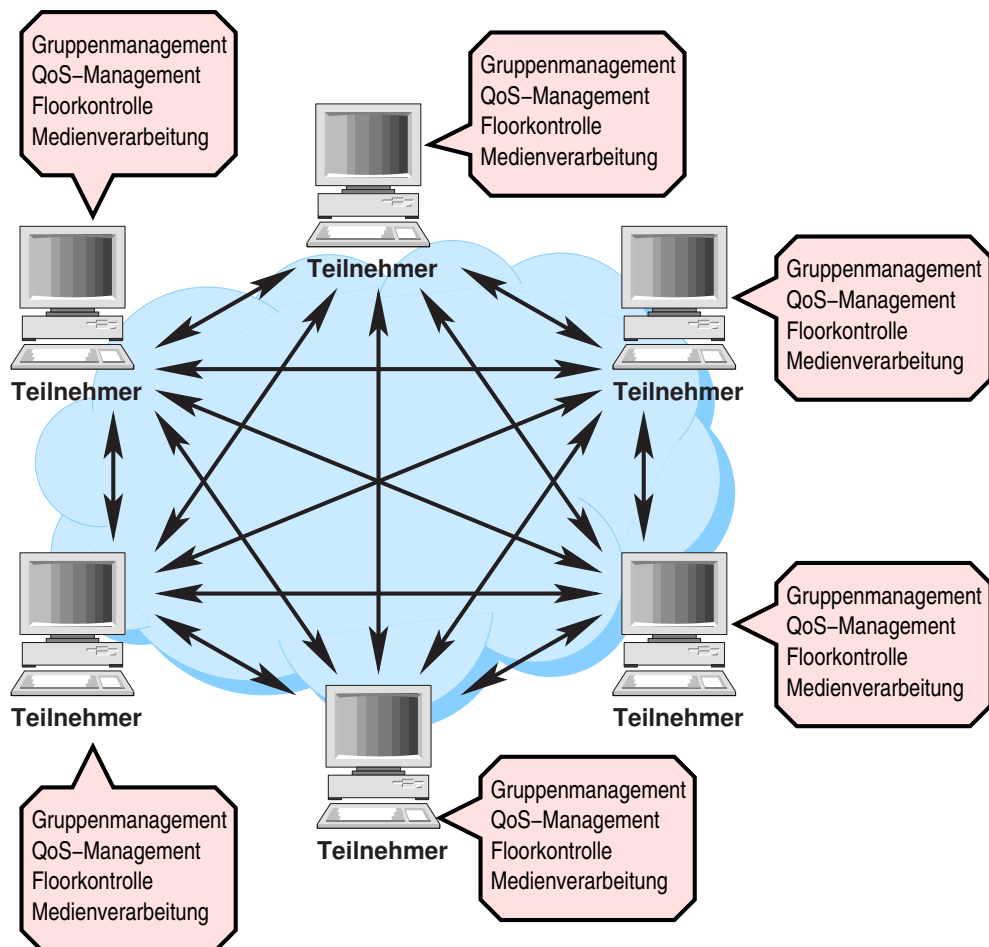


Abbildung 2.18.: Beispiel einer dezentral organisierten Topologie eines Videokonferenzsystems

Die Mbone-basierten Videokonferenzsysteme wie die Mbone-Tools, US-MINT und RENDEZVOUS sowie DAVIKO sind dezentral organisiert. Eine konsistente Verwaltung der Steuerdaten ist jedoch in diesen Systemen nicht vorgesehen und wäre auch kaum möglich. Dadurch entsteht nur ein sehr geringer Managementaufwand. COVIS ist ebenfalls ein dezentrales Videokonferenzsystem.



tem, das aber im Gegensatz zu den Mbone-basierten Systemen geschlossene Konferenzen unterstützt. Hierbei spielt die Erhaltung der Konsistenz der Daten des Gruppen- und QoS-Managements sowie der Floorkontrolle eine sehr wichtige Rolle. Der Mehraufwand für die Verwaltung der Konferenzsitzung ist also deutlich höher als bei den Mbone-basierten Systemen. Der Kern der Lösung ist ein Transportdienst, der die Konsistenz verteilter Managementdaten sichert (siehe Abschnitt 4.2).

Mbone-basierte Systeme und COVIS nutzen IP- beziehungsweise ATM-Multicast, wodurch sich der Aufwand für das Verbindungsmanagement reduziert sowie Ressourcen für die Audio- und Videoübertragung geschont werden. Allerdings haben sich, wie bereits erwähnt, Multicastdienste im Internet bisher nicht durchsetzen können und werden nur von wenigen Anwendern unterstützt. Es ist nicht zu erwarten, dass sich dies in naher Zukunft ändert. Somit stellt sich das Problem, wie sich dezentrale Konferenzsysteme in der Zukunft effizient mit Unicast-Diensten gestalten lassen.

Die Vor- und Nachteile von zentralen und dezentralen Topologien vereinen hierarchische Systeme. Sie sind jedoch nicht sehr verbreitet. Hierarchische Systeme bestehen in der untersten Ebene aus kleinen Konferenzgruppen, die über einen zentralen Punkt, ein Gateway, miteinander verbunden sind. Dieser Verbund kann dann wieder einen Zugang zu einer weiteren höheren Ebene bilden (siehe Abbildung 2.19). Somit verteilen sich die Fehlerpunkte der Kommunikation auf mehrere Gateways. Fällt beispielsweise ein Gateway zwischen Ebene 1.1.2 und Ebene 1.1 aus, können die verbleibenden Gruppen trotzdem die Kommunikation fortsetzen. Etwas kritischer ist der Ausfall eines Gateways der oberen Ebene, da somit die überregionale Kommunikation zum Erliegen kommt. Aber die Kommunikation der unteren Ebenen bleibt davon unberührt. Vorteilhaft in hierarchischen Systemen ist die Steuerung und Verwaltung von Dienstgüteeigenschaften, die für eine Untergruppe zentral in den Gateways realisiert werden kann. Ein weiterer Vorteil besteht darin, dass die verschiedenen Ebenen unterschiedliche Architekturen und Topologien enthalten können. Somit besteht die Möglichkeit, mit Systemen unterschiedlicher Art zu kommunizieren. Grundsätzlich bilden die Gateways aber auch einen Engpass der Kommunikation, der durch ein geschicktes QoS-Management in den einzelnen Ebenen vermieden werden kann.

<i>Hierarchische Systeme</i>
----------------------------------

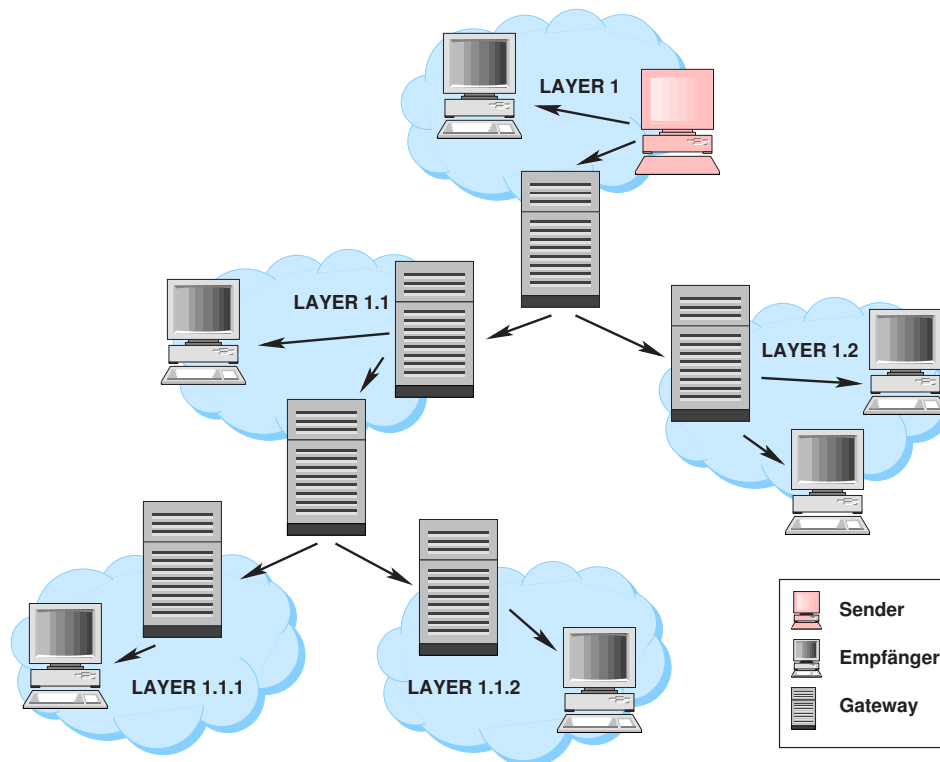


Abbildung 2.19.: Beispiel einer hierarchischen Gruppenkommunikation

Der H.32x-Standard bietet die Möglichkeit einer hierarchischen Kommunikationsstruktur, indem MCUs kaskadiert werden. Hierbei wird eine MCU mit einer oder mehreren anderen MCUs verbunden. Somit kann auf eine einfache Art ein hierarchisches System aufgebaut werden. Abbildung 2.20 zeigt als Beispiel die Struktur des DFN-Videokonferenzdienstes (DFNVC) [96, 168]. Die MCUs wirken als Gateway zwischen verschiedenen H.323-Zonen, die über das G-WiN verbunden sind. Zu beachten ist, dass die H.32x-Standards ebenfalls *Gateways (GW)* definieren. Ihre Aufgabe besteht jedoch nur darin, Umwandlungen durchzuführen, die an den Standardübergängen notwendig sind. Beispiele sind H.323/H.320-Gateways, die IP mit dem ISDN verbinden (siehe Abbildung 2.20). Eine beliebige Kaskadierung ist jedoch nicht möglich. Einerseits werden die Bilder der gemixten Videoströme sehr klein, und andererseits wird die Verwaltung zu aufwendig.

Hybride Systeme kombinieren diverse Kommunikationstopologien miteinander. Eine Kombination aus einer zentralen Steuerung und einer dezentralen

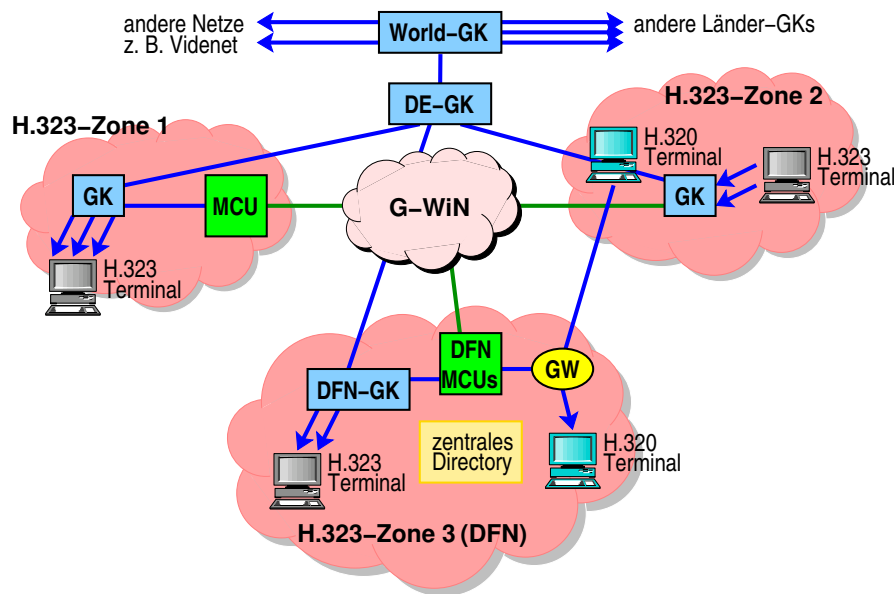


Abbildung 2.20.: Die H.323-Struktur am Beispiel des DFN-Videokonferenzdienstes DFNVideoConference nach [97]

Verteilung der Mediendaten ist beispielsweise bei HKUST und bei H.323 unter Nutzung von IP-Multicast zu finden. Dies hat den Vorteil, dass der Verwaltungsaufwand gering ist und die Mediendaten kostengünstig verteilt werden können.

Die Kombination aus einer zentralen Steuerung einer Konferenz und einer hierarchischen Verteilung der Mediendaten hat den Vorteil, dass Bedürfnisse einzelner Gruppen bezüglich der Dienstqualität gesondert befriedigt werden können, die Verwaltung aber einen geringen Aufwand hat. Mit dem H.323-Standard ist diese hybride Topologie ebenfalls möglich, indem die *Multipoint Processors (MPs)* der MCUs hierarchisch verteilt, aber durch einen *Multipoint Controller (MC)* zentral gesteuert werden.

## 2.9. Zusammenfassung

Dieses Kapitel hat einen Überblick über den aktuellen Stand der Entwicklung im Bereich der Videokonferenzsysteme gegeben. Es wurden Grundkonzepte wichtiger Videokonferenzsysteme vorgestellt und Anforderungen an ihre Gestaltung diskutiert. Darauf aufbauend wurde der Frage nachgegangen, ge-

## 2. Gestaltungsprinzipien von Videokonferenzsystemen

---

geschlossene Videokonferenzen über globale, öffentliche Netzwerke durchzuführen. Die Analyse einiger bekannter Videokonferenzsysteme in Bezug auf ihre Tauglichkeit für globale, geschlossene Konferenzen ist in Tabelle 2.1 zusammengefasst.

	H.320	H.323	MBone-Tools	USMINT	RENDEZVOUS	HKUST	ISABEL	DAVIKO	COVIS
Einsatz in globalen, öffentlichen Netzen	x) <sup>5</sup>		x) <sup>6</sup>	x) <sup>6</sup>	x) <sup>6</sup>			x	
Einsatz in heterogenen Umgebungen						x			x
Unterstützung geschlossener Gruppen	x	x				x	x		x
dezentraler Ansatz			x	x	x			x	x
Unterstützung spontaner Konferenzen				x		x	x	x	x
explizites Gruppenmanagement	x	x		x		x	x		x
explizite Floorkontrolle				x					x
integriertes QoS-Management				x		x			x

Tabelle 2.1.: Analyse bekannter Videokonferenzsysteme hinsichtlich einiger wichtiger Merkmale

Aus dieser Übersicht geht hervor, dass die derzeitigen Videokonferenzsysteme nicht oder nur unzureichend für die Durchführung von globalen Konferenzen geeignet sind. Einzig die ISDN-basierte Lösung H.320 bietet weltweit

<sup>5</sup>sofern ISDN zur Verfügung steht

<sup>6</sup>sofern IP-Multicast zur Verfügung steht

garantierte Bandbreiten. H.320 unterstützt auch geschlossene Gruppen. Jedoch sprechen die hohen Betriebskosten gegen den globalen Einsatz. Durch den zentralen Ansatz im H.320 für Mehrteilnehmerkonferenzen ist man außerdem gezwungen, sich diese teure Hardware zu kaufen oder bei einem Provider zu mieten, was das Aufsetzen spontaner Konferenzen nicht erlaubt. Der zentrale Ansatz birgt außerdem die Gefahr eines Flaschenhalses der Kommunikation und eines Single-Point-of-Failure. Durch die fehlende Floorkontrolle und das fehlende QoS-Management ist die Bedienbarkeit eines H.320-Systems beschränkt.

Die Übersicht zeigt, dass sich IP-basierte Systeme für den täglichen Gebrauch kaum für eine globale Mehrteilnehmer-Videokonferenz über öffentliche Netzwerke eignen, bei denen geschlossene Beratungen durchgeführt werden sollen. Stattdessen wird in der Praxis oft auf die Nutzung von Video verzichtet, und die Sprachkommunikation über das Telefonnetz realisiert. Dies ist darin begründet, dass zum einen derzeit kaum geeigneten automatisierten Mechanismen existieren, die die Qualität der Übertragung von Echtzeitdaten an die vorhandenen Ressourcen anpassen, wie bei H.323, den Mbone-Tools und RENDEZVOUS. Dadurch leidet die Bedienbarkeit und die Qualität eines Systems. Zum anderen werden kaum Ansätze unterstützt, die heterogene Netzwerk- und Systemumgebungen unterstützen. Nur RENDEZVOUS bietet einen manuell gesteuerten Mechanismus, der allerdings vom Nutzer ein gewisses Maß an Erfahrung erfordert. Außerdem setzen RENDEZVOUS, die Mbone-Tools und USMINT auf eine Technik, die Tendaussagen zufolge an Relevanz verliert — das IP-Multicast [8].

Aus Tabelle 2.1 ist ebenfalls ersichtlich, dass nur wenige Systeme geschlossene Gruppen unterstützen, obwohl die Mehrzahl der Beratungen einen geschlossenen Charakter haben. H.320-Systeme und ATM-basierte Systeme erlauben geschlossene Beratungen, was allerdings mit sehr hohen Betriebs- und Anschaffungskosten verbunden ist. IP-basierte H.323-Videokonferenzsysteme unterstützen geschlossene Gruppen. Dies wird jedoch bisher durch einen zentralisierten Ansatz unter Verwendung von MCUs erreicht. Im praktischen Einsatz weisen diese Systeme, insbesondere Softwarelösungen, relativ hohe Ausfallraten auf. Außerdem sind sie bezüglich der Gruppengröße und der Anzahl parallel laufender Konferenzen nur schlecht skalierbar. Wie bei H.320 können

kaum spontane Mehrteilnehmerkonferenzen abgehalten werden. Eine dezentralisierte Lösung ist für solche Fälle besser geeignet. Die zur Zeit existierenden dezentralen IP-basierten Lösungen sind jedoch nicht für die Durchführung von globalen Konferenzen über öffentliche Netzwerke geeignet, da sie zumeist auf IP-Multicast aufsetzen. Zudem werden heterogene System- und Netzumgebungen von den meisten Systemen nicht unterstützt.

Aus diesen Überlegungen heraus, leitet sich die Frage ab, wie globale, geschlossene Videokonferenzen im öffentlichen Internet realisiert werden müssen. Wie kann eine ausreichend gute Qualität in globalen Netzen gesichert werden? Wie wird die Geschlossenheit einer globalen Gruppe ohne einen zentralen Ansatz erreicht? Diese Fragen sollen in den folgenden Abschnitten beantwortet werden.

---

## 3. Gestaltung geschlossener Videokonferenzsysteme über das Internet

---

Die Analyse der derzeit vorhandenen Videokonferenzlösungen hat gezeigt, dass diese meist nur unzureichend die Anforderungen an eine Videokonferenz für geschlossene Gruppen in öffentlichen Netzen unterstützen. Um die Frage zu beantworten, wie geschlossene Videokonferenzen über das Internet gestaltet werden müssen, werden im Folgenden zwei verschiedene, neuartige Ansätze untersucht und diskutiert. Der erste Ansatz namens OCTOPUS [26, 27, 86] untersucht die Möglichkeiten, einen globalen Videokonferenzverbund durch das Zusammenschalten von lokalen Videokonferenzsystemen aufzubauen. Aufgrund der daraus gewonnenen Erkenntnisse wird ein Konzept für ein eigenständiges, dezentral organisiertes Videokonferenzsystem mit dem Namen BRAVIS entworfen und diskutiert. Dieses Kapitel gibt einen konzeptuellen Überblick über die beiden Ansätze, wie beispielsweise über die Architektur, die Funktionalität und die Topologie. Die Lösung von speziellen Teilproblemen werden in den Kapiteln 4 und 5 diskutiert.

### 3.1. Der OCTOPUS–Ansatz

#### 3.1.1. Systemarchitektur des OCTOPUS–Ansatzes

Die Grundidee von OCTOPUS ist der Aufbau einer globalen Videokonferenz durch das Verbinden verschiedener lokaler Videokonferenzen über größere Entfernungen, zum Beispiel über das Internet, basierend auf UDP/IP und RTP, oder über ATM–Netzwerke wie dem DFNATM–Dienst (siehe Abbildung 3.1).

<i>Idee des Ansatzes</i>
--------------------------

### 3. Gestaltung geschlossener Videokonferenzsysteme über das Internet

Damit sollen globale Konferenzen für geschlossene Gruppen mit einer größeren Anzahl von Teilnehmern aufgebaut werden können. Die globale Konferenz verfügt über eine gemeinsame Floorkontrolle und ein dynamisches globales Gruppenmanagement. Als lokale Videokonferenzsysteme werden hier die Systeme an den jeweiligen Standorten bezeichnet, die auf lokalen oder regionalen Netzwerken laufen. Das Besondere des verfolgten Ansatzes ist dabei, dass OCTOPUS keinerlei Einschränkungen an die Architektur des lokalen Videokonferenzsystems stellt. Es sollen, ähnlich dem Prinzip offener Systeme, Vi-

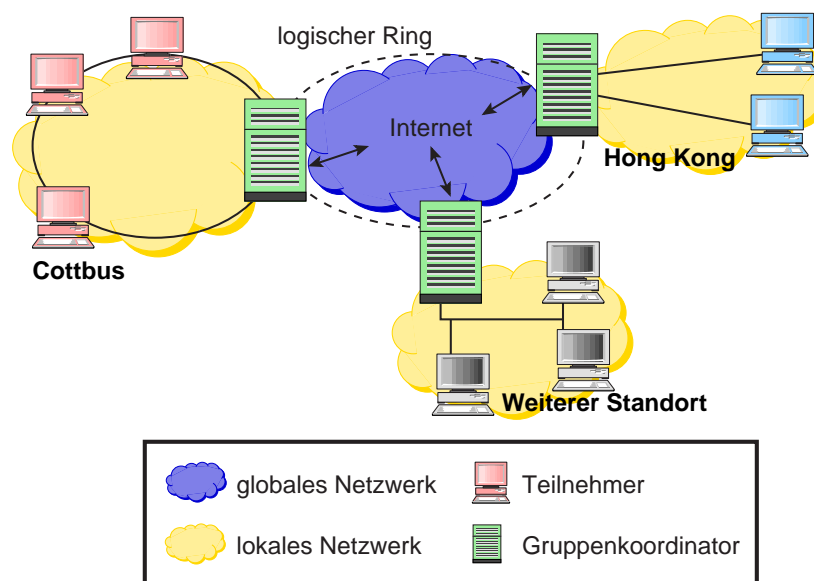


Abbildung 3.1.: Die Struktur des OCTOPUS-Systems

deokonferenzsysteme unterschiedlicher Struktur und Technologie miteinander verbunden werden können. Der Charakter der geschlossenen Gruppenkommunikation der lokalen Videokonferenzsysteme bleibt jedoch erhalten. Die durch OCTOPUS verbundenen Videokonferenzen bilden dann eine globale Videokonferenz, die alle lokalen Teilnehmer zu einer globalen Gruppe vereint.

Neuartige Funktionalitäten

Neben der Möglichkeit, verschiedenartige Videokonferenzsysteme miteinander zu koppeln, enthält der OCTOPUS-Ansatz ein Gruppenmanagement, das dynamisch den Beitritt und das Verlassen eines Teilnehmers oder einer Teilnehmergruppe unterstützt. Des Weiteren wird erstmals eine einheitliche Schnittstelle für eine verteilt organisierte Floorkontrolle realisiert. Besonders hervorzuheben ist die Entwicklung eines dynamischen QoS-Managements, das neben



den Netzwerk- und Systemressourcen auch die Wünsche des Nutzers berücksichtigt (siehe Kapitel 5).

Das OCTOPUS-Videokonferenzsystem ist in zwei Ebenen unterteilt (siehe Abbildung 3.2). Die untere Ebene umfasst die teilnehmenden lokalen Video-

2-Ebenen-  
Struktur

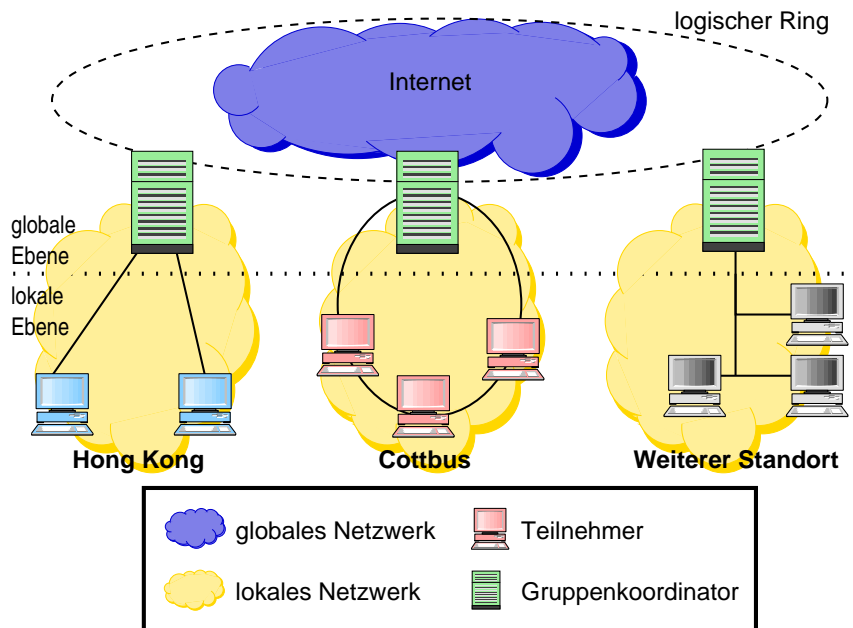


Abbildung 3.2.: Ebenenstruktur von OCTOPUS

konferenzsysteme, die autonom in ihrem jeweils vorgegebenen Operationsmodus arbeiten. Die obere Ebene realisiert die Interaktion zwischen den lokalen Konferenzsystemen. Die Schnittstelle der lokalen Systeme wird über einen so genannten *Gruppenkoordinator* realisiert, der in jedes lokale Konferenzsystem zu integrieren ist und die externe Kommunikation der lokalen Konferenz übernimmt. Der Gruppenkoordinator fungiert quasi als Gateway des lokalen Videokonferenzsystems. Er stellt eine einheitliche Schnittstelle für die globale Konferenz bereit, die die möglicherweise unterschiedliche Architektur der lokalen Konferenz verdeckt. Die globale Ebene übernimmt das Management der Konferenz. Es ist dezentral organisiert. Die Managementaktionen umfassen unter anderem die Floorkontrolle, die Dienstgüte-Regulierung und die Übertragung der Video- und Audioströme zwischen den lokalen Konferenzen.

Mit dem Aufbau einer globalen Konferenz werden die Teilnehmer der lokalen Konferenzen zu einer geschlossenen globalen Gruppe vereint. Diese Gruppe

wird global verwaltet. So wird beispielsweise der Beitritt und das Ausscheiden lokaler Teilnehmer allen Teilnehmern der Konferenz angezeigt.

Für den Austausch der Signalisierungsdaten zwischen den Gruppenkoordinatoren wird das Gruppenkommunikationsprotokoll GCP eingesetzt, das für das Gruppenmanagement des Cottbuser Videokonferenzsystems COVIS (siehe Kapitel 4) entwickelt wurde. Dabei bilden die Gruppenkoordinatoren einen logischen Ring. Das GCP kann jedoch nicht in seiner ursprünglichen Fassung genutzt werden, da es auf einem senderbasierten Multicastdienst aufbaut, der ursprünglich durch ATM bereitgestellt wurde. Aus diesem Grund musste die Protokollfunktionalität des GCP neu gestaltet werden. Der Austausch der Audio- und Videodaten zwischen den Gruppenkoordinatoren erfolgt Punkt-zu-Punkt, um den geschlossenen Charakter der lokalen Konferenzen zu erhalten. Eine Lösung mittels IP-Multicast kommt daher nicht in Frage. Zur Übertragung der Mediendaten wird in der Prototypimplementierung RTP [137] genutzt. Es werden nicht alle Videoströme der Teilnehmer übertragen. In Abhängigkeit von der zur Verfügung stehenden Bandbreite und den Darstellungsmöglichkeiten in den Endsystemen werden nur ausgewählte Videoströme übertragen. Der dafür in den lokalen Videokonferenzen verwendete Auswahlmechanismus wird in Kapitel 5 beschrieben.

#### 3.1.2. Floorkontrolle

In einer Gruppendiskussion mit einer hohen Teilnehmerzahl ist es schwierig, die Disziplin in Bezug auf den Zugriff auf gemeinsam benutzte Ressourcen zwischen den Teilnehmern aufrecht zu erhalten. Deshalb ist es in einem Videokonferenzsystem mit einer hohen Teilnehmerzahl besonders wichtig, eine Floorkontrolle zu integrieren.

Durch das Zusammenschalten der lokalen Videokonferenzen zu einer globalen Konferenz werden die lokalen Floorkontrollstrategien, die die Vergabe des Rederechts und den Zugriff auf gemeinsame Ressourcen wie Whiteboard oder Telepointer regeln, durch eine globale Floorkontrolle ersetzt, die eine einheitliche Zugriffsstrategie für die globale Konferenz festlegt. Das OCTOPUS System verwendet in der ersten Ausbaustufe den Ansatz einer gemeinsamen Sprecherwarteschlange, wie sie im COVIS-System verwendet wurde [12]. Die globale Sprecherwarteschlange, die auf den Bildschirmen aller Konferenzteilnehmer zu

sehen ist, kann von jedem Teilnehmern der globalen Konferenz geändert werden. Sie gibt dem Teilnehmer das Rederecht, der an erster Stelle eingetragen ist. Als weiteres Schema wird das einer moderierten Konferenz eingeführt, bei dem ein Konferenzleiter (Moderator) entscheidet, wer als nächster spricht.

### 3.1.3. Gruppenkoordinator

Der Gruppenkoordinator des OCTOPUS-Ansatzes realisiert die externe Kommunikation eines lokalen Videokonferenzsystems. Er muss in jedes lokale Videokonferenzsystem, das an einer Videokonferenz nach dem OCTOPUS-Ansatz teilnehmen möchte, eingefügt werden. Die Gruppenkoordinatoren stellen eine einheitliche Schnittstelle für die Kommunikation auf der globalen Ebene bereit. Damit werden die verschiedenen Architekturen der lokalen Systeme verdeckt. Die konkrete Struktur der Gruppenkoordinatoren hängt von dem jeweiligen Videokonferenzsystem ab. Beispielsweise kann in einem zentralisiert verwalteten Videokonferenzsystem wie dem HKUST-System der Gruppenserver die Aufgaben des Gruppenkoordinators mit übernehmen. In COVIS, das eine dezentrale Verwaltung verwendet, muss eine neue Komponente in das System eingeführt werden. In diesem Fall zentralisiert der Gruppenkoordinator einige Funktionen, die gewöhnlich in dezentraler Weise von dem System ausgeführt werden. Weiterhin muss der Gruppenkoordinator im Falle der Nutzung unterschiedlicher Video- und Audiostandards in den lokalen Videokonferenzsystemen eine Konvertierungsfunktion enthalten, um sie auf einen global einheitlichen Standard zu überführen.

Aufgaben
----------

Die wichtigsten Aufgaben, die ein Gruppenkoordinator zu erfüllen hat, sind:

- Aufbau der globalen Konferenz,
- Integration eines lokalen Videokonferenzsystems in eine bereits bestehende globale Konferenz,
- Management des QoS auf globaler Ebene,
- Information der anderen Gruppenkoordinatoren über das Eintreten und Verlassen lokaler Teilnehmer in die Konferenz,
- Aussenden und Empfang der Audio- und Videoströme,

### 3. Gestaltung geschlossener Videokonferenzsysteme über das Internet

- Festlegung der Anzeige von Videoströmen (siehe Kapitel 5).

Architektur

Die Architektur des Gruppenkoordinators als Schnittstelle zwischen der lokalen und der globalen Ebene entspricht dem Schema aus Kapitel 2.3. Es existiert ein Signalisierungsteil, der den Medienübertragungsteil steuert. In Abbildung 3.3 ist diese Architektur gezeigt.

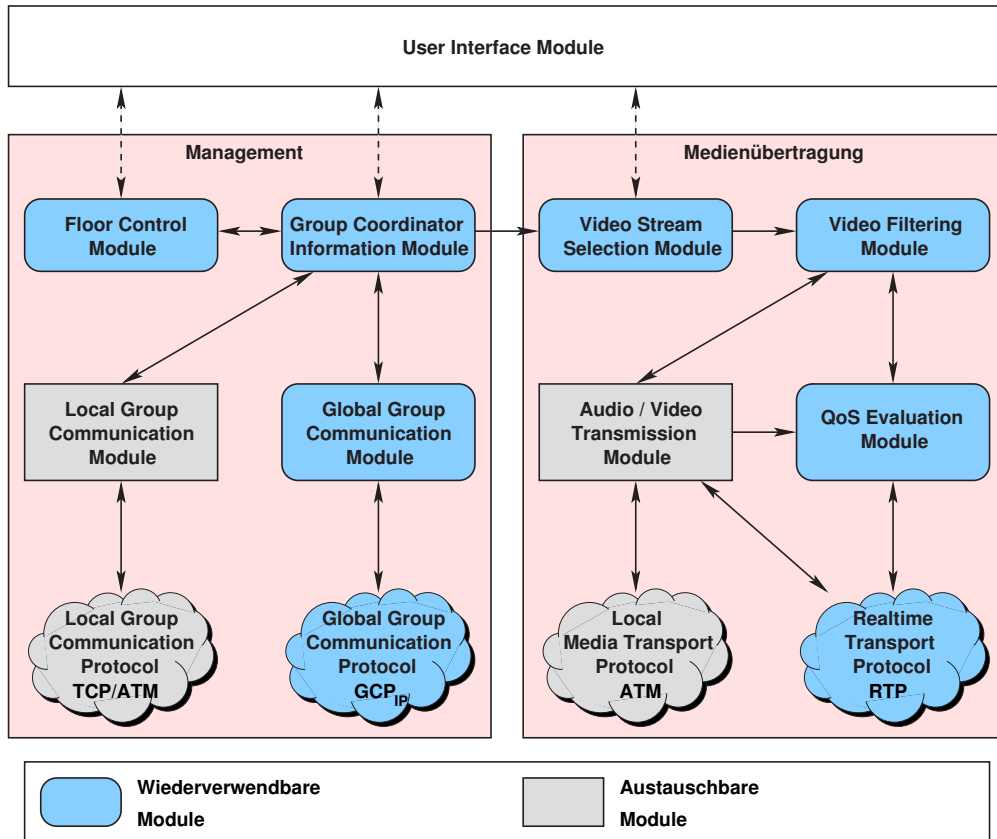


Abbildung 3.3.: Architektur des Gruppenkoordinators

Das Gruppenmanagement wird durch das *Group Coordinator Information Module* in Verbindung mit dem *Global Group Communication Module* und dem *Local Group Communication Module* realisiert. Das *Group Coordinator Information Module* bildet das Herzstück des Gruppenkoordinators. In ihm werden alle Informationen über die lokalen Teilnehmer und die entfernten Gruppen verwaltet. Er überwacht das Beitreten und Verlassen der Konferenz durch lokale Teilnehmer sowie entfernter Gruppen und informiert die anderen Gruppenkoordinatoren über Veränderungen. Das *Group Coordinator Information*

*Module* ist weiterhin für die Realisierung des QoS-Managements zuständig, das in Kapitel 5 vorgestellt wird. Die Kommunikation mit den lokalen Teilnehmern erfolgt durch das *Local Group Communication Module* entsprechend dem im lokalen Videokonferenzsystem gültigen Gruppenkommunikationsprotokoll. Das *Global Group Communication Module* kommuniziert mit den entfernten Gruppenkoordinatoren über einen globalen Gruppenkommunikationsdienst, dessen Gestaltung in Kapitel 4 diskutiert wird.

Das *Floor Control Module* ist für die Aktualisierung der gemeinsamen Sprecherwarteschlange auf den Bildschirmen der Konferenzteilnehmer verantwortlich. Lokale und entfernte Manipulationen der Warteschlange werden dem *Group Coordinator Information Module* mitgeteilt beziehungsweise von ihm angezeigt.

Die Audio- und Videoübertragung zwischen der lokalen Konferenz und den entfernten Konferenzen führt das *Audio/Video Transmission Module* aus. Es werden die Datenströme von dem *Local Transport Protocol* auf das *Realtime Transport Protocol (RTP)*[137] und umgekehrt abgebildet. Falls ein Konverter zwischen verschiedenen Audio- und Videoformaten erforderlich ist, wird dieser ebenfalls in das *Audio/Video Transmission Module* integriert. Das *Video Stream Selection Module*, das *QoS Evaluation Module* und das *Video Filtering Module* unterstützen das Bandbreitenmanagement, das im Abschnitt 5.4.2 erläutert wird.

Die Nutzerschnittstelle wird durch das *User Interface Module* bereitgestellt. Es stellt eine graphische Nutzerschnittstelle für Kontroll- und Managementzwecke bereit. Es zeigt aktuelle Informationen zum Zustand der globalen und der jeweiligen lokalen Videokonferenz an. Der Status der Floorkontrolle kann ebenfalls abgefragt werden.

Der Vorteil dieser Struktur liegt in der Wiederverwendbarkeit der Mehrzahl der Module, siehe Abbildung 3.3. So kann durch einfaches Austauschen der lokalen Kommunikationsschnittstellen ein anderes lokales Konferenzsystem ebenfalls zu dem Verbundsystem kompatibel gemacht werden.

#### 3.1.4. Kommunikationsszenarien

Das Abhalten einer globalen Konferenz in der geschilderten Art und Weise erfordert eine Reihe von Kommunikationsszenarien zwischen den Gruppenko-

ordinatoren und den lokalen Videokonferenzsystemen. Diese Kommunikationsszenarien umfassen unter anderem den Aufbau einer globalen Konferenz, das Zuschalten eines lokalen Videokonferenzsystems zur globalen Konferenz, den Beitritt bzw. das Ausscheiden lokaler Teilnehmer, den Ausfall von Teilnehmerrechnern oder des gesamten lokalen Videokonferenzsystems, das Ausscheiden einer lokalen Konferenz und das Beenden der globalen Konferenz.

#### **Aufbau der Konferenz**

Jedes lokale Videokonferenzsystem unterscheidet zwei Betriebsarten: den lokalen Modus und den globalen Modus. Im lokalen Modus läuft das lokale Videokonferenzsystem separat und ist nicht an die globale Konferenz angeschlossen. Der Gruppenkoordinator ist ausgeschaltet. Der lokale Modus ist der Ausgangsmodus jeder lokalen Konferenz. Um eine globale Konferenz aufzubauen, muss das lokale Videokonferenzsystem in den globalen Modus schalten, indem der Gruppenkoordinator in alle Aktivitäten der lokalen Gruppe einbezogen wird. Ein lokales System schaltet in den globalen Modus, wenn einer der Teilnehmer den Gruppenkoordinator explizit startet. Der Teilnehmer übergibt dabei die aktuelle Teilnehmerliste sowie die lokale Sprecherwarteschlange. Nach dem Beitritt des Gruppenkoordinators in die lokale Gruppe schalten die anderen Teilnehmer der Gruppe ebenfalls in den globalen Modus. Der Gruppenkoordinator nimmt nun aktiv an der lokalen Konferenz teil.

Um eine globale Videokonferenz aufzusetzen, müssen sich zumindest zwei lokale Videokonferenzsysteme im globalen Modus befinden. Ein Teilnehmer aus einer lokalen Gruppe ist der Initiator der globalen Konferenz, indem er eine *AppendGroup*-Nachricht an den lokalen Gruppenkoordinator sendet (siehe Abbildung 3.4 Schritt (1)). Aufgrund dieser Nachricht wird der entfernte Gruppenkoordinator auf der globalen Ebene eingeladen (Schritt (2) und (3)). Im Anschluss daran werden die Teilnehmerlisten und die Sprecherwarteschlangen ausgetauscht (Schritt (4)). Der einladende Gruppenkoordinator setzt die empfangene Sprecherwarteschlange an den Anfang der eigenen Warteschlange. Der eingeladene Gruppenkoordinator hängt die empfangene Sprecherwarteschlange entsprechend an das Ende seiner Liste, um so die Konsistenz der Warteschlange zu wahren. Die aktualisierte Teilnehmerliste und Sprecherwarteschlange wird an die lokalen Teilnehmer verschickt (Schritt (5)). Die globale Konferenz ist damit aufgebaut.

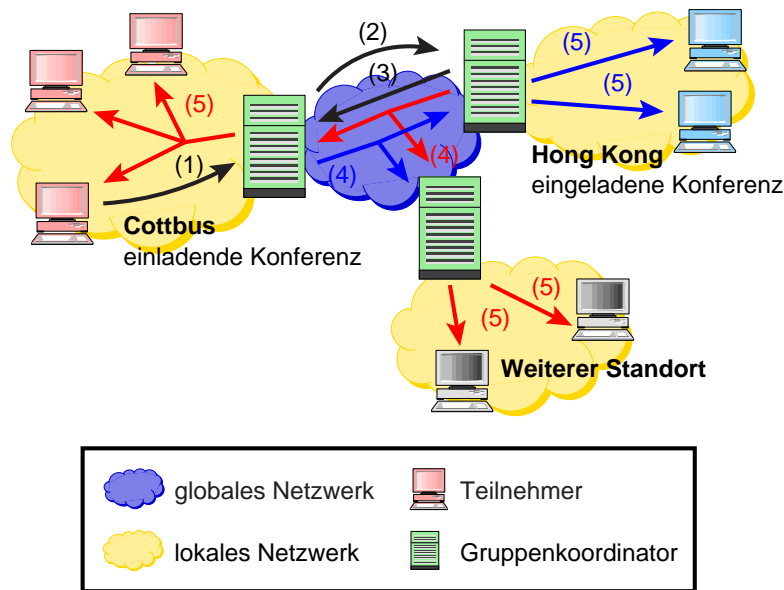


Abbildung 3.4.: Globaler Konferenzaufbau

### Beitritt zu einer bestehenden Konferenz

Eine ähnliche Prozedur wird ausgeführt, wenn ein lokales Videokonferenzsystem einer globalen Konferenz beitreten möchte. Ein Teilnehmer aus einer lokalen Konferenz sendet eine *AppendGroup*-Nachricht an den lokalen Gruppenkoordinator, der dann den Gruppenkoordinator der einzuladenden Gruppe kontaktiert (siehe Abbildung 3.4 Schritte (1) und (2)). Wird die Einladung angenommen (Schritt (3)), tauschen die Gruppenkoordinatoren die Teilnehmerlisten und die aktuellen Sprecherwarteschlangen aus (Schritt (4)). Die Warteschlangen werden wie beim Aufbau der Konferenz in den Gruppenkoordinatoren zusammengefügt. Die Gruppenkoordinatoren verteilen dann die aktualisierte Teilnehmerliste und die Sprecherwarteschlange an die lokalen Teilnehmer (Schritt (5)).

### Beitritt eines neuen Teilnehmers

Die globale Konferenz kann zu jedem Zeitpunkt durch den Beitritt eines neuen Teilnehmers vergrößert werden. Die Prozedur des Beitritts ist vom jeweiligen lokalen Videokonferenzsystem abhängig. Im COVIS kann jeder Teilnehmer einen neuen Teilnehmer einladen, der diese Einladung bestätigen muss (Schritt (1) bis (3)). Während dieser Prozedur werden dem neuen Teilnehmer

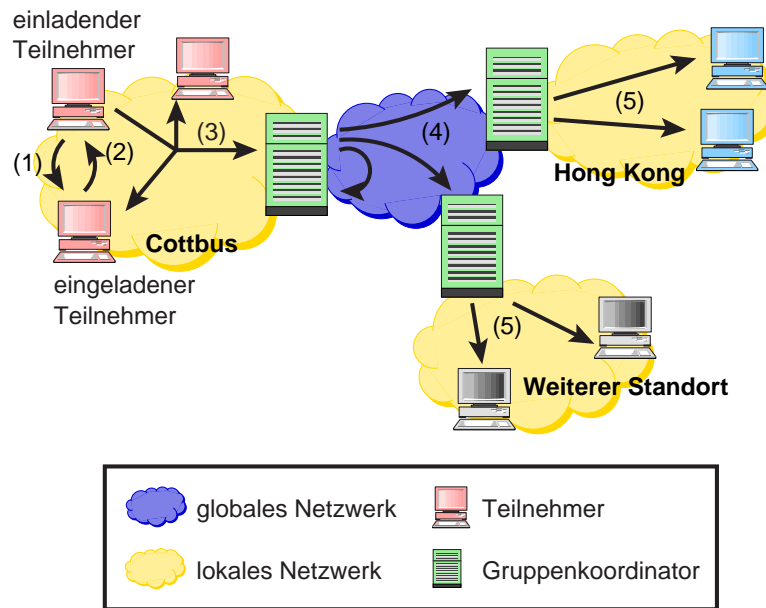


Abbildung 3.5.: Beitritt eines Teilnehmers

die aktuelle Teilnehmerliste und die aktuelle Sprecherwarteschlange mitgeteilt (Schritt (4)). Der Gruppenkoordinator wird in diesem Prozess als ein normaler Teilnehmer der Gruppe behandelt. Nachdem der Gruppenkoordinator über den Beitritt eines neuen Teilnehmers in seiner Gruppe informiert wird, teilt er dies den entfernten Gruppen über deren Gruppenkoordinatoren mit (Schritt (5) und (6)). Die entfernten Teilnehmer registrieren den neuen Teilnehmer in ihrer Teilnehmerliste.

Das Hongkonger System nutzt ein ähnliches Schema zum Einladen eines neuen Teilnehmers, wobei die Gruppenmanagementfunktionen vom Gruppenkoordinator ausgeführt werden.

In diesem Zusammenhang ist jedoch die Wahrung der geschlossenen Gruppe noch etwas problematisch. Zur Zeit wird jede lokale Gruppe in eine globale Konferenz aufgenommen, die sich im globalen Modus befindet und eingeladen wird. Einen Mechanismus, der zu einer Ablehnung einer Einladung führt, ist vorgesehen, muss aber explizit im Gruppenkoordinator und in der lokalen Konferenz integriert werden. Hierbei gibt es im Wesentlichen zwei Möglichkeiten. Die erste Möglichkeit besteht darin, dass dem Gruppenkoordinator mitgeteilt wird, von wem eine Einladung erwartet wird oder welche Einladungen



abgelehnt werden sollen. Dies ist jedoch ein sehr statisches Prinzip, das sich eigentlich kaum für die Praxis eignet. Bei der zweiten Möglichkeit wird die Einladung an alle lokalen Teilnehmer weitergeleitet. Diese müssen dann entscheiden, ob sie der laufenden globalen Konferenz beitreten möchten oder nicht. Dieser Ansatz ist für den praktischen Einsatz günstiger. Er erfordert jedoch eine Erweiterung der lokalen Videokonferenzsysteme um ein Modul, mit dem die globale Einladung verarbeitet wird.

### Änderung der Sprecherwarteschlange

Die Erhaltung der Konsistenz der Sprecherwarteschlange bei ihrer Änderung ist eines der zu lösenden Probleme in einem hierarchischen Ansatz wie dem OCTOPUS. Während einer Konferenz kann sich jeder Teilnehmer zu jeder

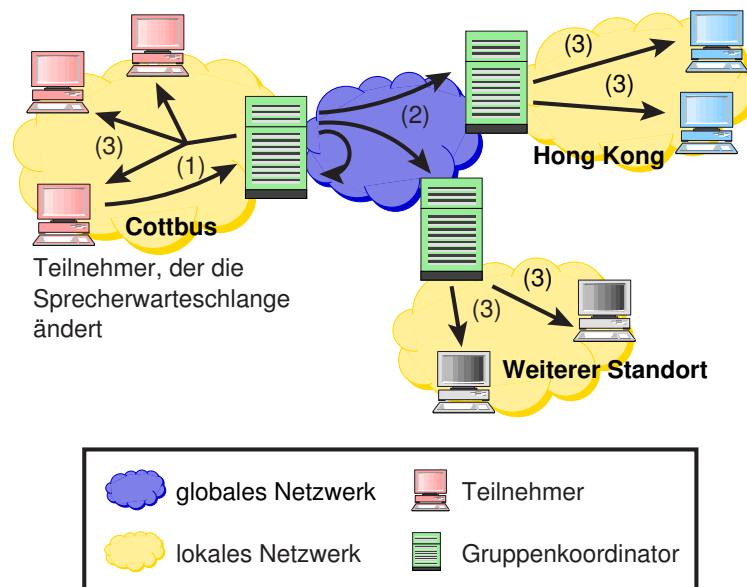


Abbildung 3.6.: Änderung der Sprecherwarteschlange

Zeit in die Sprecherwarteschlange ein- und austragen. Die Anfrage an den Floor wird zunächst an den Gruppenkoordinator geschickt (siehe Abbildung 3.6 Schritt (1)). Um die Konsistenz der Sprecherwarteschlange zu sichern, muss der Gruppenkoordinator zunächst die Anfrage an die globale Gruppe verschicken (Schritt (2)). Erst wenn sicher ist, dass die Anfrage alle anderen Gruppenkoordinatoren empfangen haben, darf sie an die lokalen Teilnehmer geschickt werden, um die lokalen Sprecherwarteschlangen zu aktualisieren (Schritt (3)).

### Austritt eines Teilnehmers

Ein Teilnehmer einer lokalen Gruppe kann zu jeder Zeit aus der Konferenz austreten. Dazu benachrichtigt er die lokale Gruppe und deren Gruppenkoordinator. Der Gruppenkoordinator informiert die anderen Gruppenkoordinatoren über den Austritt des Teilnehmers, die dann ihre eigenen lokalen Teilnehmer informieren. Daraufhin wird der Teilnehmer aus der Teilnehmerliste und gegebenenfalls aus der Sprecherwarteschlange entfernt.

### Austritt einer lokalen Gruppe

Das Verlassen der Konferenz durch eine Gruppe wird durch einen Teilnehmer der lokalen Konferenz eingeleitet. Er schickt eine *Disconnect*-Nachricht

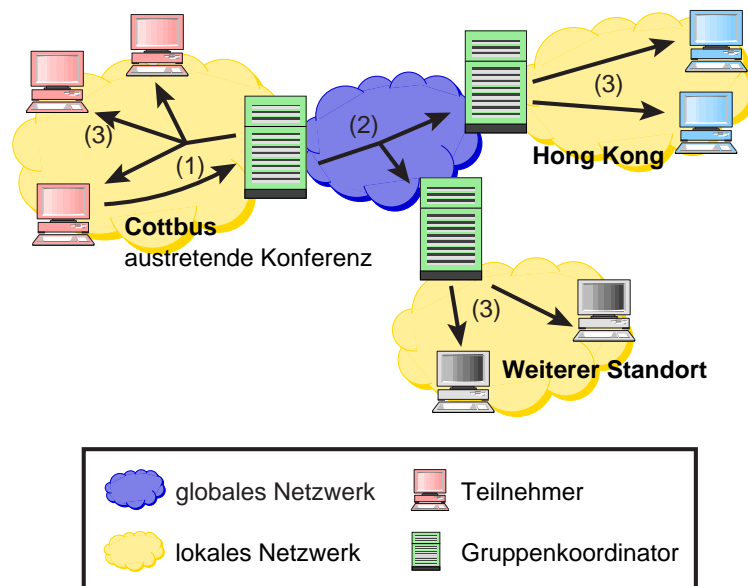


Abbildung 3.7.: Verlassen einer Gruppe

an den Gruppenkoordinator, der daraufhin die Gruppe der Gruppenkoordinatoren verlässt. Jeder Gruppenkoordinator informiert die lokalen Teilnehmer über die ausgetretene Gruppe, wodurch alle Teilnehmer aus dieser Gruppe aus der Teilnehmerliste und der Sprecherwarteschlange entfernt werden können. Der Gruppenkoordinator der ausgetretenen Gruppe bestätigt das Verlassen mit einer *DisconnectSuccess*-Nachricht an die Gruppe. Daraufhin können die lokalen Teilnehmer die Teilnehmer der entfernten Gruppen aus Teilnehmerliste und Sprecherwarteschlange entfernen.

### Ausfall einer lokalen Gruppe

Der Ausfall einer Gruppe stellt ein ernsthaftes Problem dar, weil es den logischen Ring unterbricht, auf dem sich die Gruppenkoordinatoren befinden. Es wird durch den Ausschluss der fehlerhaften Gruppe aus der Konferenz gelöst. Die verbleibenden Gruppenkoordinatoren werden darüber informiert und schließen den logischen Ring wieder. Die ausgefallenen Teilnehmer werden aus der Teilnehmerliste sowie aus der Sprecherwarteschlange entfernt.

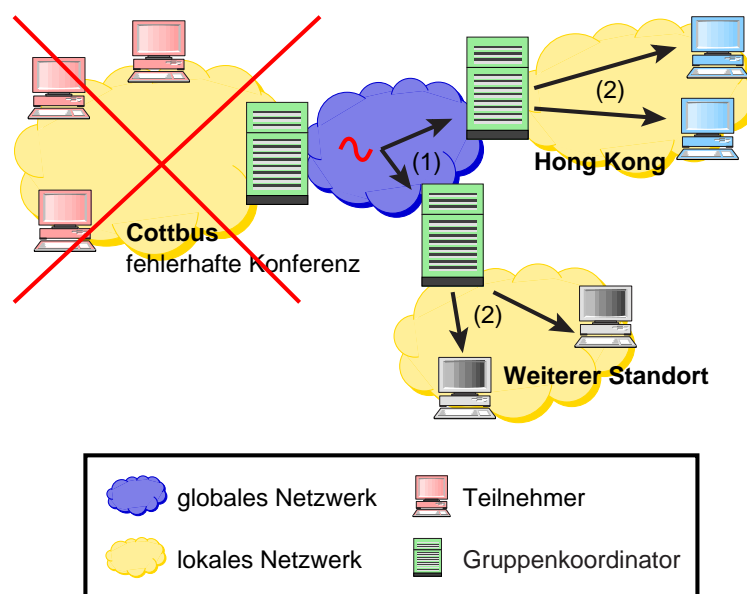


Abbildung 3.8.: Ausfall einer Gruppe

Dieser Ausschluss der fehlerhaften Konferenzgruppe klingt sehr hart. Jedoch wird durch das Signalisierungsprotokoll mehrfach versucht, einen Übertragungsfehler zu beheben. Ein schwerwiegender Fehler, wie beispielsweise der Ausfall des Gruppenkoordinators oder der Netzwerkverbindungen zwischen den Gruppenkoordinatoren, kann jedoch nicht durch das Protokoll behoben werden. Um den verbleibenden Teilnehmern trotzdem eine stabile Kommunikation zu ermöglichen, wird die Fehlerstelle aus der Konferenz entfernt. Ein späteres erneutes Einladen der ausgefallenen Gruppe ist möglich, um die Konferenz fortzusetzen.

### 3.1.5. Bewertung des OCTOPUS–Ansatzes

In diesem Abschnitt wurden die Grundzüge des Entwurfs für das Videokonferenzsystems OCTOPUS vorgestellt, das in einem gemeinsamen Projekt der Hong Kong University of Science and Technology und der BTU Cottbus entwickelt wurde. Mit OCTOPUS ist es nunmehr möglich, verschiedene lokale Videokonferenzsysteme mit unterschiedlichen Architekturen miteinander zu verbinden, um spontan globale, geschlossene Videokonferenzen aufzusetzen. Weitere Kernelemente des OCTOPUS–Ansatzes sind das dynamische Gruppenmanagement, die integrierte Floorkontrolle sowie ein Prioritäten–basiertes Bandbreitenmanagement. Es wurde ein OCTOPUS–Gruppenkoordinator als Prototyp für das Cottbuser Videokonferenzsystem COVIS unter SUN Solaris implementiert und getestet. Erweiterungen, die für den globalen Betrieb des COVIS notwendig sind, wurden in die COVIS–Endsysteme integriert. Durch den Aufbau von Konferenzen mit der Hong Kong University of Science and Technology und durch Simulationen wurde das Verhalten des Verbundsystems getestet.

## 3.2. Der BRAVIS–Ansatz

In diesem Abschnitt wird das Videokonferenzsystem BRAVIS (BRAndenburg Video conferencing Service) vorgestellt, das am Lehrstuhl Rechnernetze und Kommunikationssysteme der BTU Cottbus entwickelt wurde [123]. Im Rahmen dieser Arbeit wurden dafür Konzepte entworfen, Alternativen zu bisherigen Lösungen gesucht und gestalterische Aspekte untersucht, um geschlossene Videokonferenzsysteme über das Internet zu realisieren. Während OCTOPUS das Problem der Geschlossenheit von Konferenzgruppen durch den Verbund verschiedener lokaler Videokonferenzsysteme untersucht hat, war das Ziel der Entwicklung von BRAVIS die Entwicklung eines eigenständigen Systemansatzes, der auf einem verteilten Managementprinzip beruht und geschlossene Gruppen in IP–Netzen unterstützt. Dabei wird von typischen Gruppengrößen von 2 bis 10 Teilnehmern ausgegangen, wobei als Obergrenze für diese Entwicklung 25 Teilnehmer angenommen werden. BRAVIS wurde im Rahmen eines Landesprojekts entwickelt und soll Teleteaching–Anwendungen wie Teleseminare, Telekonsultationen, Prüfungen und Spezial–Vorlesungen unterstützen. Er ist jedoch auch für weitere Einsatzfälle nutzbar. BRAVIS basiert auf den

zuvor entwickelten Videokonferenzsystemen COVIS und OCTOPUS. Im Folgenden werden die Grundzüge des BRAVIS-Ansatzes vorgestellt und wichtige Entwurfsentscheidungen diskutiert.

### 3.2.1. BRAVIS-Funktionalität

BRAVIS bietet einen Desktop-Videokonferenzdienst für geschlossene Gruppen in IP-Netzen, der von Workstation und PCs gestartet werden kann und auf einem verteilten Gruppen- und QoS-Management beruht. Konferenzen können mit allen Teilnehmern geschaltet werden, die über die BRAVIS-Software verfügen. Die Floorkontrolle wie auch die Dienstgüteeinstellungen werden dezentral in den Endsystemen vorgenommen. BRAVIS nutzt ebenfalls das für COVIS entwickelte Prinzip der dynamischen Skalierung der Videoströme, das die Empfangsrate der Videoströme an die Belastung der Endsysteme automatisch anpasst. Dadurch wird der Dekompressionsaufwand der Videoströme bei den Empfängern reduziert und eine bessere Darstellungsqualität erreicht. Die dafür notwendigen Berechnungen werden nach dem in [13] beschriebenen Prinzip ausgeführt. Des Weiteren werden aus Nutzersicht die nachfolgend beschriebenen Funktionen bereitgestellt.

*Dienstbeschreibung*

BRAVIS unterstützt ein dynamisches Beitreten und Verlassen der Konferenz. Eine Konferenz kann von einem beliebigen Teilnehmer spontan initiiert werden. BRAVIS sieht eine explizite Einladung neuer Teilnehmer durch einen der Konferenzteilnehmer vor. Dies ist eine wesentliche Voraussetzung für die Gewährleistung der Geschlossenheit der Gruppe. Die Aufforderung zur Einladung basiert auf der Abstimmung zwischen den aktuellen Konferenzteilnehmern (soziales Protokoll). Die Konferenz kann von einem Teilnehmer zu jedem Zeitpunkt verlassen werden.

*Dynamisches Gruppenmanagement*

In BRAVIS werden alle Teilnehmer der Konferenz auf den Bildschirmen der Teilnehmer dargestellt. Dabei kann das in COVIS eingeführte Prinzip der Sprecherpriorisierung eingestellt werden, das dem aktuellen Sprecher eine bessere Dienstgüte (Framerate, Bildgröße) zuweist als den Zuhörern. Der Sprecher wird in diesem Fall auch bevorzugt in den Skalierungsprozeduren behandelt. Da die Darstellung aller Teilnehmer häufig nicht notwendig bzw. gewünscht ist sowie auch durch die beschränkte Bildschirmgröße begrenzt ist, wird in der nächsten Ausbaustufe eine modifizierte Variante des für OCTOPUS entwickel-

*Differenzierte Gruppendarstellung*

### 3. Gestaltung geschlossener Videokonferenzsysteme über das Internet

---

ten Abstimmungsprinzips ergänzt [27], [184], das eine Auswahl der darzustellenden Teilnehmer erlaubt, die zugleich für die Bandbreitenzuweisung ausgenutzt werden kann.

#### Geregelte Floor-Vergabe

BRAVIS unterstützt gegenwärtig zwei Floor-Vergabestrategien: die FIFO-Zuweisung über eine gemeinsame Sprecherwarteschlange und die Moderation. Beide Modi können beim Konferenzstart und auch während der Konferenz eingestellt werden. Die Sprecherwarteschlange ist auf allen Bildschirmen sichtbar und kann von den Teilnehmern unabhängig manipuliert werden (Eintragen und Austragen von Sprechwünschen). Der Erste in der Warteschlange besitzt das Rederecht und zugleich den exklusiven Zugriff auf das Whiteboard. Die Moderation ist basierend auf dem Warteschlangenkonzept realisiert, indem einem Teilnehmer, dem Moderator, exklusiv das Recht zur Manipulation der Warteschlange und damit zur Vergabe des Rederechts gegeben wird. Das Umschalten in den moderierten Modus während einer laufenden Konferenz muss durch eine Abstimmung der Teilnehmer erfolgen.

#### Whiteboard

BRAVIS enthält eine Whiteboard-Komponente. Die Nutzung des Multimedia Lecture Board (mlb) der Universität Mannheim [165] wurde angestrebt. Die Integration erwies sich jedoch aufgrund der unterschiedlichen verwendeten Signalisierungsprotokolle und Floorkontroll-Strategien im ersten Schritt als zu komplex. Deshalb wird vorerst eine einfache Whiteboard-Komponente integriert, die neben der graphischen Ein- und Ausgabe eine textuelle Schnittstelle zum Nutzer für kleine Textnachrichten hat, ähnlich einem Internet-Chat. Die Textnachrichten können sowohl an einen als auch an mehrere Teilnehmer gleichzeitig gesendet werden. Eine Regelung des Zugriffs auf den Chat über die Floorkontrolle gibt es nicht, um kleine Diskussionen per Text unter den Zuhörern zuzulassen.

#### Graphische Nutzerschnittstelle

Für den Zugang zu BRAVIS steht eine speziell gestaltete graphische Nutzerschnittstelle zur Verfügung, die dem Nutzer neben der Darstellung der Konferenzteilnehmer eine einfache und einheitliche Bedienung der Steuerelemente für die Einladung, die Anforderung des Rederechts und den Zugriff auf das Whiteboard bietet.

### 3.2.2. Topologie

BRAVIS soll Teilnehmer an verschiedenen Standorten im Internet miteinander verbinden. Obwohl primär eine regionale Anwendung vorgesehen ist, existieren genau genommen keine (regionalen) Obergrenzen. Der BRAVIS-Ansatz nimmt keine einheitliche Netztopologie an, wie sie zum Beispiel in einem lokalen Netz gegeben ist. Er geht davon aus, dass sich die Teilnehmer in verschiedenen Netzen befinden, die über eine unterschiedliche Dienstgüte verfügen.

BRAVIS nutzt vom Grundansatz her die gleiche Topologie wie OCTOPUS (siehe Abbildung 3.9). Es verzichtet jedoch auf die Integration eines speziellen, relativ aufwendig zu programmierenden Gateways. Die Vermeidung eines

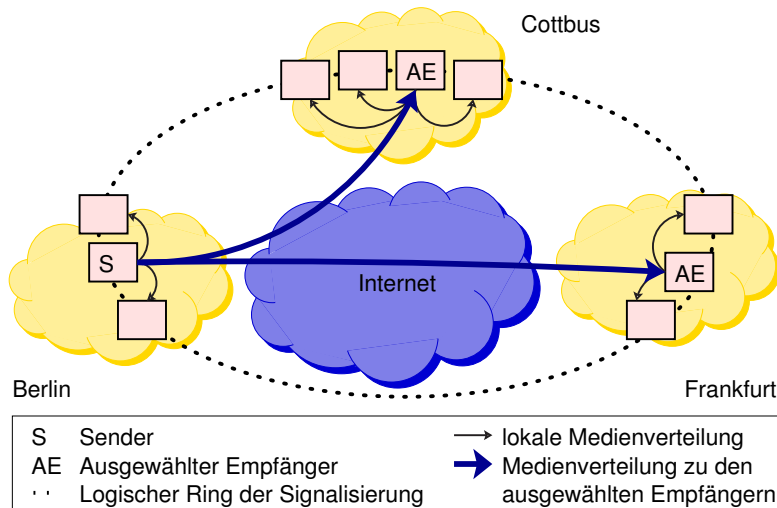


Abbildung 3.9.: Die BRAVIS Topologie

expliziten Gateways wird möglich, weil BRAVIS im Gegensatz zum OCTOPUS-Ansatz davon ausgeht, dass alle Teilnehmer über die gleiche Videokonferenzschnittstelle verfügen. Des Weiteren sind die Endsysteme in den letzten Jahren um ein Vielfaches leistungsfähiger geworden, so dass sie die Aufgaben des OCTOPUS-Gruppenkoordinators bewältigen können. BRAVIS lehnt sich hierbei an die in den letzten Jahren verstärkt entwickelten Peer-to-Peer-Konzepte an, bei denen jeder Konferenzteilnehmer sowohl Sender als auch Empfänger ist. Da IP-Multicast nur einen empfängerbasierten Multicastdienst zur Verfügung stellt und somit nur offene Gruppen unterstützen kann, müssen im

BRAVIS-Ansatz die Signalisierungs- sowie die Audio- und Videodaten über Unicastverbindungen übertragen werden.

#### Signalisierung

BRAVIS nutzt einen logischen Ring für das Konferenzmanagement. Die Kommunikationstopologie wird jedoch maßgeblich durch das Signalisierungsprotokoll bestimmt, das in Kapitel 4 eingehend diskutiert wird.

#### Audio- und Videodaten

Unabhängig von Overlay-Multicast-Ansätzen [8] wurde in BRAVIS ein zweistufiges Konzept für die Verteilung der Audio- und Videodaten entwickelt, bei dem alle Teilnehmer in Untergruppen aufgeteilt werden. In jeder Untergruppe existiert ein ausgewählter Empfänger, der die Multimediaströme der entfernten Untergruppen dupliziert und an die anderen Teilnehmer der lokalen Gruppe weiterleitet (siehe Abbildung 3.9). Die Kommunikation zwischen dem ausgewählten Empfänger und den anderen lokalen Gruppenmitgliedern erfolgt ebenfalls Punkt-zu-Punkt, um eine geschlossene Gruppenkommunikation zu ermöglichen. Verfügt die lokale Gruppe über einen senderbezogenen Multicastdienst, kann dieser genutzt werden. Als ausgewählter Empfänger kann jeder Teilnehmer der lokalen Gruppe fungieren. Die Funktion wird von dem Konferenzteilnehmer der lokalen Gruppe übernommen, der die meisten Ressourcen zur Verfügung stellen kann. Die Aushandlung, welcher Teilnehmer der Leistungsstärkste ist, wird durch das QoS-Management der Teilnehmersysteme durchgeführt und dynamisch angepasst. Dabei können zwei Sonderfälle eintreten: (1) alle Empfänger befinden sich in der lokalen Gruppe oder (2) alle Empfänger gehören unterschiedlichen Untergruppen (bzw. Netzen) an. Die Lösung dieses Problems wird in Abschnitt 5.4.4 diskutiert.

Trotz der veränderten Topologie bleibt das in COVIS und OCTOPUS [12], [27] verwendete Prinzip der dynamischen Skalierung der Videoströme erhalten. Jeder Teilnehmer sendet mit der maximalen Bandbreite an eine entfernte Gruppe, die von dieser akzeptiert werden kann. Die maximale Bandbreite bestimmt sich aus der Bandbreite, die aktuell vom Sender zum Teilnetz vorhanden ist, und aus der Bandbreite, die das leistungsstärkste Teilnehmersystem in diesem Netz verarbeiten kann. Leistungsschwächere Teilnehmer müssen die ankommenden Videoströme entsprechend ihrer Leistungsfähigkeit skalieren, bevor sie diese Daten weiter verarbeiten. Dafür werden Frame-Dropping-Filter verwendet, die einige der noch komprimierten Videoframes verwerfen.



### 3.2.3. Architektur

Aus der in den vergangenen Abschnitten beschriebenen Funktionalität von BRAVIS sowie den topologischen und organisatorischen Randbedingungen leitet sich die in Abbildung 3.10 dargestellte Architektur des Videokonferenzdienstes ab, die bei jedem Teilnehmer vorhanden sein muss. Die Architek-

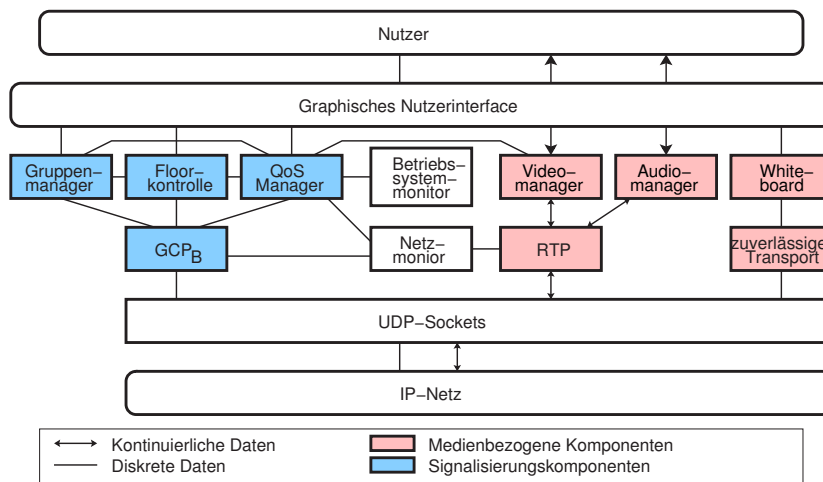


Abbildung 3.10.: BRAVIS-Architektur

tur gliedert sich, wie bei multimedialen Anwendungen üblich, in einen Datenübertragungs- und Signalisierungsteil, die jeweils auf den Protokollen RTP und GCP aufsetzen.

Den Kern des Signalisierungsteils bilden der Gruppenmanager und der QoS-Manager. Sie realisieren zusammen mit dem Floorkontroll-Modul die Funktionen, die im Abschnitt 3.2.1 beschrieben wurden. Die Konsistenz der Managementdaten wird durch das Gruppenkommunikationsprotokoll GCP sichergestellt, das in Kapitel 4 detailliert diskutiert wird.

Das Senden und Empfangen der Audio-, Video- und Datenströme realisiert der Datenübertragungsteil. Die Verteilung der Audio- und Videodaten wird auf Basis der in Abschnitt 3.2.2 beschriebenen Topologie durchgeführt. Als Transportprotokoll für die Audio- und Videodaten wird RTP [137] genutzt.

Die Daten für das Whiteboard werden durch ein einfaches zuverlässiges Transportprotokoll, das eigens dafür entwickelt wurde, übertragen. Dieses Protokoll nutzt ebenfalls Unicast-Verbindungen, um den geschlossenen Charakter

*Konferenz-  
management*

*Medien-  
bearbeitung*

der Konferenz zu bewahren. Im Zusammenhang mit der in das BRAVIS-System integrierten Floorkontrolle kann die Übertragungsreihenfolge der Whiteboard-Daten sichergestellt werden, da dann nur der aktuelle Sprecher den Zugriff auf das Whiteboard besitzt. Somit ist keine Synchronisation im Whiteboard oder dem Transportprotokoll notwendig.

Die Möglichkeiten der QoS-Regulierung des Videokonferenzdienstes sind beschränkt, da der Dienst letztlich nur eine Anwendung ist, die nicht Einfluss auf die Netzsituation nehmen kann. BRAVIS geht davon aus, dass sich die Teilnehmer in getrennten Netzen befinden, die unterschiedliche QoS-Modelle nutzen. Deshalb beschränkt sich BRAVIS vorerst auf die Überwachung der Systemlast in den Endsystemen und der Netzsituation. Dafür werden zwei Monitore genutzt. Der Betriebssystem-Monitor beobachtet die aktuelle Prozessorlast und den Ressourcenverbrauch. Diese Beobachtungen werden genutzt, um die Filterparameter für die dynamische Skalierung der Systemlast anzupassen [13]. Die Überwachung der Netzlast dient der dynamischen Anpassung der kontinuierlichen Datenströme an die aktuelle Netzwerksituation. Die Resultate aus beiden Beobachtungen fließen außerdem in die Festlegung des ausgewählten Empfängers für eine Untergruppe ein. Ein anderer Gesichtspunkt des QoS-Managements ist die Dienstgütesicherung durch Nutzung von QoS-Modellen wie IntServ und DiffServ, die auf ihre dynamische Konfigurierbarkeit durch die eben genannten Monitore hin untersucht wurden. In Kapitel 5 wird auf dieses Problem näher eingegangen.

#### 3.2.4. Bewertung des BRAVIS-Ansatzes

BRAVIS bietet erstmals einen Videokonferenzdienst, der geschlossene Gruppen im Internet ohne einen Videokonferenzserver oder eine MCU unterstützt. Dabei stellt der Ansatz nur sehr geringe Anforderungen an das Netzwerk, da weder Multicast- noch QoS-sichernde Dienste für den Betrieb notwendig sind. Durch diese geringen Anforderungen ist BRAVIS ideal für spontane Beratungen in Firmen und im Privatleben geeignet. Mit der Nutzung der IP-Technologie ist dieser Ansatz zudem sehr preiswert. Die grundlegenden Elemente des Ansatzes sind die Unterteilung der Konferenzgruppe in lokale Untergruppen, die Punkt-zu-Punkt-Übertragung zu ausgewählten Empfängern einer Untergruppe, die Entwicklung eines geschlossenen Gruppenmodells durch GCP und das verteilte

Gruppen- und QoS-Management. Die beschriebene BRAVIS-Architektur wurde für SOLARIS- und LINUX-Umgebungen implementiert und im Campus der BTU Cottbus an ausgewählten Standorten, die bis zu 1,5 km entfernt sind, installiert und getestet.

### 3.3. Zusammenfassung

Dieses Kapitel hat einen Überblick über die Ansätze OCTOPUS und BRAVIS gegeben, die geschlossene Videokonferenzen über das Internet ermöglichen. Beide Konzepte versprechen dabei eine Reihe von Vorteilen und Merkmalen, die mit bisherigen Systemlösungen nicht erreicht werden können.

Die Wahrung der Geschlossenheit der Gruppe stellt in globalen öffentlichen Netzen wie dem Internet ein Problem dar, weil senderbasierte Multicastdienste im Internet nicht verbreitet sind. Somit können keine effektive Verteilmechanismen des Netzwerks für eine Gruppenkommunikation genutzt werden. Durch neuartige Ansätze für die Signalisierung und die Medienverteilung in OCTOPUS und BRAVIS wird sowohl eine effektive Verteilung der Konferenzdaten als auch die Geschlossenheit der Gruppe erreicht. Diese Ansätze werden im Kapitel 4 beziehungsweise im Kapitel 5 vorgestellt und diskutiert.

*geschlossene  
Gruppen*

OCTOPUS enthält ein neuartiges QoS Management, das heterogene Netzwerkumgebungen in der Weise unterstützt, dass die Medienströme zwischen den entfernten Gruppen individuell an die vorhandenen Ressourcen angepasst werden. Somit erhalten Gruppen mit einer guten Netzwerkanbindung eine höhere Qualität als solche mit einer schlechten Anbindung. Auch in BRAVIS wird dieses Konzept genutzt, da die Topologien beider Systemansätze für die Übertragung der Video- und Audiodaten sehr ähnlich sind. Heterogene Endsysteme werden ebenfalls durch OCTOPUS und BRAVIS unterstützt.

*QoS Management*

Mittels OCTOPUS können spontane Konferenzen aufgesetzt werden, wenn sich die lokale Konferenz im globalen Modus befindet. In einem dezentral organisierten lokalen System wie COVIS muss dazu der Gruppenkoordinator in die laufende Konferenz eingebunden werden. Erst dann ist die Gruppe bereit, eine entfernte Gruppe zu kontaktieren beziehungsweise eine Einladung einer entfernten Gruppe anzunehmen. Diese Beschränkung der Spontaneität kann aber durch ein automatisches Integrieren des Gruppenkoordinators in eine laufen-

*Spontaneität*

### 3. Gestaltung geschlossener Videokonferenzsysteme über das Internet

---

de Sitzung aufgehoben werden. In zentralisierten Lösungen, wie dem HKUST, wird die Software für den Gruppenkoordinator direkt auf dem Gruppenserver gestartet. Dadurch können Systeme mit einem zentralisierten Ansatz spontan eingeladen werden, da der Gruppenkoordinator nicht explizit der Gruppe beitreten muss.

Das spontane Aufsetzen einer Videokonferenz wird mit BRAVIS erleichtert, da nur die Endsysteme ohne zusätzliche Zwischensysteme miteinander kommunizieren. Es ist keine Konfiguration eines Konferenzservers, Gruppenkoordinators oder Netzwerks notwendig, um eine Konferenz zu starten.

globale Konferenzen

Durch den OCTOPUS-Ansatz ist es erstmals möglich, lokale Videokonferenzsysteme für den Aufbau von globalen Konferenzen zu nutzen. Der Vorteil liegt darin, dass OCTOPUS keine Bedingungen an das lokale System stellt, und somit quasi jedes System über OCTOPUS kommunizieren kann. Somit können Investitionen für ein bereits gekauftes lokales Videokonferenzsystem erhalten werden, und die Nutzer müssen nicht auf ein neues System umgeschult werden. OCTOPUS stellt dazu Schnittstellen zur Verfügung, die auch für andere gängige lokale Systeme, wie H.323 oder ISABEL genutzt werden können.

BRAVIS stellt einen eigenständigen und unabhängigen Ansatz für den Aufbau von globalen Konferenzen im Internet bereit. Im Gegensatz zu bisherigen Systemen berücksichtigt BRAVIS die Besonderheiten des Internets, wie beispielsweise die Heterogenität des Netzes und der Endsysteme sowie die schwankenden zur Verfügung stehenden Ressourcen des Netzwerks.

Robustheit

Der hierarchische Ansatz von OCTOPUS bietet außerdem eine höhere Ausfallsicherheit als zentralisierte Systeme, wie beispielsweise H.323. Auch gegenüber einer hierarchischen Anordnung von H.323 MCUs, wie sie im DFN-Videokonferenzdienstes (DFNVC) angeboten wird, ist OCTOPUS robuster gegen Fehlerfälle. Bei einem Fehlerfall einer MCU, die sich in der Hierarchie recht weit oben befindet, wird die gesamte Konferenz gestört, wobei jede Teilgruppe in den unteren Hierarchieebenen betroffen ist. Dies liegt darin begründet, dass der gesamte Audio- und Videoverkehr einer Konferenzschaltung über diese MCU verteilt wird. Fällt ein Gruppenkoordinator im OCTOPUS-System aus, so wird nur die lokale Gruppe dieses Gruppenkoordinators aus der Konferenz entfernt. Die verbleibenden Gruppen können ungestört weiter kommunizieren.

Der Ausfall eines Teilnehmers einer lokalen Gruppe hat keine Folgen für die verbleibenden Teilnehmer.

BRAVIS ist durch den dezentralen Ansatz sehr robust gegenüber Fehlerfällen. Tritt ein Fehler bei einem der Teilnehmer auf, kann die Kommunikation unter Ausschluss dieses Teilnehmers fortgesetzt werden.

Zur Zeit werden in OCTOPUS und BRAVIS die Daten nicht verschlüsselt übertragen. Durch die Geschlossenheit der Gruppe ist jedoch eine Vertraulichkeit erreichbar, die mit der des Telefondienstes vergleichbar ist. Trotzdem kann durch dedizierte Angriffe Dritter die Kommunikation mit speziellen Software-Werkzeugen belauscht werden. Besonders bei Broadcastmedien wie dem Ethernet- oder den Funk-LANs in den lokalen Bereichen besteht diese Gefahr. Deshalb wird gegenwärtig in einer weiteren Arbeit ein Sicherheitskonzept für den BRAVIS-Ansatz entwickelt, wodurch auch streng vertrauliche Beratungen mit einer verschlüsselten Datenübertragung möglich werden. Dieses Konzept sieht verschiedene Sicherheitsstufen vor, die von den Teilnehmern in Abhängigkeit von der angestrebten Vertraulichkeit der Beratung vereinbart werden können. Die Sicherheitsarchitektur soll in BRAVIS integriert werden und kann gegebenenfalls auch für OCTOPUS übernommen werden.

*Vertrauliche Beratungen*

Firewalls stellen oft auch ein Problem für Videokonferenzdienste dar, da sowohl MBone-basierte Systeme als auch H.323-Systeme Multicastadressen beziehungsweise UDP-Ports dynamisch für die Übertragung der Medienströme zuweisen. Systeme, die einen breiten Bereich von offenen Portnummern benötigen, gelten als Sicherheitsrisiko für das jeweilige Netzwerk. Ansätze für eine dynamische Konfiguration von Firewalls durch die Anwender eines Videokonferenzsystems sind auf eine ähnliche Weise sicherheitskritisch, da dadurch quasi jedem Benutzer eine Schnittstelle zur Konfiguration der Firewall bereitsteht. Mit OCTOPUS und BRAVIS lassen sich diese Probleme weitgehend vermeiden, da nur eine sehr geringe Anzahl von freien Ports zur Verfügung stehen muss. Ein fester Port wird für die Signalisierung über das Gruppenkommunikationsprotokoll benötigt. Jeweils ein weiterer Port muss für die Video- und Audiokommunikation über RTP geöffnet werden. Da auf eine Rückmeldung über einen RTCP-Kanal verzichtet werden kann, ist die Reduzierung auf einen festen geöffneten Port pro Medium möglich. Ein weiterer Vorteil im OCTOPUS besteht darin, dass nur der Host, auf dem der Gruppenkoordinator installiert

*Firewall*

### *3. Gestaltung geschlossener Videokonferenzsysteme über das Internet*

---

ist, von der Firewall freigegeben werden muss. Somit ist das verbleibende Netzwerk weitgehend geschützt.

---

## 4. Sicherung der Geschlossenheit der Gruppe in globalen Videokonferenzsystemen

---

Kern eines Videokonferenzsystems für geschlossene Gruppen ist die Sicherung der Geschlossenheit der Konferenzgruppe. Geschlossenheit bedeutet dabei, dass die ausgetauschten Audio-, Video- und Datenströme nur die Partner erreichen, die aktuell an der Konferenz teilnehmen. Geschlossene Videokonferenzen erfordern deshalb ein senderbezogenes Sendeparadigma, das die Empfänger der Daten kennt. Für den Beitritt zur Konferenz ist somit eine explizite Einladung oder Anmeldung notwendig.

Wie im Kapitel 2 bereits diskutiert, existieren zur Zeit keine zufrieden stellenden Lösungen für einen senderbezogenen Multicastdienst im Internet. Zentralistische Ansätze wie die H.32x-Systeme lösen das Problem der Geschlossenheit der Konferenzgruppe durch eine explizite Anmeldung beim Gruppenserver, dem Gatekeeper, bei dem sich alle Konferenzteilnehmer an- und abmelden müssen.

Videokonferenzlösungen, die auf Mbone aufbauen, können geschlossene Gruppen nur unter Verwendung zusätzlicher Sicherheitsmechanismen wie der Verschlüsselung der Teilnehmer-Videoströme sichern, weil das zugrunde liegende Übertragungsprinzip des IP-Multicast keine geschlossenen Gruppen unterstützt. Jeder Teilnehmer, der die Multicast-Adresse kennt, kann sich (unbemerkt von allen anderen Teilnehmern) zu der Konferenz zuschalten.

In einem verteilten Ansatz, wie ihn OCTOPUS<sup>1</sup> und BRAVIS verfolgen, sind zusätzliche Maßnahmen erforderlich, um die Geschlossenheit der Gruppe zu gewährleisten. Es ist außerdem erforderlich, dass für alle Teilnehmer stets ei-

---

<sup>1</sup>OCTOPUS ist an sich ein hierarchisches System. Die Kommunikation der Gruppenkoordinatoren in der globalen Ebene basiert jedoch auf einem verteilten Ansatz.

ne konsistente Sicht auf die Konferenz, insbesondere ihre Managementdaten gegeben ist.

Dem allgemeinen Architekturbild eines Videokonferenzsystems in Abbildung 2.6 folgend, wird in diesem Kapitel das Konferenzmanagement in Bezug auf die Wahrung der Geschlossenheit einer Gruppe untersucht. Diese Diskussion bildet einen Schwerpunkt der im Rahmen dieser Arbeit entwickelten Mechanismen.

### 4.1. Anforderungen an einen dezentralen Managementansatz

OCTOPUS und BRAVIS basieren auf einem verteilten Managementansatz, bei dem jedes Endsystem die Funktionen der Gruppen- und QoS-Verwaltung eigenständig ausführt. Somit muss der Zustand der Konferenz in jedem teilnehmenden Endsystem identisch sein, um einheitliche Managemententscheidungen treffen zu können. Zustandsänderungen bei einem Teilnehmer müssen allen beteiligten Endsystemen mitgeteilt werden, um die Konsistenz der Konferenz zu wahren. Das Management der Systeme OCTOPUS und BRAVIS erfordert dafür ein geeignetes Signalisierungsprotokoll, das den Austausch und die Konsistenz der Managementdaten sichert. In [11] wurden die Anforderungen spezifiziert, die ein Dienst erfüllen muss, um die Konsistenz der Konferenzdaten zu sichern. Diese sind:

**Zuverlässigkeit:** Verfälschungen und Verluste von Signalisierungsnachrichten müssen erkannt und durch Neuübertragung behoben werden. Damit wird gewährleistet, dass die Teilnehmer stets eine korrekte Managemententscheidung treffen können.

**Ordnung:** Die Nachrichten müssen in der Reihenfolge ausgeliefert werden, in der sie gesendet wurden. Dies sichert, dass Ereignisse wie das Beitreten zur Konferenz, das Verlassen der Konferenz oder die Änderung der Floorkontrolle allen Teilnehmern in der Reihenfolge angezeigt werden, in der sie aufgetreten sind. Es können kausale Ordnung oder totale Ordnung unterschieden werden. Kausale Ordnung bedeutet, dass die Nachrichten in der Reihenfolge ausgeliefert werden, wie sie von *einem* Sender erzeugt



wurden. Dies ist bei der Verteilung von QoS-spezifischen Daten ausreichend. Als Erweiterung dazu steht die totale Ordnung, bei der zusätzlich die Ereignisse *aller* Sender in eine einheitliche Reihenfolge gebracht werden. Dies ist bei der Floorkontrolle zwingend notwendig.

**Atomarität:** Die Nachrichten müssen an alle Gruppenmitglieder ausgeliefert werden oder an keines. Damit wird gewährleistet, dass die Managementdaten aller Teilnehmer in der gleichen Weise aktualisiert werden und somit die Konsistenz der Sicht auf die Konferenz gewahrt wird. Wenn eine Nachricht nicht an alle Teilnehmer gesendet werden konnte, wird diese Übertragung als „nicht stattgefunden“ betrachtet und muss gegebenenfalls wiederholt werden.

Am Beispiel einer verteilt organisierten Floorkontrolle mittels einer Sprecherwarteschlange wird die Notwendigkeit der Erfüllung dieser harten Anforderungen deutlich. Wird eine Anforderung an die Floorkontrolle nicht atomar an alle Teilnehmer übertragen, kann auf jedem Endsystem eine andere Sprecherwarteschlange mit unterschiedlicher Länge und Einträgen existieren. Somit ist es nicht möglich, den Sprecher eindeutig zu identifizieren. Die Atomarität sichert, dass auf allen beteiligten Endsystemen, die gleichen Einträge in der Sprecherwarteschlange enthalten sind. Die eindeutige Reihenfolge der Einträge in der Warteschlange auf allen beteiligten Hosts wird durch die totale Ordnung der Flooranforderungen gesichert. Sonst könnten mehrere Teilnehmer gleichzeitig davon ausgehen, dass sie der aktuelle Sprecher sind. Die Zuverlässigkeit sichert hierbei, dass nicht beispielsweise durch Bitverfälschungen eine Flooranforderung in eine Floorabgabe umgewandelt wird.

## 4.2. $GCP_{ATM}$

Für das Videokonferenzsystem COVIS wurde mit dem Gruppenkommunikationsprotokoll  $GCP_{ATM}$  [11] ein Protokoll entworfen, das die geforderten Kriterien nach Zuverlässigkeit, Atomarität und Ordnung erfüllt.  $GCP_{ATM}$  basiert auf der Nutzung eines senderbasierten Multicastdienstes, der durch ATM (Asynchronous Transfer Mode) unterstützt wird. In der Entwurfsphase des  $GCP_{ATM}$  wurden verschiedene Protokolle auf ihre Eignung als Signalisierungsprotokoll

mit den genannten Eigenschaften für einen Videokonferenzdienst untersucht. Dazu gehören unter Anderem RAMP, SRM, TMTP, RMTP, RMP, Totem und MTP.

**RAMP** Die Zuverlässigkeit des *Reliable Adaptive Multicast Protocol (RAMP)* [85] basiert auf negativen Bestätigungen und auf extrem geringen Fehlerraten (kleiner  $10^{-12}$ ). RAMP kann des Weiteren keine totale Ordnung und Atomarität für eine Gruppenkommunikation gewährleisten, da es nur eine 1:n-Kommunikation unterstützt.

**SRM** Das *Scalable Reliable Multicast (SRM)* [43] wurde für das Whiteboard der Mbone-Tools [88] entwickelt. Bestätigungen der Datenpakete und die Neuübertragung nach einem Paketverlust werden an die gesamte Gruppe gesendet, wodurch ein hoher Overhead entsteht. Es sichert keine Ordnung und keine Atomarität.

**SMP** Die Protokollprinzipien vom *Scalable Multicast Protocol (SMP)* [51] wurden von SRM abgeleitet. SMP findet Anwendung im *Digital Lecture Board* der Universität Mannheim [46]. Jedoch gewährleistet SMP keine totale Ordnung und auch keine Atomarität der Datenübertragung, die für die Sicherung der Konsistenz der Verwaltungsdaten eines verteilten Videokonferenzsystems wichtig ist.

**TMTP und RMTP** Das *Tree-based Multicast Transport Protocol (TMTP)* [178] und das *Reliable Multicast Transport Protocol (RMTP)* [91] basieren auf einem hierarchischen Ansatz, bei dem jeder Sender einen logischen Verteilbaum zu den Empfängern aufbaut. Die Datenpakete werden dem Knoten in der nächst höheren Hierarchiestufe bestätigt. Verloren gegangene Pakete werden gegebenenfalls von diesem wiederholt gesendet. Totale Ordnung und Atomarität werden von diesen Protokollen nicht gewährleistet.

**RMP** Das *Reliable Multicast Protocol (RMP)* [173] sichert Zuverlässigkeit, Atomarität und totale Ordnung in verschiedenen Graden. Dadurch ist es sehr komplex. Des Weiteren geht es von einem eher kontinuierlichen Datenfluss aus, was in einem Signalisierungsprotokoll nicht der Fall ist. RMP liegt dem IP-Multicast zugrunde, wodurch es nicht für geschlossene Gruppen geeignet ist.

**Totem** Das *Totem* Protokoll [2] basiert auf einem Tokenprinzip. Nur dem Tokenhalter ist es gestattet, Daten zu senden. Dies sichert totale Ordnung und Zu-

verlässigkeit der Datenübertragung. Totem ist jedoch sehr komplex und beschränkt auf ein lokales Netzwerk. Gateways, wie in [3] vorgeschlagen, ermöglichen den Einsatz in globalen Umgebungen, erhöhen aber gleichzeitig die Komplexität des Protokolls. Totem setzt außerdem einen eher kontinuierlichen Datenstrom voraus.

Das *Multicast Transport Protocol (MTP)* [7] beruht auf einem Client–Server–Prinzip. Die Clients fordern vom Server einen Token mit einer eindeutigen Sequenznummer an, um Daten senden zu dürfen. Dadurch wird eine totale Ordnung der Datenauslieferung möglich. Von Nachteil ist die zentrale Rolle des Servers. Fällt er aus oder ist er überlastet, kommt die Kommunikation zum Erliegen.

MTP

$GCP_{ATM}$  [11] wurde aufgrund der beobachteten Einschränkungen dieser Protokolle entwickelt. Es enthält einige Ideen der untersuchten Lösungen, wie beispielsweise den tokenbasierten Mechanismus ähnlich dem in [2]. Alle Teilnehmer bilden einen logischen Ring, auf dem der Token kreist. Der Besitzer des Tokens hat jeweils das Senderecht, wodurch die totale Ordnung der Datenauslieferung sichergestellt wird, siehe Abbildung 4.1.

 $GCP_{ATM}$ 

Das sehr einfache Übertragungsprinzip des  $GCP_{ATM}$  arbeitet wie folgt. Der  $GCP$ –Protokollinstanz können zu einem beliebigen Zeitpunkt Daten übergeben werden, die zwischengespeichert werden, bis die Instanz den Token bekommt. Nur der Tokenbesitzer darf PDUs senden. Alle Teilnehmer haben den Empfang der PDUs zu bestätigen. Unbestätigte PDUs werden wiederholt gesendet, um die Zuverlässigkeit in  $GCP_{ATM}$  zu erreichen. Gibt es nach einer gewissen Zeit noch ausstehende Bestätigungen, wird der Teilnehmer, dessen Bestätigung ausblieb, aus der Gruppe ausgeschlossen (er muss die Konferenz verlassen). Damit haben alle anderen Teilnehmer, die sich noch in der Konferenz befinden, die Nachricht erhalten, wodurch Atomarität gewährleistet wird. Der Token wird nach Erhalt aller Bestätigungen an den nächsten Teilnehmer übergeben. Der Empfang des Tokens muss ebenfalls bestätigt werden. Hat der Tokenhalter keine Nachrichten zu versenden, gibt er den Token weiter.

Die für  $GCP_{ATM}$  entwickelte Lösung setzt auf der darunter liegenden Ebene eine existierende senderbasierte Multicast–Struktur voraus. Das ist für ATM gegeben. In  $GCP_{ATM}$  ist jeder Teilnehmer über eine überlagerte Multicaststruktur mit den anderen Teilnehmern verbunden (siehe Abbildung 4.1).

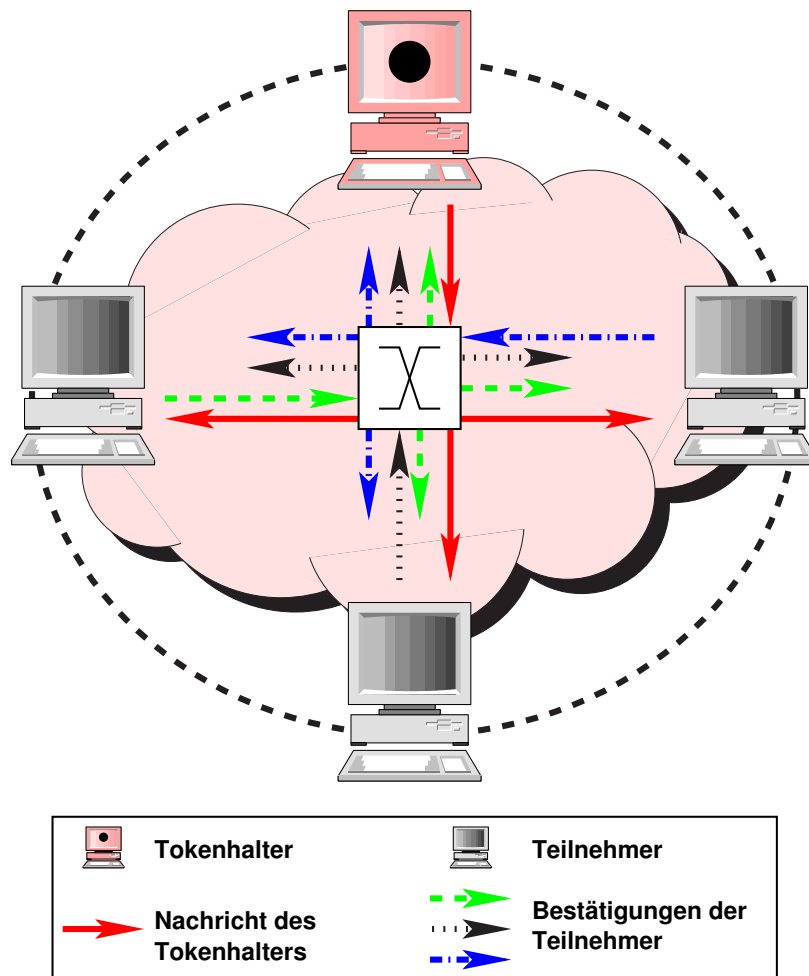


Abbildung 4.1.: Das Prinzip des GCP<sub>ATM</sub>

Da sich die ATM-Technologie, anders als lange prognostiziert, nicht durchgesetzt hat, bleibt der GCP<sub>ATM</sub>-Ansatz in seiner Anwendung beschränkt. Die IP-Technologie unterstützt Multicast, jedoch ist das verwendete Übertragungsprinzip empfängerbasiert, das heißt, Empfänger können jederzeit der Gruppe beitreten. Der Sender muss sie nicht unbedingt kennen. Damit werden also primär offene Gruppen unterstützt. Es stellt sich damit die Frage, wie die durch das GCP<sub>ATM</sub> bereitgestellten Eigenschaften Zuverlässigkeit, totale Ordnung und Atomarität innerhalb einer geschlossenen Gruppe auf IP übertragen werden können.

## Abbildung des GCP auf IP-Netzwerke

Mit dem *Explicit Multicast (XCast)* [16] wurde ein senderbasiertes Multicast über IP vorgeschlagen. Es befindet sich gerade in der Erprobung. Mit diesem Protokoll wäre es möglich, das Prinzip des GCP<sub>ATM</sub> direkt auf eine IP-Variante zu übernehmen. XCast ist jedoch, wie in Abschnitt 2.6 geschildert, noch nicht ausgereift. Durch die Nutzung einer definierten IP-Multicastgruppe, in der sich alle XCast-Router befinden müssen, ist mit XCast trotz des senderbasierten Ansatzes keine geschlossene Gruppenkommunikation möglich.

XCast

Das GCP<sub>ATM</sub>-Prinzip könnte weiterhin durch eine entsprechende IPv4 Option einer senderbasierten Multicastübertragung übernommen werden, wie es der RFC 1770 [48] vorschlägt. Es können jedoch nur maximal 9 Teilnehmer in einer Konferenz teilnehmen, da die Länge des Optionfeldes nur 40 Byte groß ist und 2 Byte für Kontrollinformationen genutzt werden. Des Weiteren ist diese Lösung ungenügend verbreitet, da kaum ein Router im Kernnetz des Internet in der Lage ist, diese IPv4-Option auszuwerten.

IPv4 Option

Die Abbildung des Multicastdienstes auf einen Unicastdienst wäre ein weiterer Ansatz, um das Prinzip des GCP<sub>ATM</sub> auf IP abzubilden. Von Vorteil ist, dass die Reaktionszeiten nicht proportional mit der Anzahl der Teilnehmer wächst. Aber ein Host muss  $n - 1$  Verbindungen verwalten. Hinzu kommt, dass sich um den Sender herum ein Bandbreitenengpass bildet, da er  $n - 1$  Pakete quasi gleichzeitig abschicken muss und ebenso  $n - 1$  Bestätigungen empfängt. Dies führt zu der so genannten *Quittungsimplosion*, die unerwünschte Paketverluste bewirken kann. Aus diesem Grund ist es im Allgemeinen vorteilhaft, keine direkte Unicast-Abbildung eines Multicastdienstes zu nutzen<sup>2</sup>.

Unicastabbildung

Die Forschung und Entwicklung von *Overlay-Multicast*-Ansätzen begannen erst etwa zwei Jahre nach der Entwicklung von OCTOPUS, da allgemein auf den Erfolg von IP-Multicast spekuliert wurde. Somit konnte für Überführung des GCP<sub>ATM</sub> nicht auf solche Techniken zurückgegriffen werden. Des Weiteren unterstützen die derzeit verfügbaren Overlay-Multicast-Ansätze kein senderbasiertes Multicast [9, 29, 176].

Overlay Multicast

Das Prinzip des GCP<sub>ATM</sub> kann somit nicht ohne Weiteres auf IP-Netzwerke übertragen werden, da die notwendige Voraussetzung eines senderbasierten

<sup>2</sup>Mit BRAVIS wurde trotzdem dieser Ansatz zu Beginn des Projekts umgesetzt. Die Hintergründe werden dafür in C erläutert.

Multicastdienstes fehlt. Die Nutzung des empfängerbasierten IP-Multicast ist aufgrund der offenen Gruppenverwaltung nicht für geschlossene Gruppenkommunikation geeignet. Die Anforderungen nach Zuverlässigkeit, totaler Ordnung und Atomarität müssen durch eine Unicaststruktur erfüllt werden. Dafür wurde ein anderes Realisierungskonzept des GCP –  $GCP_{IP}$  – entworfen, das im nächsten Abschnitt vorgestellt wird.

### 4.3. $GCP_{IP}$

Das Signalisierungsprotokoll *Gruppenkommunikationsprotokoll über IP* (*Group Communication Protocol over IP*,  $GCP_{IP}$ ) wurde im Rahmen des OCTOPUS-Projekts für die Signalisierung der Gruppenkoordinatoren entwickelt. Es kann auch für andere Anwendungen, die eine geschlossene, dezentral organisierte Gruppenkommunikation benötigen, eingesetzt werden. Oberhalb  $GCP_{IP}$  können zum Beispiel wie in COVIS das Gruppen- und QoS-Management sowie die Floorkontrolle aufgesetzt werden. Es ist aber auch denkbar, dass komplexere Gruppenkommunikationsprotokolle wie SCCP (*Simple Conference Control Protocol*) [20] oder SCCS (*Scalable Conferencing Control Service*) [162] auf  $GCP_{IP}$  aufgesetzt werden.

$GCP_{IP}$  wurde ursprünglich für den Einsatz über UDP/IP entworfen und in der ersten Realisierung prototypisch so implementiert. In einer zukünftigen Entwicklung kann  $GCP_{IP}$  auch direkt auf IP aufgesetzt werden (siehe Abbildung 4.2), wobei dann Multiplexfähigkeiten hinzugefügt werden sollten. Dadurch ergäbe sich der Vorteil, dass mehrere GCP-Instanzen gleichzeitig auf einem Host genutzt werden können, um mehrere, voneinander unabhängige Konferenzen zu unterstützen. Dies wird in der Zukunft eine möglicherweise

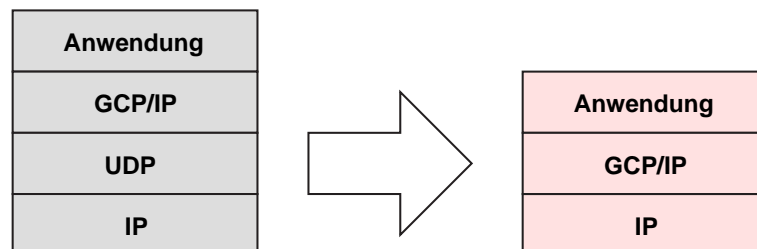


Abbildung 4.2.: Der  $GCP_{IP}$ -Protokollstack

wichtigere Rolle spielen, wenn Anwendungen nur noch von einem Terminal aus auf einem Applikationsserver gestartet werden. Den Vorteilen gegenüber steht eine Programmierung im Betriebssystemkern, da für IP in der Regel keine Programmierschnittstelle im Benutzeradressraum zur Verfügung gestellt wird.

### 4.3.1. Der $GCP_{IP}$ -Dienst

$GCP_{IP}$  stellt wie  $GCP_{ATM}$  einen zuverlässigen, atomaren und total ordnenden Gruppenkommunikationsdienst bereit. Es wurden lediglich kleine Änderungen an den Dienstprimitiven vorgenommen. Somit können bereits existierende Anwendungen, wie beispielsweise das QoS-Management oder die Floorkontrolle des COVIS, durch einfaches Ersetzen des Signalisierungsprotokolls von einem ATM-basierten System auf ein IP-basiertes System übertragen werden.

$GCP_{IP}$  unterscheidet drei Phasen: Aufbau und Beitritt zur Konferenz, Datenaustausch sowie Verlassen der Konferenz. Im Anschluss werden die drei Teildienste und die zugehörigen Dienstprimitive vorgestellt, die für das Verständnis der nachfolgenden Protokollbeschreibung notwendig sind. Des Weiteren wird auf die Unterschiede zum  $GCP_{ATM}$ -Dienst eingegangen.

#### Beitritt zur Konferenz (JOIN)

Der Teildienst JOIN umfasst alle Aktivitäten, die mit dem dynamischen Beitritt eines neuen Teilnehmers zu einer existierenden Gruppe verbunden sind. Dazu gehört auch der Aufbau einer Sitzung durch einen Teilnehmer, der den Ausgangspunkt für den weiteren Konferenzaufbau bildet. JOIN beinhaltet die folgenden Dienstprimitive:

##### **JOINrequest:** JOINreq(Host)

Einladung eines neuen Teilnehmers. Der Parameter (Host) gibt die IP-Adresse des Hosts des neuen Teilnehmers an. Methoden zum Finden eines Teilnehmers in einem globalen Netzwerk sind nicht Bestandteil von  $GCP_{IP}$ , sondern werden anderen Protokollen wie beispielsweise dem *Session Initiation Protocol (SIP)* [55] überlassen.

##### **JOINindication:** JOINind(User@Host)

Anzeige einer Einladung beim Eingeladenen. Der Anwendung über dem  $GCP_{IP}$  bleibt es überlassen, ob sie durch die zwei Primitive JOINresp

oder `JOINrej` aktiv auf die Nachricht reagiert. Als Parameter wird im Gegensatz zu `GCPATM` der Name des einladenden Teilnehmers und dessen Standort übergeben.

**JOINrespond:** `JOINresp`

Der neue Teilnehmer oder die Anwendung auf dem eingeladenen Hosts stimmt einer Einladung zu einer Konferenz zu.

**JOINreject:** `JOINrej`

Der neue Teilnehmer oder die Anwendung auf dem eingeladenen Hosts lehnt eine Einladung ab.

**JOINdisconnect:** `JOINdis`

Anzeige eines Fehlers während der Beitrittsphase bei dem Einladenden und dem Eingeladenen. Dies führt zum Abbruch des Beitritts. `JOINdis` wird dem Einladenden angezeigt, wenn der Eingeladene die Einladung abgelehnt hatte.

**JOINnotify:** `JOINntf(User@Host)`

Anzeige bei allen Gruppenteilnehmern, dass ein neuer Teilnehmer der Konferenz beigetreten ist. Zusätzlich zum `GCPATM` wird der Name des Teilnehmers und der benutzte Host als Parameter übergeben. Aus Sicht des Eingeladenen ist jeder Teilnehmer der Gruppe ein neuer Teilnehmer. Deshalb wird im Gegensatz zum `GCPATM` jeder Teilnehmer beim Eingeladenen angezeigt (siehe Abbildung 4.4).

In den Abbildungen 4.3 und 4.4 sind die Zeitablaufdiagramme der JOIN-Phase mit den dazugehörigen Dienstprimitiven dargestellt. Im Anhang B.1 werden mögliche Fehlerfälle und deren Behebung dargestellt.

### Datenaustausch (DAT)

Der Teildienst DAT dient der zuverlässigen, atomaren und total geordneten Übertragung der Management- und Steuerdaten. Abbildung 4.5 zeigt das Zeitablaufdiagramm der DAT-Phase mit den dazugehörigen Dienstprimitiven:

**DATrequest:** `DATreq(data)`

Übergabe der Daten, die übertragen werden sollen.



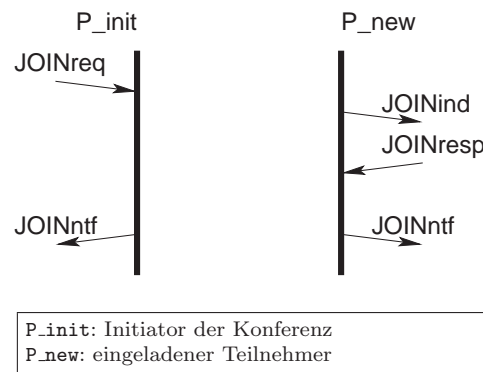


Abbildung 4.3.: Erfolgreiche Initialisierung einer Konferenz

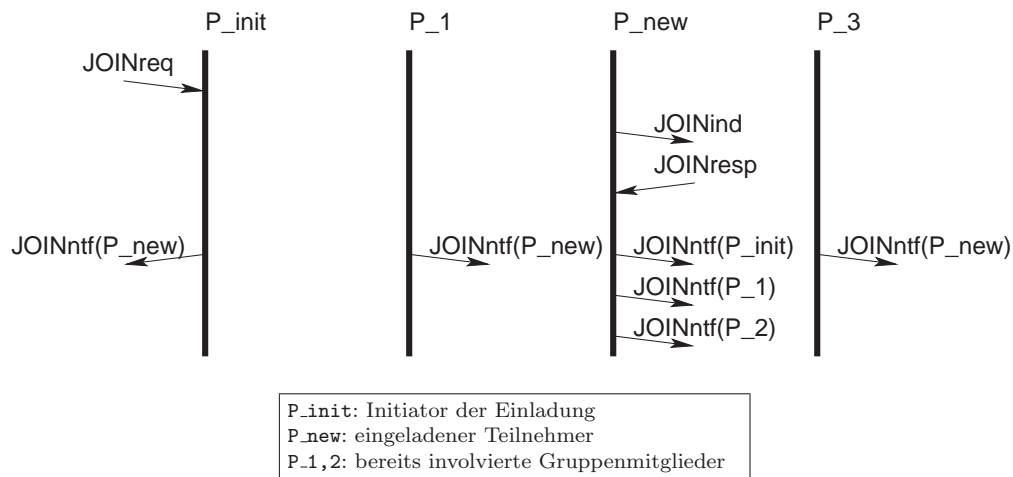


Abbildung 4.4.: Erfolgreicher Beitritt eines Teilnehmers zur Gruppe

**DATindication:** DATind(data)

Anzeige des Empfangs neuer Daten beim Teilnehmer. Anders als im  $GCP_{ATM}$  wird die erfolgreiche Datenübertragung auch beim Absender durch ein DATind angezeigt<sup>3</sup>.

**Verlassen der Konferenz (LEAVE)**

LEAVE realisiert das Verlassen der Konferenz durch einen Teilnehmer. Das Verlassen der Konferenz ist zu jedem Zeitpunkt möglich. LEAVE realisiert

<sup>3</sup>Ohne diesen Zusatz wird keine totale Ordnung erreicht, wenn zwei Teilnehmer quasi gleichzeitig den Dienst anfordern.

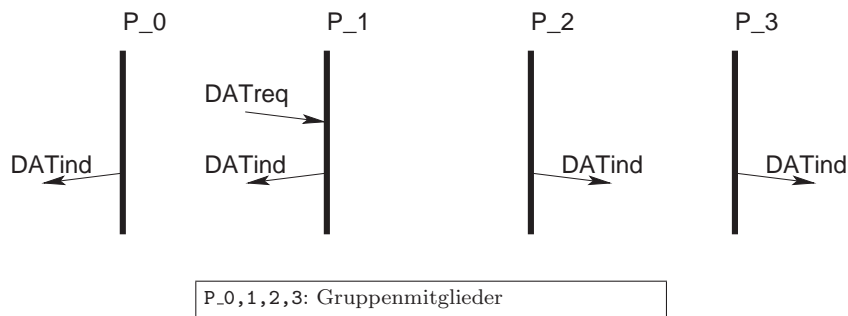


Abbildung 4.5.: Erfolgreiche Übertragung eines Datenpakets

ebenfalls das Beenden einer Konferenz, wenn ein Teilnehmer einer Zweiergruppe die Gruppe verlässt. LEAVE umfasst die Dienstprimitive, die in Abbildung 4.6 als Zeitablaufdiagramm dargestellt sind:

**LEAVErequest:** LEAVEreq

Ein Teilnehmer zeigt an, dass er die Konferenz verlassen möchte.

**LEAVEnotify:** LEAVEntf (User@Host)

Anzeige bei allen Teilnehmern, dass ein Teilnehmer die Konferenz verlassen hat. Zusätzlich zu  $GCP_{ATM}$  werden der Name und der Hostname des ausgetretenen Teilnehmers als Parameter des Dienstprimitivs übergeben. Wird LEAVEntf mit dem eigenen Namen oder dem Parameter **self** übergeben, so wird für diesen Teilnehmer die Konferenz beendet. Dadurch wird das Dienstprimitiv LEAVEconfirm des  $GCP_{ATM}$  eingespart.

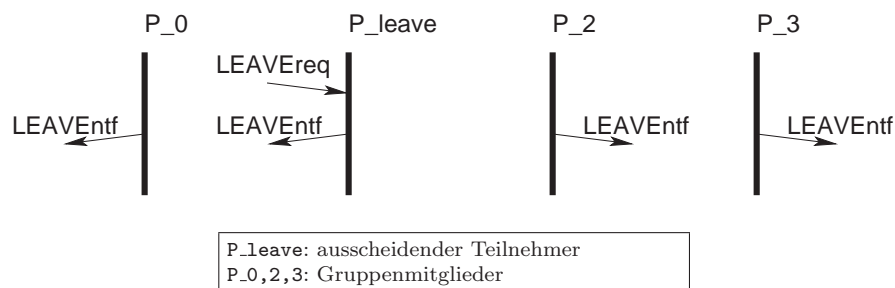


Abbildung 4.6.: Austritt eines Teilnehmers aus der Gruppe

### 4.3.2. Das GCP<sub>IP</sub>-Protokoll

Um einen zuverlässigen, atomaren und total geordneten Dienst mittels Punkt-zu-Punkt-Verbindungen zu realisieren, wurde im GCP<sub>IP</sub> das Prinzip des logischen Rings verwendet (siehe Abbildung 4.7). Anders als im GCP<sub>ATM</sub>, bei dem die Stationen über mehrere senderbasierte Multicast-Verbindungen kommunizieren, wird im GCP<sub>IP</sub> jede Station über Punkt-zu-Punkt Verbindungen mit den Nachbarsystemen verbunden. Auf diesem Ring kreist ein Token, das wie im GCP<sub>ATM</sub> das Senderecht vergibt. Die Daten werden ebenfalls über diesen Ring verteilt. Aus der Verwendung des logischen Rings ergeben sich Unterschiede für den Beitritt beziehungsweise Austritt der Konferenz sowie für den Datenaustausch, die später erläutert werden.

Die drei Eigenschaften des GCP-Dienstes werden in diesem Ansatz wie folgt gesichert:

Die totale Ordnung gewährleistet GCP<sub>ATM</sub> durch ein Tokenprinzip, das auf Punkt-zu-Punkt-Verbindungen beruht<sup>4</sup>. Somit ist dies auf GCP<sub>IP</sub> anwendbar. Der Token rotiert auf einem logischen Ring, den die Teilnehmer bilden. Nur dem Tokenhalter ist es erlaubt, auf Anfragen des Dienstnutzers wie JOINreq, DATreq oder LEAVEreq zu reagieren. Er darf Datenpakete verschicken, Einladungen an noch nicht involvierte Teilnehmer versenden und aus der Konferenz eigenständig austreten. Er darf außerdem die von ihm gesendeten Datenpakete wieder von dem Ring nehmen. Die verbleibenden Teilnehmer sind zum Datenempfang bereit. Auf keinen Fall dürfen sie eigene Pakete kreieren und verschicken oder Daten des Tokenhalters nicht weitergeben. Sie müssen Anfragen des Dienstnutzers zwischenspeichern, bis sie selbst zum Tokenhalter werden.

<i>totale Ordnung</i>
-----------------------

Die in [10] diskutierten Protokolle sichern die Zuverlässigkeit entweder durch positive oder negative Bestätigungen. GCP<sub>ATM</sub> verwendet positive Bestätigungen, die bei einer Gruppenkommunikation im Allgemeinen zu einer Quittungsimplosion bei dem Sender der Nachricht führen. Negative Bestätigungen werden geschickt, wenn ein Datenpaket nicht empfangen wurde. Dies wird hauptsächlich durch Überprüfung der Sequenznummern festgestellt. Dieses

<i>Zuverlässigkeit</i>
------------------------

<sup>4</sup>Die PDUs für die Tokenweitergabe im GCP<sub>ATM</sub> werden per Multicast übertragen. In den PDUs wird der Empfänger eingetragen. Teilnehmer, die nicht der eingetragene Empfänger sind, ignorieren die PDUs, wodurch dies einer Punkt-zu-Punkt-Verbindung gleicht.

#### 4. Sicherung der Geschlossenheit der Gruppe in globalen Videokonferenzsystemen

Verfahrens funktioniert aber nur unter der Voraussetzung, dass relativ kontinuierlich Daten gesendet werden müssen, um rechtzeitig fehlende Dateneinheiten festzustellen.

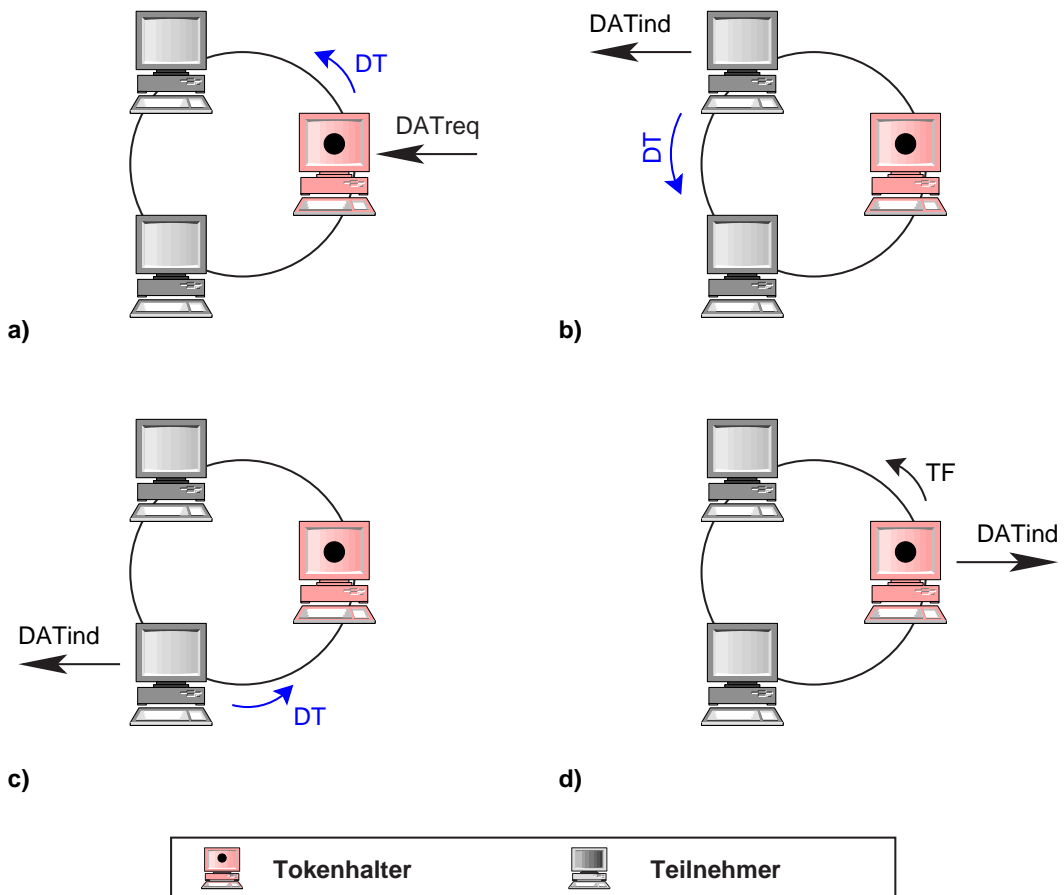


Abbildung 4.7.: Prinzip des Datenaustauschs im GCP<sub>IP</sub>

In einem Signalisierungsprotokoll wie dem GCP<sub>IP</sub> werden jedoch nur selten Daten gesendet, so dass Verluste erst sehr spät erkannt werden. Somit ist diese Methode der Fehlererkennung hierfür nicht geeignet. Um die Nachteile von positiven und negativen Bestätigungen zu umgehen, wurde die Verteilung der Daten über den logischen Ring konzipiert. Das Prinzip ist in Abbildung 4.7 dargestellt. Der Tokenhalter möchte das **DATreq** versenden und bildet das Datenpaket **DT** (a)). Das Datenpaket **DT** wird durch jeden Teilnehmer empfangen. Diese zeigen die **DATind**-PDU beim Dienstanutzer an und leiten danach das **DT** zum nächsten Teilnehmer weiter (b) und c)). Das Datenpaket **DT** wird

durch den Tokenhalter wieder empfangen, der ebenfalls eine  $DATind$ -SDU an den Dienstonutzer als Bestätigung für den erfolgreichen Datentransfer übergibt. Der Tokenhalter gibt danach das Token über die  $TF$ -PDU weiter (d)).

Jedes Datenpaket wird implizit dadurch bestätigt, dass es den Ring durchlaufen hat und wieder beim Tokenhalter angekommen ist. Durch dieses Prinzip reduziert sich der Aufwand für die Verwaltung von nur zwei Verbindungen pro Teilnehmer, anstelle von  $n - 1$  Verbindungen bei einer vollständigen Abbildung des Multicast auf Unicast.

Bei der Gestaltung des  $GCP_{ATM}$  wurde davon ausgegangen, dass bitweise Verfälschungen der Nutzdaten im ATM-Netzwerk nicht auftreten. In einem globalen IP-Netzwerk unter Nutzung von UDP kann diese Annahme nicht getroffen werden. Um Übertragungsfehler zu erkennen, wurde im  $GCP_{IP}$  ein CRC (*Cyclic Redundancy Check*) nach CCITT integriert. Ein fehlerhaftes Paket wird verworfen und muss erneut übertragen werden.

$GCP_{IP}$  realisiert die Atomarität wie  $GCP_{ATM}$  durch den so genannten *Forced Leave* Mechanismus, der nach einem dreimaligen missglückten Wiederholen einer Paketübertragung gestartet wird. Die Zahl 3 für die Anzahl der Wiederholungen begründet sich einerseits auf den im TCP vorgeschlagenen Wert von 3 [22] sowie auf Messungen der Fehlerrate auf verschiedenen globalen Internetverbindungen, die in Abbildung 4.16 gezeigt werden. Fehlerhafte Gruppenteilnehmer werden aus der Konferenz ausgeschlossen und müssen erneut eingeladen werden. Das Prinzip wird am Ende dieses Abschnitts beschrieben.

Atomarität

Die Phasen des  $GCP_{IP}$ -Dienstes, Aufbau und Beitritt zur Konferenz, Datenaustausch und Verlassen der Konferenz spiegeln sich auch im Protokoll wider. Hinzu kommt die Tokenweitergabe.

### Initialisierung der Konferenz

Eine Konferenz wird dadurch initialisiert, dass ein Teilnehmer, der Initiator, einen anderen Teilnehmer einlädt. Er übergibt dem  $GCP_{IP}$  ein  $JOINreq$ , aus dem eine  $JR$ -PDU gebildet und zu dem einzuladenden Teilnehmer geschickt wird (siehe Abbildung 4.8a). Der eingeladene Teilnehmer muss diese Einladung bestätigen. Das Ausbleiben der Bestätigung, weil beispielsweise der Teilnehmer nicht erreichbar ist, führt zum Abbruch der Einladung mit der

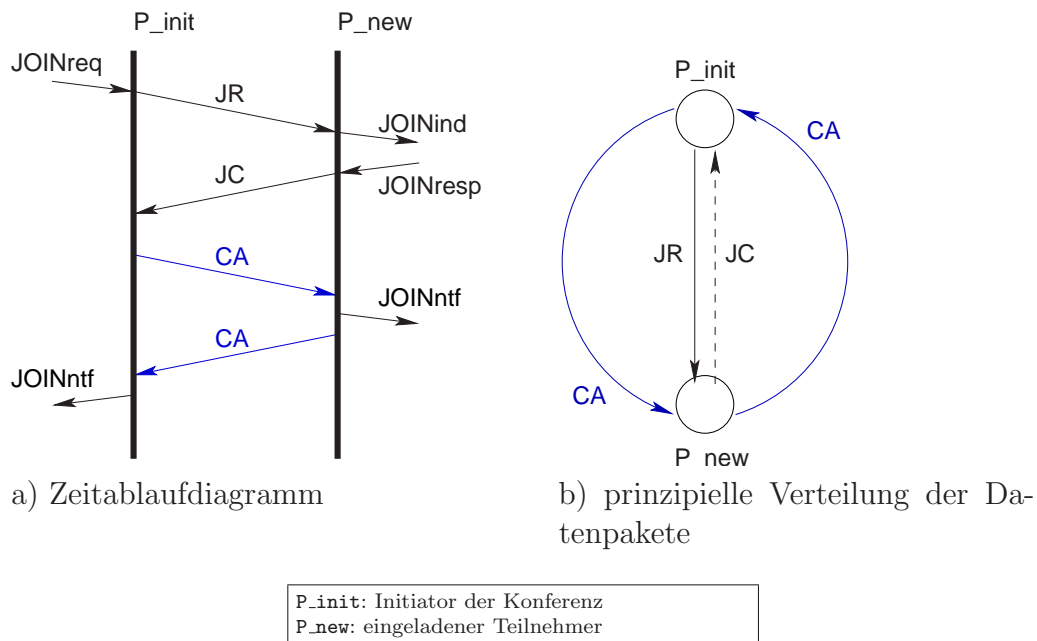


Abbildung 4.8.: Initialisierungsphase der Konferenz

Anzeige eines `JOINdis`. Sie muss dann explizit durch den Nutzer wiederholt werden.

Nach erfolgreicher Übertragung der `JR-PDU` (*JoinRequest*) wird beim eingeladenen Teilnehmer das `JOINind`-Primitiv an die Anwendung übergeben. `GCPIP` wartet jetzt auf eine Bestätigung (`JOINresp`) oder eine Ablehnung (`JOINrej`) der Einladung.

Zustimmung

Der eingeladene Teilnehmer stimmt dem Beitritt zur Konferenz mit einem `JOINresp` zu. Die Bestätigung wird dem Initiator durch eine `JC-PDU` (*JoinConfirm*) mitgeteilt. Abbildung 4.8b) zeigt den prinzipiellen Verlauf der Datenpakete während dieser Protokollphase. Der Initiator erstellt daraufhin seine Sicht auf den logischen Ring, das heißt er fügt den neuen Teilnehmer in eine Liste `hostlist` ein, in der alle Teilnehmer der Konferenz gespeichert werden, und konfiguriert die bereitgestellten UDP-beziehungswise IP-Sockets. Der Socket zum linken Nachbarn dient im `GCPIP` zur Übertragung von Nachrichten (entgegen dem Uhrzeigersinn), die über den Socket zum rechten Nachbarn bestätigt werden können (im Uhrzeigersinn)<sup>5</sup>. Ausgehend vom Initiator wird danach ein

<sup>5</sup>Die Richtung der Datenpakete spielt konzeptionell keine Rolle und könnte somit auch getauscht werden.

*CA-Paket (Connection Addition)* entgegen dem Uhrzeigersinn verschickt. *CA* dient zur Bekanntmachung des neuen Teilnehmers bei schon involvierten Teilnehmern und zur Rekonfiguration des logischen Rings. Des Weiteren bekommt der neue Teilnehmer eine Liste mit den bereits involvierten Teilnehmern. Es ist zu beachten, dass in der Initialisierungsphase der Initiator als Mitglied der Gruppe betrachtet wird.

Lehnt der eingeladene Teilnehmer die Einladung mittels *JOINrej* ab, wird dem Initiator eine *JD-PDU (JoinDisconnect)* geschickt. Dem Dienstonutzer des Initiators wird daraufhin ein *JOINdis* für die abgelehnte Einladung angezeigt.

Ablehnung

Bei Abwesenheit des eingeladenen Teilnehmers wird ebenfalls eine Ablehnung *JD* an den Initiator zurück geschickt. Dem Eingeladenen wird anders als im GCP<sub>ATM</sub> in diesem Fall auch ein *JOINdis* übergeben, um eventuelle graphische oder akustische Mitteilungen ausschalten zu können. Dies war in GCP<sub>ATM</sub> nicht vorgesehen.

Entsprechende Zeitablaufdiagramme über die Ausnahmesituationen sind im Anhang B.1 gezeigt und kurz erläutert.

### Beitritt zur Konferenz

Das Beitreten zu einer Konferenz verläuft analog zur Initialisierung. Um eine konsistente Gruppenzusammensetzung auf allen beteiligten Stationen zu gewährleisten, ist es nur dem Tokenhalter gestattet, eine Einladung einzuleiten. Wenn ein *JOINreq* vom Teilnehmer anliegt, wird ein *JR* an Eingeladenen gesendet. Nach dem Abschicken des *JR* durch den einladenden Tokenhalter gibt es zwei Möglichkeiten fortzufahren.

Die erste Möglichkeit ist, den Token nach dem Versenden des *JR* weiterzugeben, um den Kommunikations- und Signalisierungslauf nicht zu unterbrechen. Der Nachteil ist, dass dann jeder Teilnehmer eine Liste mit den Teilnehmern halten muss, denen eine Einladung geschickt wurde. Dies würde den Protokollablauf erheblich komplizierter gestalten. Deshalb wurde die zweite Möglichkeit implementiert, die den Protokollablauf bis zum Eintreffen einer Antwort des Eingeladenen unterbricht<sup>6</sup>. Dies stellt sicher, dass ein neuer Teilnehmer nicht gleichzeitig durch mehrere bereits involvierte Teilnehmer eingeladen werden kann. Des Weiteren wird so ein Beitritt eines unerwünschten Teilnehmers ver-

<sup>6</sup>Durch eine Timerüberwachung wird ein unendlich langes Warten verhindert.

#### 4. Sicherung der Geschlossenheit der Gruppe in globalen Videokonferenzsystemen

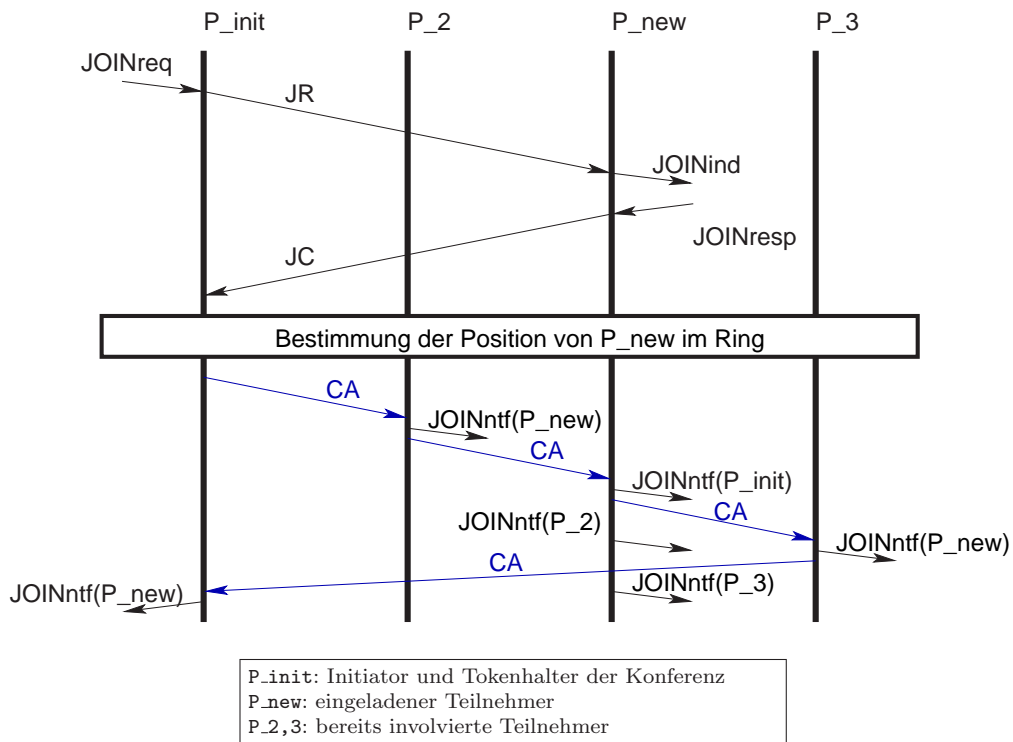


Abbildung 4.9.: Zeitablaufdiagramm für den Beitritt eines Teilnehmers

hindert, da der Tokenhalter nur auf die Antwort des Eingeladenen wartet. Der Nachteil, dass während dieser Zeit keine Steuerinformationen ausgetauscht werden können, spielt keine große Rolle. Durch die Kommunikation über den Audiokanal kann es allen Teilnehmern bekannt gemacht werden, dass ein neuer Teilnehmer eingeladen wird und dann beispielsweise die Floorkontrolle zeitweilig nicht auf Anfragen reagiert.

##### Ablehnung

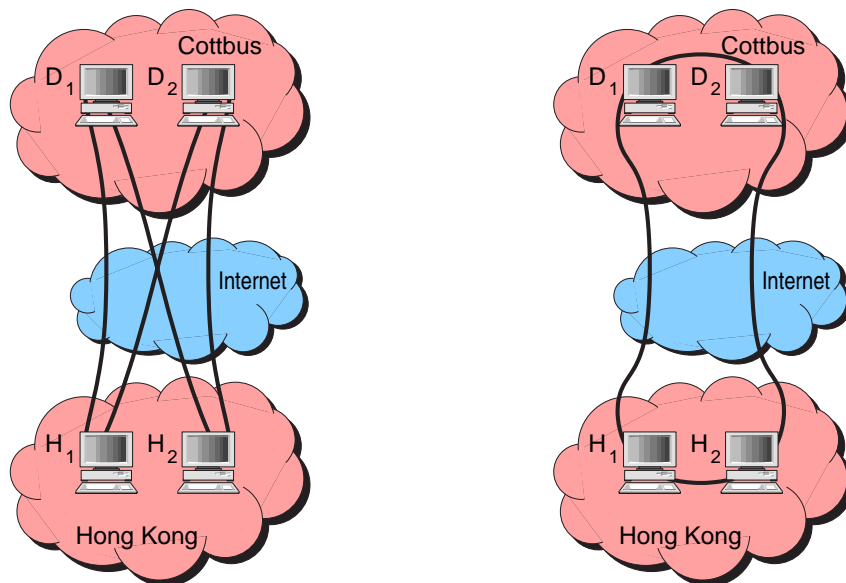
Lehnt der eingeladenen Teilnehmer die Einladung ab oder tritt ein Kommunikationsfehler auf, so wird wie bei der Initialisierung der Konferenz die Einladungsprozedur mit einem `JOINdis` beendet.

##### Zustimmung und Positionsbestimmung im logischen Ring

Empfängt der Initiator der Einladung eine Beitrittsbestätigung `JC`, muss die Position des neuen Teilnehmers in dem logischen Ring bestimmt werden. Aufgrund der Verteilung der Daten über die logische Ringstruktur im `GCPIP` dürfen die Teilnehmer nicht an einer beliebigen Stelle oder am Anfang beziehungsweise am Ende des logischen Rings eingeordnet werden. Besonders in globalen Netzen könnten die Verzögerungszeiten bei einer beliebigen Anordnung unnötig ansteigen. `GCPATM` bestand dieses Problem aufgrund der



Nutzung des Multicastdienstes nicht. Zur Veranschaulichung dieses Problems soll folgendes Beispiel dienen (siehe Abbildung 4.10).



a) nicht optimierte Ringkonfiguration

b) optimierte Ringkonfiguration

Abbildung 4.10.: Optimierung der Paketlaufzeiten im  $GCP_{IP}$ -Ring

Der Initiator  $D_1$  einer Konferenz, der sich in Deutschland befindet, hat einen Teilnehmer  $H_1$  in Hong Kong eingeladen. Zu einem späteren Zeitpunkt wurden die Teilnehmer  $D_2$  und  $H_2$  aus Deutschland und Hong Kong in die Konferenz eingeladen. Die globalen Paketlaufzeiten von Deutschland nach Hong Kong und zurück sind natürlich weitaus größer als die lokalen Verzögerungen innerhalb eines Gebietes. Bei einem Aufbau der logischen Ringe mit einer fest definierten Position, beispielsweise dem Ende der Liste, würde sich eine Sendefolge von  $D_1 \rightarrow H_1 \rightarrow D_2 \rightarrow H_2 \rightarrow D_1$  ergeben. Somit gibt es vier interkontinentale Verbindungen, die zu einer hohen Verzögerung führen (siehe Abbildung 4.10a). Günstiger ist es, die Lokalität der Teilnehmer zu bestimmen und auszunutzen. So kann das eben genannte Beispiel optimiert werden, wenn Teilnehmer aus einem Gebiet auch Nachbarn im logischen Ring sind, wie zum Beispiel  $D_1 \rightarrow D_2 \rightarrow H_2 \rightarrow H_1 \rightarrow D_1$ . Hier zeigt sich, dass nur noch zwei globale, verzögerungsintensive Verbindungen genutzt werden (siehe Abbildung 4.10b).

*Problem der Position im Ring*

Zur Bestimmung der Position des neuen Teilnehmers im logischen Ring des

*Algorithmus der Positionsbestimmung*

$GCP_{IP}$  sucht der Initiator in seiner Liste `hostlist` einen Teilnehmer, der dieselbe IP-Netzadresse wie der Eingeladene besitzt. Ist ein solcher Teilnehmer gefunden, so kann der neue Teilnehmer nach dem gefundenen Teilnehmer in dem logischen Ring integriert werden. Verläuft die Suche erfolglos, muss der neue Teilnehmer zwischen zwei IP-Netzbereichen in den logischen Ring eingefügt werden. Die sich im logischen Ring am Rand eines Netzbereichs befindlichen Teilnehmer werden dazu vom Initiator aufgefordert, die Verzögerungszeit zu dem Eingeladenen zu messen (siehe Abbildung 4.11). Der Initiator ent-

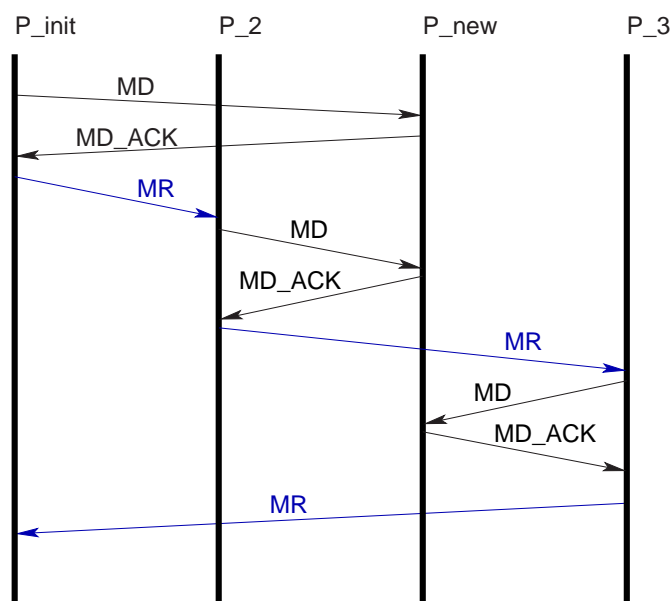


Abbildung 4.11.: Rekonfiguration des logischen Rings bei einem Beitritt eines neuen Teilnehmers

scheidet dann aufgrund der geringsten Summe aus den gemessenen Zeiten von benachbarten Teilnehmern, an welcher Stelle der neue Teilnehmer in den logischen Ring integriert wird. Dieser recht hoch erscheinende Aufwand in der Einladungsphase reduziert jedoch merklich die Tokenumlauf- und Verzögerungszeiten.

Um den Ring unter Berücksichtigung der aktuellen Netzsituationen dynamisch zu optimieren, wären ständige Messungen und die Rekonfiguration des logischen Rings notwendig. Dies würde jedoch die Komplexität des Protokollablaufs erhöhen. Da sich innerhalb einer Videokonferenzsitzung die Netzsitua-

tion höchstwahrscheinlich nicht sehr drastisch ändert, wurde auf diese komplexe Erweiterung verzichtet.

Ist die Position des neuen Eingeladenen bestimmt, wird allen Teilnehmern der Beitritt mit einer CA-PDU bekannt gemacht (siehe Abbildung 4.9). Um den Eingeladenen an die berechnete Position im logischen Ring einzuordnen, überprüft jeder Teilnehmer, ob er ein Nachbar des neuen Teilnehmers werden soll. Ist dies der Fall, werden seine Sockets so konfiguriert, dass er auf den Eingeladenen zeigt. Bei jedem Durchlauf des CA bei einem Teilnehmer wird der Eingeladene in der Liste `hostlist` registriert und ein `JOINntf` angezeigt. Der Tokenhalter nimmt die von ihm gesendete CA-PDU vom Ring und zeigt ein `JOINntf` für einen erfolgreichen Beitritt des Eingeladenen an. Bei einem Verlust des CA wird es bis zu dreimal wiederholt. Danach wird die Fehlererkennung *Error Detection* gestartet, die später erläutert wird.

Bekanntmachung

$GCP_{IP}$  verfügt nicht über eine Speicherung aller bisher ausgelieferten Datenpakete. Somit ist  $GCP_{IP}$  nicht in der Lage, die Anwendung des Eingeladenen während der Einladungsphase mit den anderen Teilnehmern zu synchronisieren. Vielmehr ist es der Anwendung selbst überlassen, dass der neue Teilnehmer die Konsistenz mit den anderen Teilnehmern herstellt. So muss zum Beispiel bei dem neuen Teilnehmer die Sprecherwarteschlange mit den gleichen Einträgen eingerichtet werden, um eine konsistente Floorkontrolle bei allen Teilnehmern zu ermöglichen.  $GCP_{IP}$  bietet hierbei die Möglichkeit, den Token für eine weitere Tokenhaltezeit zu behalten, um auf eine mögliche Synchronisationsnachricht der Anwendung zu warten. Wenn diese ausbleibt, so wird der Token weitergegeben.

Später Beitritt  
(Late Join)

### Tokenübergabe

Die Tokenübergabe im  $GCP_{IP}$  wird im Gegensatz zum  $GCP_{ATM}$  durch einem Zwei-Wege-Handshake gesichert. Der Tokenhalter übergibt mittels einer TF-PDU (*TokenForward*) den Token an den linken Nachbarn, das durch eine TF\_ACK-PDU (*Token Forward Acknowledge*) bestätigt wird (siehe Abbildung 4.12). Während der Tokenhaltezeit, die durch einen Timer  $t_{Token}$  gesteuert wird, können Daten von der Anwendung übergeben werden. Liegen keine Daten vor, wird der Token weitergereicht.

Die Tokenübergabe muss bestätigt werden. Trifft keine Betätigung ein, wird

Fehlersituationen

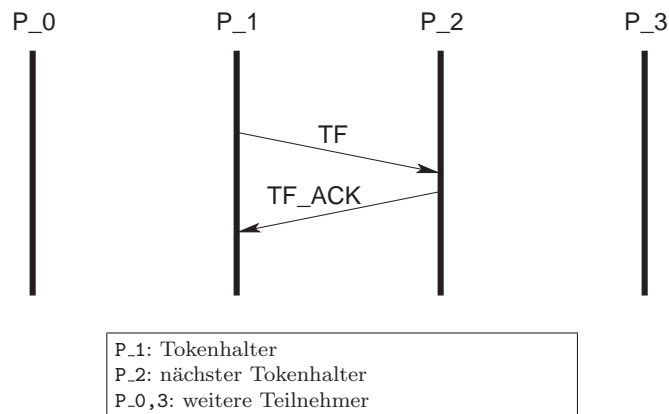


Abbildung 4.12.: Tokenübergabe im GCP<sub>IP</sub>

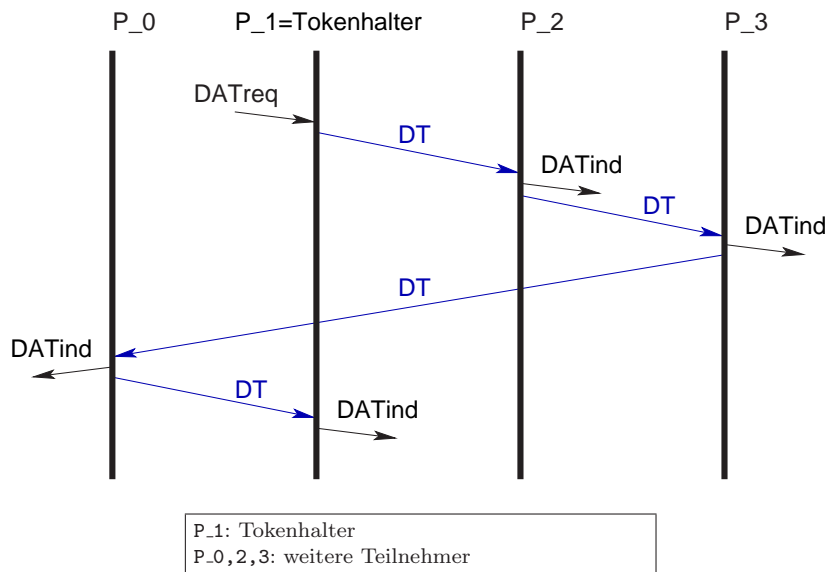
die Tokenübergabe bis zu dreimal wiederholt. Danach wird der *Forced Leave* Mechanismus eingeleitet, der den nicht antwortenden Teilnehmer aus der Konferenz ausschließt. Nach Beendigung des *Forced Leave* wird der Token dem nächstfolgenden Teilnehmer in der Liste übergeben. Um Tokenduplikate zu verhindern, enthält der Token eine Sequenznummer, die bei jedem Empfang eines Token auf seine Korrektheit überprüft wird. Ein Tokenverlust kann nur auftreten, wenn der Tokenhalter selbst ausfällt. Um diesen Fehler abzufangen, wird der *Error Detection* Mechanismus gestartet, der am Ende dieses Abschnitts erläutert wird.

### Datenübertragung

Die Datenübertragung im GCP<sub>IP</sub> unterscheidet sich grundsätzlich von der Multicast-Übertragung in GCP<sub>ATM</sub>. GCP<sub>IP</sub> verteilt die Daten über den logischen Ring, wie in der Abbildung 4.7 gezeigt wurde. Das dazugehörige Zeitablaufdiagramm der Datenübertragung enthält Abbildung 4.13.

Nur der Tokenhalter darf aus einem DATreq ein DT (*Data Transmission*) bilden und über den logischen Ring verteilen. Bei jedem Teilnehmer, bei dem ein DT empfangen wurde, wird ein DATind angezeigt und an den Nachbarn weitergeleitet. Der Tokenhalter nimmt das Paket vom Ring und gibt den Token weiter.

Geht das DT auf dem Weg verloren, so wird es vom Tokenhalter bis zu dreimal wiederholt. Bleibt die Übertragung erfolglos, wird der fehlerhafte Host im

Abbildung 4.13.: Zeitablaufdiagramm einer Datenübertragung im GCP<sub>IP</sub>

logischen Ring lokalisiert und mittels Forced Leave aus der Konferenz ausgeschlossen. In Abschnitt 4.3.2 wird auf die Prozedur zur Behebung der Fehlersituation näher eingegangen. Danach wird erneut versucht, das DT zu verteilen. Um eine doppelte Auslieferung ein und derselben Daten zu vermeiden, wird in DT auch der Absender übertragen. Da pro Tokenumlauf nur ein Datenpaket von einem Host verschickt werden darf, ist diese Identifikation ausreichend, um eine Übertragungswiederholung zu erkennen. In dieser Situation wird kein DATreq angezeigt.

### Verlassen der Gruppe

Das Verlassen der Gruppe in GCP kann entweder regulär durch einen Teilnehmer begonnen (*Regular Leave*) oder durch eine Fehlersituation erzwungen werden (*Forced Leave*). Forced Leave ist ein Teilablauf des Regular Leave. Beide Fälle sind in Abbildung 4.14 als Zeitablaufdiagramm dargestellt.

Das reguläre Austreten aus der Gruppe wird durch ein LEAVEreq des Teilnehmers eingeleitet. Um die Konsistenz des logischen Rings zu gewährleisten, darf nur der Tokenhalter die Konferenz verlassen. Der Austrittswunsch wird dem linken Nachbarn durch ein LR-Paket (*LeaveRequest*) mitgeteilt, mit dem gleichzeitig eine Tokenübergabe verbunden ist. Das LR muss durch ein LC (*Lea-*

Regular Leave

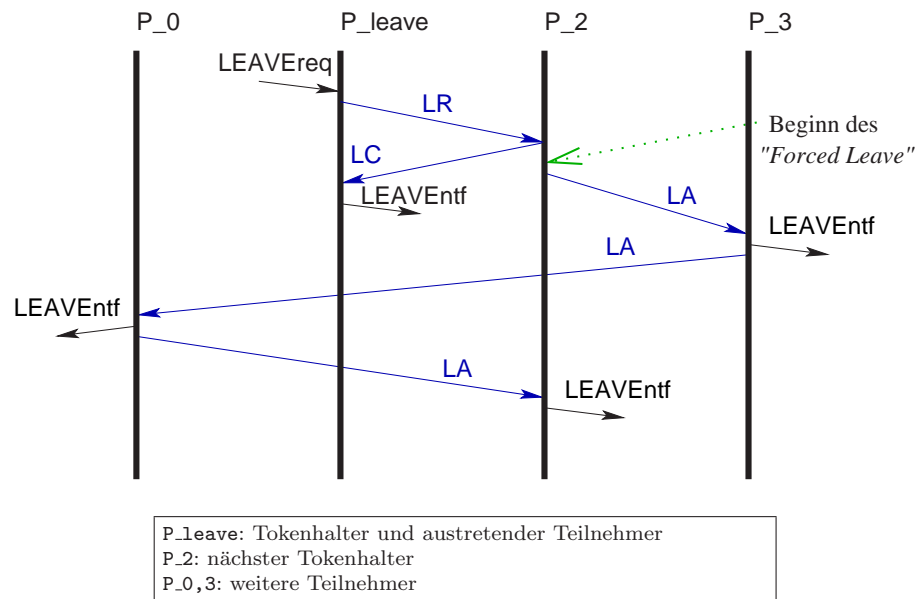


Abbildung 4.14.: Zeitablaufdiagramm für das Verlassen der Gruppe

*veConfirmation*) bestätigt werden. Der erfolgreiche Austritt wird durch ein `LEAVEntf(self)` angezeigt.

Beim Ausbleiben der Bestätigung wird das LR bis zu dreimal wiederholt. Danach wird die Gruppenzugehörigkeit des verlassenden Teilnehmers ohne Weiteres mit einem `LEAVEntf(self)` beendet. Es liegt die Vermutung nahe, dass der Teilnehmer in diesem Augenblick durch eine Netzwerkstörung von der Gruppe getrennt wurde. Die Wahrscheinlichkeit, dass dieser Fehler auftritt, ist eher gering einzuschätzen. Somit ist es ausreichend, dass dieser Fehler von der Gruppe erkannt und behoben wird.

Nachdem der Austrittswunsch von dem neuen Tokenhalter bestätigt wurde, gibt dieser mit einer *LA-PDU (LeaveAnnouncement)* den Austritt des Teilnehmers bekannt. Das LA kreist ebenfalls auf dem Ring und bewirkt die Anzeige des `LEAVEntf` sowie das Entfernen aus der `hostlist`. Bei einem Übertragungsfehler wiederholt der Tokenhalter die Übertragung des LA bis zu dreimal. Danach wird auf die Fehlererkennung des folgenden Abschnitts zurückgegriffen. Die Nachbarn des ausgetretenen Teilnehmers schließen den logischen Ring durch eine einfache Rekonfiguration der Sockets wieder. Im Gegensatz zum `GCPATM` müssen nur die Nachbarn den logischen Ring wieder schließen. Da der ATM-Multicast Dienst nur das dynamische Hinzufügen von einzel-

nen Teilnehmern in die Multicastgruppe erlaubt, aber nicht das dynamische Verlassen eines Einzelnen, war in  $GCP_{ATM}$  jeder Teilnehmer gezwungen, die ATM-Multicastgruppe zu schließen und danach wieder erneut aufzubauen. Dies führte teilweise zu Problemen im  $GCP_{ATM}$ .

Das Forced Leave wird vom Tokenhalter gestartet, nachdem ein oder mehrere fehlerhafte Teilnehmer oder Verbindungen erkannt wurden. Dazu sendet dieser ein LA-Paket durch den Ring, das alle Teilnehmer enthält, die die Konferenz aufgrund des Fehlverhaltens verlassen müssen.

*Forced Leave*

Durch den Forced Leave Mechanismus wird die Atomarität der Datenübertragung im  $GCP_{IP}$  gewahrt. Forced Leave scheint ungewöhnlich zu sein. Aber wenn ein Teilnehmer nicht mehr reagiert, sei es durch einen schwerwiegenden Fehler<sup>7</sup> im Netzwerk oder auf dem Host, macht eine weitere Teilnahme an einer multimedialen Gruppenkonferenz kaum noch Sinn, da auch die Video-, Audio- und Whiteboarddaten nicht mehr ausgetauscht werden können. Deshalb ist es günstiger, den entsprechenden Teilnehmer gezielt aus der Konferenz zu entfernen und später erneut einzuladen.

### Fehlererkennung (Error Detection)

Im  $GCP_{IP}$  können verschiedene Fehlersituationen eintreten. Zum einen kann der Tokenhalter ausfallen und der Token verloren gehen. Zum anderen arbeitet  $GCP_{IP}$  über verbindungslosen Protokollen (UDP beziehungsweise IP). Es sind deshalb Verluste beim Austausch von Datenpaketen, beispielsweise CA, DT oder LA, möglich. Diese Fälle müssen erkannt und behoben werden. Dazu dienen die nachfolgenden Mechanismen zur Fehlererkennung im  $GCP_{IP}$ .

Nicht-Tokenhalter erwarten in einer durch die Round-Trip-Time (RTT) des Tokens bestimmten Zeit  $t_{xx}$ , dass sie wieder Tokenhalter werden. Ist dies nicht der Fall, ist ein Tokenverlust aufgetreten. Daraufhin sendet einer der Teilnehmer ein ED-Paket (*Error Detection*) durch den Ring (siehe Abbildung 4.15). Dieser Teilnehmer wird zum vorläufigen Tokenhalter erklärt. Die ED-PDU muss bestätigt werden, um die Fehlerstelle des logischen Rings zu lokalisieren. Wird ED nicht bestätigt, ist ein Fehler im Ring gefunden worden.

*Tokenverlust*

<sup>7</sup>Es sei darauf hingewiesen, dass Forced Leave erst gestartet wird, wenn mindestens 4 Tokenübergaben oder 4 Datenübertragungen mit anschließender Fehlererkennung (wobei auch hier mindestens 4 Pakete verloren gehen müssen) fehlgeschlagen sind. Dies deutet also auf einen schweren Fehler entweder im Netzwerk oder auf dem entsprechenden Host.

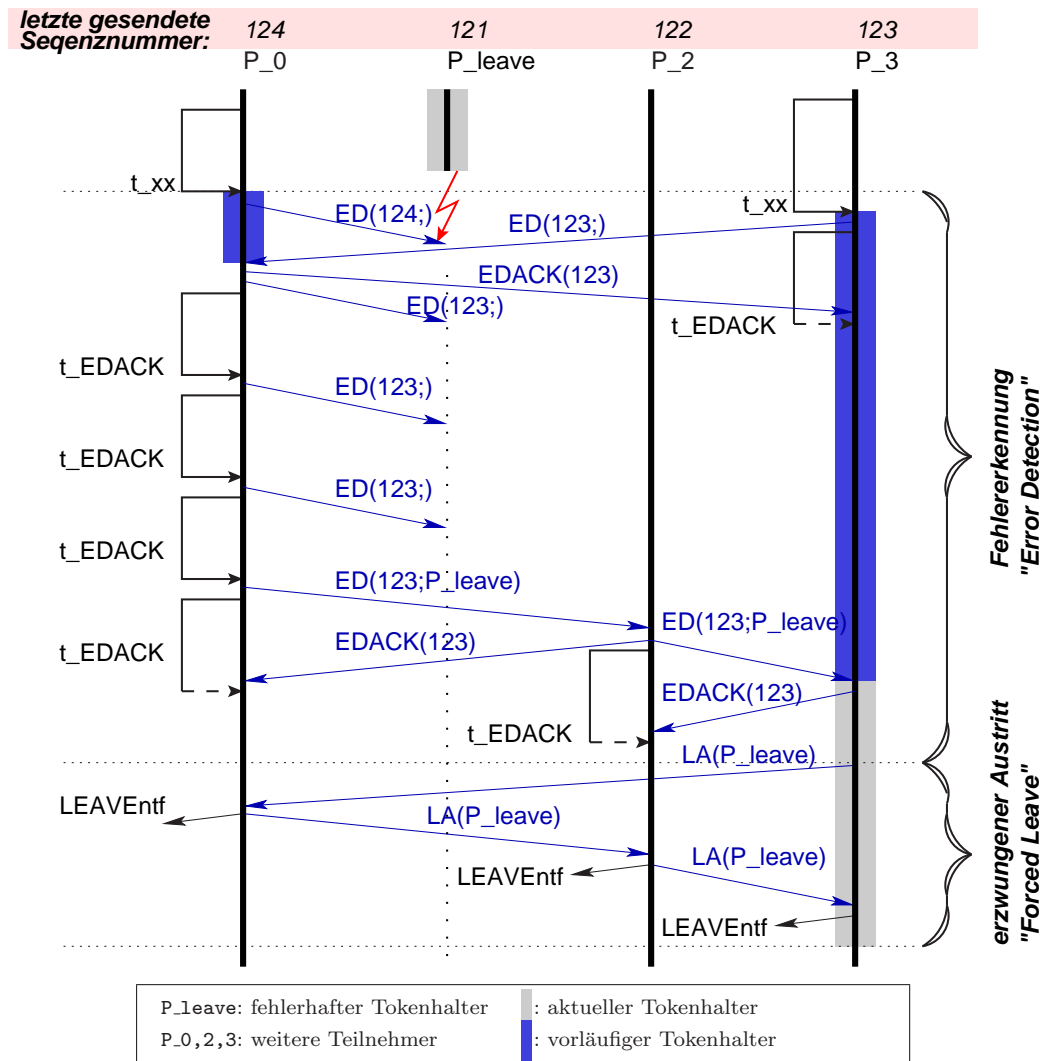


Abbildung 4.15.: Fehlererkennung

Der nicht antwortende Teilnehmer wird in dem ED-Paket gespeichert und dem nächsten Teilnehmer im logischen Ring das Paket zugeschickt. Empfängt der vorläufige Tokenhalter sein eigenes ED-Paket, wird er zum Tokenhalter und startet den Forced-Leave-Mechanismus für alle Teilnehmer, die in dem ED-Paket enthalten sind. Somit wird der Ring wieder konsistent rekonfiguriert.

Zum Schutz davor, dass zwei Teilnehmer die Fehlersituation quasi gleichzeitig entdecken und zum vorläufigen Tokenhalter werden, enthält das ED die zuletzt durch den Teilnehmer versendete Sequenznummer des Token. Bekommt ein vorläufiger Tokenhalter in dieser Situation ein ED-Paket mit einer höheren



Sequenznummer als der eigenen, wird es nicht weitergeleitet. Erhält dieser im umgekehrten Fall ein ED mit einer kleineren Sequenznummer, verliert er den Status als vorläufiger Tokenhalter. Somit wird der Tokenhalter wieder eindeutig (siehe Abbildung 4.15).

Wenn Pakete des Typs CA, DT oder LA auf dem Weg durch den Ring verloren gehen, wird die Übertragung der Pakete durch den Tokenhalter bis zu dreimal wiederholt. Sind sie trotz mehrfacher Wiederholung nicht wieder beim Tokenhalter angekommen, muss dieser den Fehlerpunkt im logischen Ring lokalisieren. Dazu wird vom Tokenhalter eine ED-PDU verschickt, die nach dem gleichen Prinzip wie bei einem Tokenverlust den Fehlerpunkt lokalisiert. Durch den Forced Leave Mechanismus werden auch hier die fehlerhaften Teilnehmer aus dem Ring entfernt. Im Anschluss daran wird erneut das CA-, DT- oder LA-Paket im Ring verteilt, das vor der Fehlersituation verschickt werden sollte.

Weitere Fehlersituationen können durch den Verlust der PDUs JR, JC, LR und LC auftreten. Diese können durch einfache Timerfunktionalitäten erkannt und daraufhin behoben werden (siehe Anhang B.1).

*Fehlerhafte Datenübertragung*

### 4.3.3. Diskussion der $GCP_{IP}$ Lösung

Protokolle können von vielen Standorten aus betrachtet und bewertet werden. Die wichtigsten Kriterien für die Bewertung eines Signalisierungsprotokolls sind jedoch die Verfügbarkeit und Nutzbarkeit, die Robustheit gegenüber Fehlerfällen, die Verzögerung einer Datenübertragung sowie der Bandbreitenverbrauch beziehungsweise der Overhead. Der Durchsatz des Protokolls ist bei der Signalisierung nicht von großer Bedeutung, da davon ausgegangen wird, dass nur selten Nachrichten gesendet werden, die zudem kurz sind.

#### **Nutzbarkeit**

$GCP_{IP}$  stellt einen zuverlässigen, atomaren und total geordneten Datenübertragungsdienst für die Signalisierung in geschlossenen Mehrteilnehmer-Videoferenzsystemen unter Nutzung von Unicastverbindungen im Internet bereit. Da keine Multicastdienste für das Funktionieren des Protokolls notwendig sind, kann  $GCP_{IP}$  uneingeschränkt an allen Endsystemen benutzt werden, die den UDP/IP-Protokollstack installiert haben.

### Robustheit

Das GCP<sub>IP</sub> wurde aus der Notwendigkeit entworfen, die Gruppenkoordinatoren des OCTOPUS-Systems über das Internet zu verbinden. Diese Umgebung ist jedoch fehleranfällig, so dass Pakete verloren gehen, dupliziert oder verfälscht werden können. In der Ringtopologie des GCP<sub>IP</sub> können Fehler sporadisch an unterschiedlichen Stellen auftreten. Bei der Gestaltung des Protokolls, stellte sich die Frage, ob jede Übergabe der DT-, CA- und LA-PDUs von Host zu Host bestätigt werden muss oder ob es ausreicht, dass der Tokenhalter eine implizite Bestätigung durch den Umlauf des Pakets bekommt. Letzteres würde das Protokoll stark vereinfachen. Zudem können Bandbreite gespart und Verzögerungszeiten minimiert werden. Zur Beantwortung dieser Frage wurden im Vorfeld der GCP<sub>IP</sub>-Entwicklung Messungen der Paketverluste zu verschiedenen Standorten hin durchgeführt. In Abbildung 4.16 sind die Resultate dieser Messungen dargestellt.

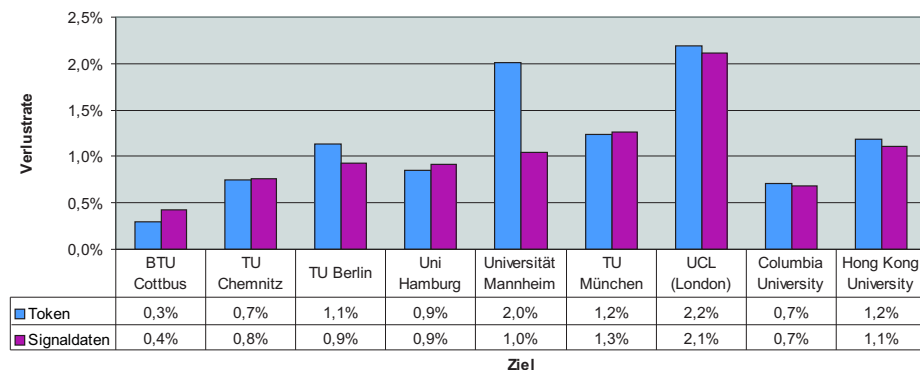


Abbildung 4.16.: Messungen der Verlustrate zu verschiedenen Standorten

Messbeschreibung:  
Verlustrate

Vom der BTU Cottbus aus wurden jeweils 8000 Messungen über einen Zeitraum von 4 Wochen durchgeführt. Dabei wurden die Antwortzeiten mit dem UNIX-Programm *ping* gemessen. Es wurden Paketlängen von 12 und 208 Byte benutzt, um den Einfluss der Paketlänge auf die Verlustwahrscheinlichkeit zu erkennen. Die Paketlängen entsprachen einem Paket für die Tokenübergabe TF beziehungsweise für die Signaldaten DT im GCP<sub>ATM</sub><sup>8</sup>. Wie die Resultate zei-

<sup>8</sup>Es wurden die Paketgrößen des GCP<sub>ATM</sub> genutzt, da die Messungen im Vorfeld der GCP<sub>IP</sub> Entwicklung durchgeführt wurden und zu diesem Zeitpunkt die Paketgrößen für GCP<sub>IP</sub> noch nicht feststanden.

gen, liegen die Fehlerraten bei etwa einem Prozent. Somit lässt sich abschätzen, dass sporadisch verteilte Fehlerquellen zwischen den Teilnehmern des logischen Rings nicht sehr häufig auftreten. Es ist also ausreichend, dass eine Datenübertragung nur beim Tokenhalter implizit bestätigt wird. Bei den Messungen hat sich außerdem gezeigt, dass Verluste von mehr als drei aufeinander folgenden Paketen kaum aufgetreten sind. Dadurch wurde bestätigt, dass die Wahl für die Anzahl der Übertragungswiederholungen von 3 gerechtfertigt ist. Zwischen den Standorten BTU Cottbus und Hong Kong wurden weiterhin mehrere erfolgreiche Tests über jeweils 24 Stunden mit dem GCP<sub>IP</sub> durchgeführt, bei denen eine Gruppe mit ein bis zwei Teilnehmern pro Standort aufgebaut wurde. Es hat sich gezeigt, dass während der Tests zwar Paketverluste aufgetreten sind, aber kein einziger Forced Leave ausgeführt wurde. In weit fehleranfälligen Umgebungen wie einem *Wireless LAN* wurde das Protokoll und sein Verhalten jedoch noch nicht getestet.

### Performanz des GCP<sub>IP</sub>

Das GCP<sub>IP</sub> ist aufgrund seiner Ringtopologie nicht beliebig in Bezug auf die Anzahl der Gruppenmitglieder skalierbar. Eine höhere Anzahl von Teilnehmern bewirkt einen Anstieg der Token- und Signalumlaufzeiten und somit der Reaktionszeiten  $t_{\text{Reaktion}}$  auf eine Anforderung aus der Nutzerumgebung. Werden diese Anforderungen mit einer starken Verzögerung bearbeitet, leidet darunter die Bedienbarkeit des Systems. Beispielsweise sollte in einer Anwendung eine Reaktion auf einen Tastendruck innerhalb einer Zeit von 100 ms bis 2 s [139], [93], [103] erfolgen, damit die Anwendung beim Nutzer nicht als abgestürzt gilt. In [10] wird von einer maximalen Verzögerung von 1s ausgegangen. Es stellt sich die Frage, bis zu welcher Zahl von Gruppenmitgliedern der Ansatz des GCP<sub>IP</sub> akzeptabel ist.

Allgemeingültig lässt sich diese Frage nicht beantworten, da der Einflussfaktor *Entfernung zwischen den Standorten* in jeder Sitzung sehr unterschiedlich sein kann. Eine Konferenz, bei der sich die Konferenzteilnehmer über den gesamten Globus verteilen, wird nur eine kleinere Gruppengröße zulassen als eine Konferenz, die in einer Region oder auf einem Universitätscampus stattfindet. Aus diesem Grund wurden die Messungen innerhalb eines lokalen Netzwerks durchgeführt, um zuerst die Eigenschaften des Protokolls unabhängig von Entfernungseinflüssen zu untersuchen.

Für die Messungen wurde ein lokales 100 Mbit/s Ethernet-Netzwerk genutzt, an dem unterschiedlich leistungsstarke Rechner angeschlossen waren, wie beispielsweise mehrere SUN SparcStation 20, SUN Ultra 10 und SUN LX. Da keine aufwendigen Rechenoperationen durch das GCP<sub>IP</sub> notwendig sind, hat sich jedoch die Rechenleistung der beteiligten Rechner kaum bemerkbar gemacht. Die Reaktionszeit  $t_{\text{Reaktion}, l}$ <sup>9</sup> auf ein Ereignis des Dienstnutzers wird durch das Zeitintervall zwischen dem Absenden des DATreq und dem Empfang des DATind bestimmt und hängt maßgeblich von der Summe aus der Wartezeit  $t_W$  für ein DATreq-Primitiv, bis zum Eintreffen des Token, und der Datenübertragungszeit  $t_{\text{DAT}, l}$  ab:

$$t_{\text{Reaktion}, l} = t_{\text{DAT}, l} + t_W \quad (4.1)$$

Je nachdem, ob der Host gerade den Token besitzt oder bis zu einem vollen Tokenumlauf warten muss, liegt  $t_W$  im Intervall  $[0, t_{\text{RTT-Token}, l}]$ , mit  $t_{\text{RTT-Token}, l}$  als Tokenumlaufzeit im LAN.

Die Reaktionszeit  $t_{\text{Reaktion}, l}$  liegt demnach im Intervall

$$t_{\text{DAT}, l} \leq t_{\text{Reaktion}, l} \leq t_{\text{DAT}, l} + t_{\text{RTT-Token}, l}. \quad (4.2)$$

Um den Einfluss beider Größen getrennt einzuschätzen, wurde einerseits die Tokenumlaufzeit sowie die Verzögerung eines Datenpakets vom Zeitpunkt des Übernahme durch das GCP<sub>IP</sub> bis hin zum Ausliefern an den Dienstnutzer gemessen. Beide Messungen wurden getrennt durchgeführt, um gegenseitige Abhängigkeiten auszuschließen.

Die Messergebnisse der Tokenumlaufzeit aus drei Messreihen á 5000 Einzelmessungen sind in Abbildung 4.17 dargestellt. Die durchschnittliche Tokenumlaufzeit bei 8 Teilnehmern lag bei etwa 55 ms. Mittels Microsoft Excel wurde die lineare Approximation<sup>10</sup> des Messergebnisses:

$$t_{\text{RTT-Token}, l} = 7.42n - 5.78 \quad (4.3)$$

bestimmt, wobei  $n$  die Anzahl der Teilnehmer des GCP<sub>IP</sub>-Rings darstellt.

<sup>9</sup>Der Index  $l$  steht für lokal.

<sup>10</sup>Die Gleichungen werden hier zur besseren Übersicht ohne Einheiten angegeben. Die Einheit der Zeit in den Gleichungen ist jeweils Millisekunden.

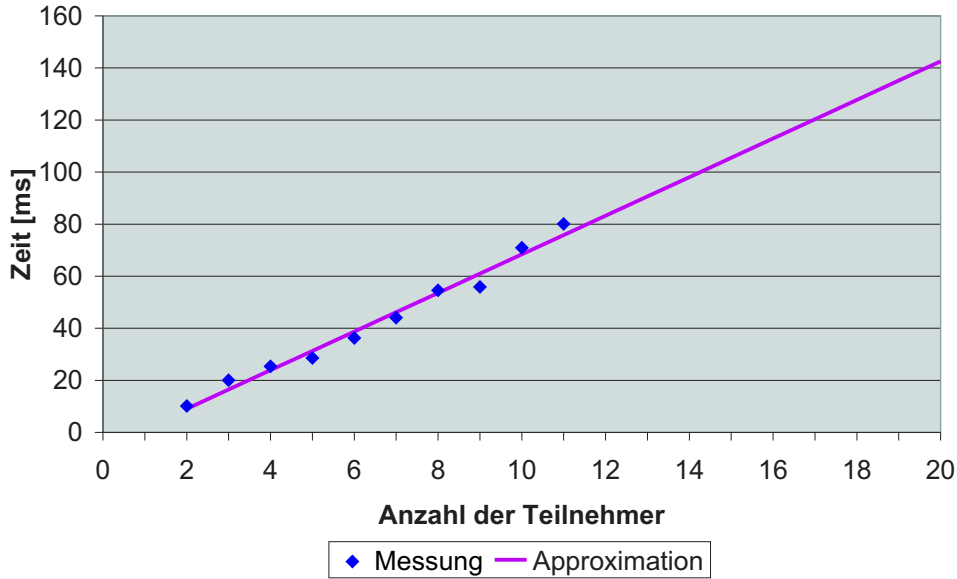


Abbildung 4.17.: Messung und Approximation der Tokenumlaufzeit  $t_{\text{RTT-Token}, 1}$

Die Messungen der Datenübertragungszeiten wurden ebenfalls dreimal mit jeweils 5000 Messwerten durchgeführt. Da die Uhren trotz des *Network Time Protocols (NTP)* [104] nicht hundertprozentig synchron laufen, wird die Empfangszeit des `DATind` nur auf dem Sendehost des `DATreq` registriert und daraus die Zeitdifferenz berechnet. Die Ergebnisse sind in Abbildung 4.18 dargestellt. Auch hier ist eine Approximation der Verzögerungszeiten angegeben, die nach folgender Gleichung berechnet wird:

$$t_{\text{DAT}, 1} = 1.29n - 1.25 \quad (4.4)$$

Diese Gleichung entspricht auch der minimalen Reaktionszeit  $\check{t}_{\text{Reaktion}, 1}$  auf ein Ereignis des Dienstinutzers. Beide Messungen zusammen ergeben die maximale Reaktionszeit  $\hat{t}_{\text{Reaktion}, 1}$ :

$$\hat{t}_{\text{Reaktion}, 1} = 8.71n - 6.83 \quad (4.5)$$

In Abbildung 4.19 ist diese Approximation als *Maximum* dargestellt. Die Abbildung enthält außerdem die Messung und Approximation der tatsächlichen

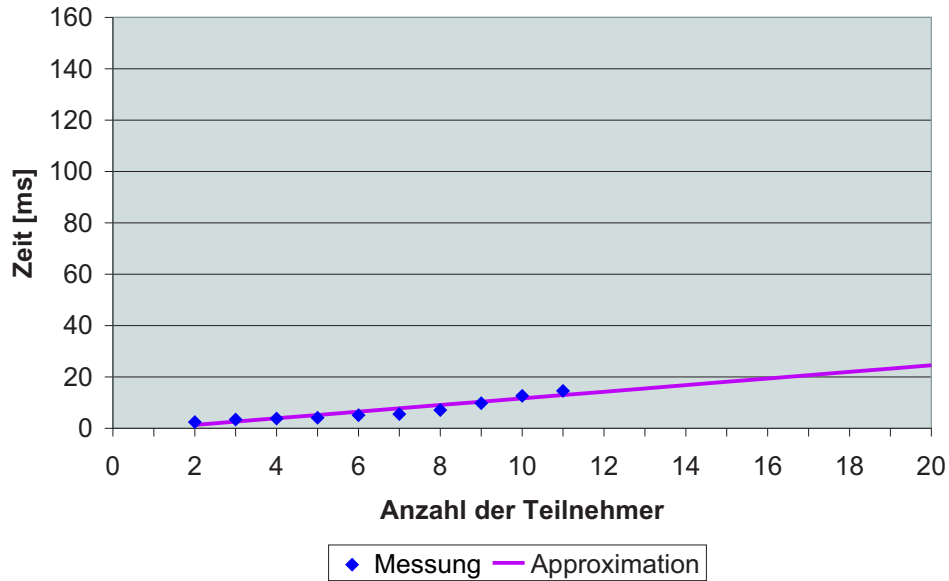


Abbildung 4.18.: Messung und Approximation der Datenübertragungszeiten  $t_{\text{DAT}, 1}$

Verzögerung vom Absenden eines  $\text{DATreq}$  bis zum Empfang des  $\text{DATind}$ <sup>11</sup>. Das *Minimum* stellt in der Abbildung die Approximation der minimalen Reaktionszeit  $\bar{t}_{\text{Reaktion}, 1}$  dar. Daraus ist ersichtlich, dass die tatsächliche Reaktionszeit bei nur etwa einem Drittel zwischen dem Maximum und dem Minimum liegt. Die Approximation der Messergebnisse lautet:

$$\bar{t}_{\text{Reaktion}, 1} = 4.55n - 9.11 \quad (4.6)$$

Unter der Annahme, dass die maximale Reaktionszeit unter 1000 ms liegen muss, ist aus Gleichung 4.5 die maximale Anzahl von Teilnehmern bestimmbar:

$$\hat{n} = \frac{1000 + 6.83}{8.71} \approx 116 \quad (4.7)$$

Demnach sind 116 Teilnehmer innerhalb einer  $\text{GCP}_{\text{IP}}$ -Sitzung möglich. Unter der Annahme des gemessenen Durchschnitts aus Gleichung 4.6 ist eine Gruppengröße von etwa 221 Teilnehmern vorstellbar. Zu beachten ist, dass es sich hierbei um eine lokale Abschätzung handelt. Wie schon erwähnt, verrin-

<sup>11</sup>Diese Messung wurde auch für  $\text{GCP}_{\text{B}}$  durchgeführt, das Multicastverbindungen auf Unicastverbindungen abbildet. Das Ergebnis ist in Anhang C kurz diskutiert.

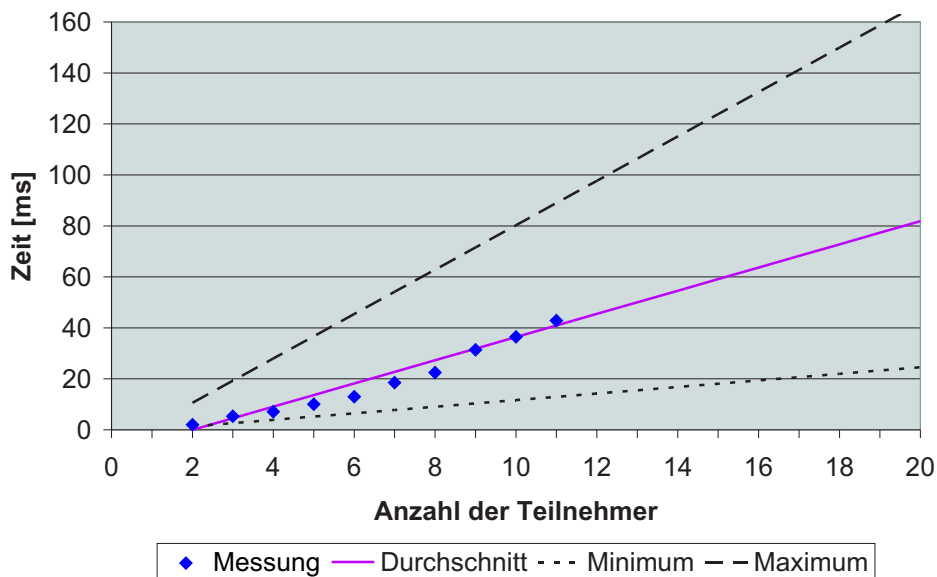


Abbildung 4.19.: Approximation der maximalen Gesamtverzögerung und Messung der tatsächlichen, durchschnittlichen Verzögerung durch  $GCP_{IP}$

gert sich diese Zahl, wenn ein globales Umfeld mit zusätzlichen Leitungs- und Routingverzögerungen als Einsatzumgebung des  $GCP_{IP}$  dient.

Um eine Aussage über die Größe dieser Verzögerungen und deren Einfluss auf die Einsatzfähigkeit des  $GCP_{IP}$  zu erhalten, wurden die Entfernungen zu verschiedenen Standorten von Cottbus aus mit dem Programm *ping* gemessen. Die Resultate sind in Abbildung 4.20 dargestellt. Die Messungen zeigen, dass im Verbund der deutschen Hochschulen und sogar innerhalb Europas Antwortzeiten von unter 40 ms realistisch sind. Interkontinentale Verbindungen, beispielsweise in die USA, können mit weniger als 100 ms Round-Trip-Time abgeschätzt werden. Der Ursache für die hohen Verzögerungen nach Hong Kong liegt in der Nutzung einer Satellitenverbindung zwischen Frankreich und Singapur<sup>12</sup>.

Globale  
Abschätzung

<sup>12</sup>Mittels des UNIX-Programms *traceroute* wurde diese Verbindung nachgewiesen

#### 4. Sicherung der Geschlossenheit der Gruppe in globalen Videokonferenzsystemen

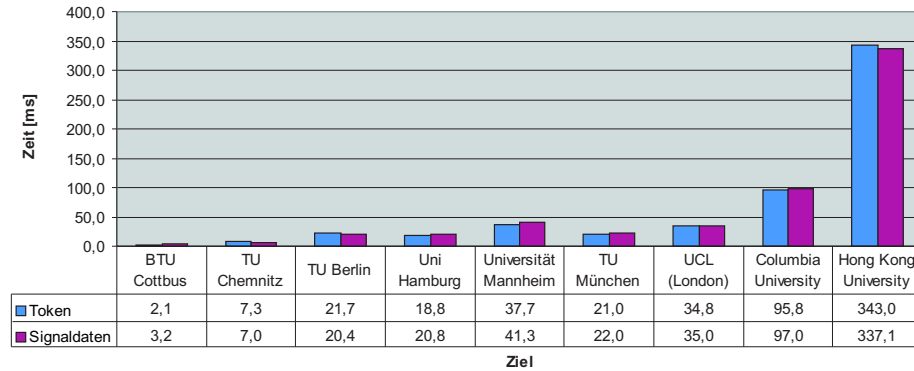


Abbildung 4.20.: Messung der Verzögerungen zu verschiedenen Standorten

Um die Anzahl der Teilnehmer in einem globalen Umfeld abschätzen zu können, kann der Ansatz aus Gleichung 4.2 verwendet werden. Somit ist die maximale Reaktionszeit im globalen Umfeld:

$$\hat{t}_{\text{Reaktion, g}} = t_{\text{RTT-Token, g}} + t_{\text{DAT, g}} \quad (4.8)$$

Unter Vernachlässigung der Leitungsverzögerung  $t_L$  in lokalen Netzwerken beträgt die globale Token-RTT:

$$t_{\text{RTT-Token, g}} = t_{\text{RTT-Token, l}} + \underbrace{\sum_{i=1}^{n-1} t_{L\{i,i+1\}} + t_{L\{n,1\}}}_{\text{globale Leitungsverzögerungen}} \quad (4.9)$$

mit  $n$  als Anzahl der Teilnehmer. Das Tupel  $\{i, i+1\}$  bezeichnet die Verbindung zwischen Teilnehmer  $i$  und seinem Nachbarn  $i+1$  in einem  $\text{GCP}_{\text{IP}}$ -Ring.

Für die globale Verzögerung der Datenübertragung gilt:

$$t_{\text{DAT, g}} = t_{\text{DAT, l}} + \sum_{i=1}^{n-1} t_{L\{i,i+1\}} + t_{L\{n,1\}} \quad (4.10)$$



Somit ergibt sich aus den Gleichungen 4.9 und 4.10 für eine globale Sitzung mit dem  $GCP_{IP}$  folgende Abschätzung der Reaktionszeit auf ein Nutzerereignis:

$$t_{\text{Reaktion, g}} = t_{\text{DAT, 1}} + t_{\text{RTT-Token, 1}} + 2 \left( \sum_{i=1}^{n-1} t_{L\{i,i+1\}} + t_{L\{n,1\}} \right) \quad (4.11)$$

Um die maximale Anzahl der Teilnehmer zu bestimmen, muss nun noch die Ungleichung

$$t_{\text{Reaktion, g}} \leq 1000 \text{ ms} \quad (4.12)$$

erfüllt sein.

Dies soll anhand eines Beispiels erläutert werden. Für Europa wurden RTTs von unter 40 ms gemessen. Daraus folgt, dass die einfache Verzögerung etwa 20 ms beträgt (siehe Abbildung 4.20). Unter der Annahme, dass alle Strecken eine Verzögerung von 20 ms besitzen ist, lässt sich die Gleichung 4.11 folgendermaßen vereinfachen:

$$t_{\text{Reaktion, g}} = t_{\text{DAT, 1}} + t_{\text{RTT-Token, 1}} + 2n \cdot 20 \text{ ms} \quad (4.13)$$

mit  $n$  als Anzahl der Teilnehmer. Somit lautet die Ungleichung 4.12 unter Nutzung der Approximation aus Gleichung 4.5:

$$40n + 8.71n - 6.83 \leq 1000 \text{ ms}, \quad (4.14)$$

und die Lösung dessen:

$$\hat{n} = \frac{1000 + 6.83}{48.71} \approx 21 \quad (4.15)$$

Somit können europaweit etwa 21 Teilnehmer gleichzeitig an einer  $GCP_{IP}$ -Sitzung teilnehmen, ohne dass sich Verzögerungen störend bemerkbar machen. Die Anzahl der Teilnehmer steigt auf etwa 23 unter Berücksichtigung der durchschnittlichen Reaktionszeit aus Gleichung 4.6.

Für Videokonferenzen über mehrere Kontinente hinweg ist leider nur eine sehr geringe Anzahl<sup>13</sup> von Teilnehmern pro Gruppe möglich, da hohe Leitungsverzögerungen sich zu stark auf die Zeiten für den Tokenumlauf und die Datenübertragung auswirken. Sind sich die Teilnehmer jedoch über die

<sup>13</sup>Konkret lässt sich die Anzahl jedoch nicht hinreichend genau bestimmen, da die Leitungsverzögerungen weltweit zu stark variieren.

physikalisch bedingten, starken Verzögerungen bei einer globalen Konferenz bewusst, so kann der Erwartungswert von 1000 ms und somit die maximale Teilnehmerzahl erhöht werden.

### Durchsatz

Die Übertragung von sehr vielen Steuerinformationen von allen Teilnehmern gleichzeitig, scheint für  $GCP_{IP}$  problematisch zu sein, da der Tokenumlauf und die Datenübertragung Zeit beansprucht. Aus diesem Grund wird jetzt der Durchsatz des Protokolls untersucht, also wie viele Pakete pro Sekunde vom  $GCP_{IP}$  verteilt werden können, ohne bei den Sendern einen Stau zu verursachen. Anders ausgedrückt, wie oft dürfen Steuerinformationen ausgetauscht werden.

Für diese Analyse wird angenommen, dass sobald ein Teilnehmer im  $GCP_{IP}$  der Tokenhalter wird, ein Paket zur Übertragung bereitliegt. Demzufolge werden pro Tokenumlauf  $n$  Pakete übertragen, wenn  $n$  die Anzahl der Teilnehmer ist. Somit setzt sich die Zeit, die zwischen dem Absenden von zwei Paketen vergeht, aus  $n$  Datenübertragungszeiten  $t_{DAT}$  und einer Tokenumlaufzeit  $t_{RTT-Token}$  zusammen.

$$t = n \cdot t_{DAT} + t_{RTT-Token} \quad (4.16)$$

lokale  
Abschätzung

Unter Berücksichtigung der lokalen Approximationen in Gleichung 4.3 und 4.4 kann diese Zeit wie folgt berechnet werden:

$$t = n \cdot t_{DAT, 1} + t_{RTT-Token, 1} = 1.29n^2 - 6.17n + 5.78 \quad (4.17)$$

Daraus ergibt sich ein geschätzter lokaler Durchsatz<sup>14</sup>, den ein Teilnehmer über das  $GCP_{IP}$  übertragen kann:

$$B_{Eff., 1} = \frac{1000}{1.29n^2 - 6.17n + 5.78} \quad (4.18)$$

globale  
Abschätzung

Mit den Gleichungen 4.9 und 4.10 lässt sich die Zeit zwischen dem Absenden

---

<sup>14</sup>Zur Erhaltung der Übersichtlichkeit werden die Gleichungen auch hier ohne Einheiten dargestellt. Der Durchsatz wird in Pakete pro Sekunde angegeben.

von zwei Paketen in globalen Netzen abschätzen:

$$t = n \cdot \left( t_{\text{DAT}, 1} + \sum_{i=1}^{n-1} t_{L\{i,i+1\}} + t_{L\{n,1\}} \right) + t_{\text{RTT-Token}, 1} + \sum_{i=1}^{n-1} t_{L\{i,i+1\}} + t_{L\{n,1\}} \quad (4.19)$$

Wird jetzt wieder von der Annahme ausgegangen, dass innerhalb von Europa eine Leitungsverzögerung von 20 ms gilt, kann folgende Approximation für den Durchsatz abgeleitet werden:

$$B_{\text{Eff., g}} = \frac{1000}{n(1.29n - 1.25) + (7.42n - 5.78) + (n + 1) \cdot 20} \quad (4.20)$$

$$= \frac{1000}{1.29n^2 + 26.17n + 20} \quad (4.21)$$

Die Gleichungen 4.18 und 4.21 sind in Abbildung 4.21 graphisch dargestellt. Daraus ist ersichtlich, dass bei 20 Teilnehmern im lokalen Bereich etwa 1,6

Bewertung

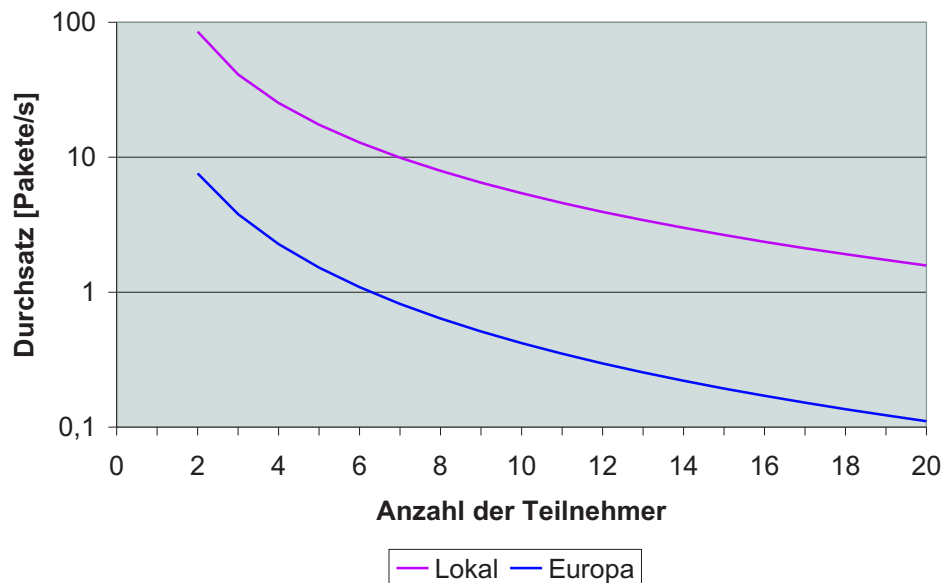


Abbildung 4.21.: Maximaler Durchsatz von  $GCP_{IP}$

Pakete pro Sekunde<sup>15</sup>, also alle 600 ms, abgeschickt werden können. Für das Management eines Videokonferenzsystem ist dies völlig ausreichen, da ein Spre-

<sup>15</sup>Die Einheit Byte pro Sekunde würde weniger aussagekräftig sein, da die Pakete eine feste Größe von etwa 200 Byte haben.

cherwechsel oder eine Anfrage an die Floorkontrolle nicht so häufig stattfindet. Die Zeit zwischen zwei Anfragen liegt in der Regel im Minutenbereich. Auch das QoS-Management tauscht nicht so häufig Daten aus. Hier werden im Abstand 5 bis 10 Sekunden die aktuellen QoS-Zustandsinformationen ausgetauscht. In einer globalen Umgebung müssen die Pakete eines Teilnehmers einen Abstand von etwa 9 Sekunden haben, um sich nicht bei diesem zu stauen. Somit ist für die Floorkontrolle auch kein Engpass zu erwarten. Für das QoS-Management kann es allerdings problematisch werden, wenn alle 5 Sekunden der aktuelle QoS-Zustand an die anderen Teilnehmer verteilt werden soll. Deshalb muss hier die QoS-Anwendung die Übertragungshäufigkeit in Abhängigkeit der Teilnehmerzahl reduzieren. Ein Ansatz wie in RTCP kann hier genutzt werden [137]. GCP<sub>IP</sub> ist somit in der Lage, den Anforderungen gerecht zu werden, die Managementkomponenten eines Videokonferenzsystems an den Durchsatz stellen.

#### **Bandbreitenverbrauch**

Die Signalisierung in einem Videokonferenzsystem erzeugt einen Mehraufwand (*Overhead*) zu der eigentlich interessierenden Übertragung der multimedialen Daten. Dieser Mehraufwand sollte möglichst gering sein. Der Overhead, der durch das GCP<sub>IP</sub> entsteht, kann nicht unmittelbar berechnet werden, da dieser von der Anwendung über dem GCP<sub>IP</sub> und dessen Sendeverhalten abhängt. Es kann lediglich der Bandbreitenbedarf für den Tokenmechanismus bestimmt werden.

Um die Bandbreite zu berechnen, die ein Teilnehmer durch die Signalisierung erzeugt, wird das Ergebnis der Tokenumlaufzeit herangezogen. Das gesendete Datenvolumen leitet sich aus den Paketlängen  $L$  der TF- und TF\_ACK-PDU von jeweils 12 Byte ab. Die Gleichung:

$$B_i = \frac{L_{\text{TF-PDU}} + L_{\text{TF\_ACK-PDU}}}{t_{\text{RTT-Token}}} \quad (4.22)$$

ergibt die Bandbreite, die ein Teilnehmer  $i$  in der GCP<sub>IP</sub>-Sitzung verbraucht. Der Zähler der Gleichung gibt an, wieviele Daten pro Zeiteinheit, hier pro Tokenumlauf  $t_{\text{RTT-Token}}$ , von dem Sender verschickt werden. Dies ist die Summe aus der Länge der TF-PDU  $L_{\text{TF-PDU}}$  und der TF\_ACK-PDU  $L_{\text{TF\_ACK-PDU}}$ . Beide PDUs sind 12 Byte lang. Die Tokenumlaufzeit ist abhängig von der Anzahl

der Teilnehmer und wird nach Gleichung 4.3 bestimmt. Die Gesamtbandbreite pro Sitzung ergibt sich somit zu:

$$B_{\text{gesamt}} = nB_i \quad (4.23)$$

Die Gesamtbandbreite ist in Abbildung 4.22 dargestellt. Des Weiteren ist das Verhältnis dieser Bandbreite zu der gesamten Bandbreite einer Videokonferenz<sup>16</sup> aufgezeichnet.

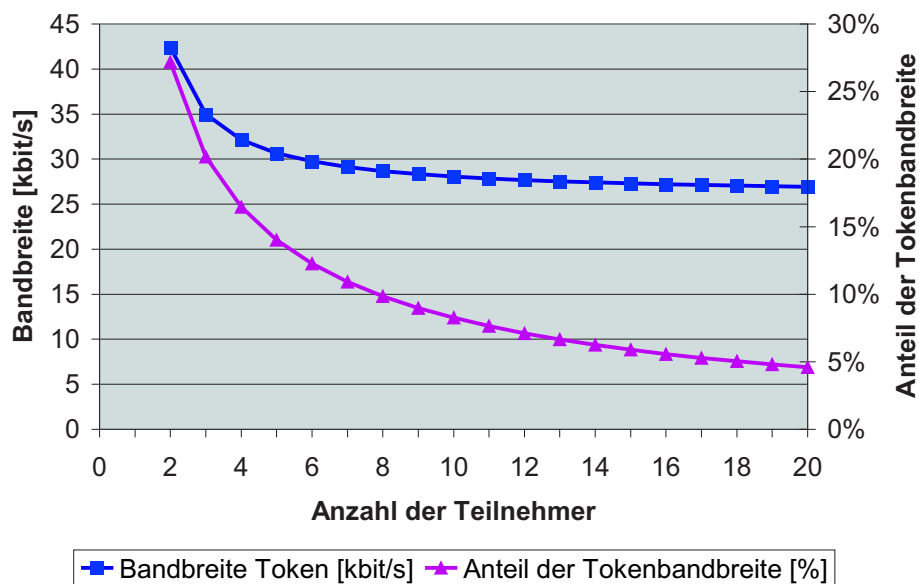


Abbildung 4.22.: Bandbreitenabschätzung für den Signalisierungsoverhead ohne Anpassung

In Abbildung 4.22 ist zu sehen, dass die benötigte Bandbreite für die Signalisierung bei zwei Teilnehmern sehr hoch ist und erst mit steigender Anzahl der Teilnehmer sinkt. Dies ist darauf zurückzuführen, dass die Tokenhaltezeit einen entscheidenden Einfluss auf die benutzte Bandbreite hat. Bei zwei Teilnehmern wird von einem Teilnehmer etwa alle 10 ms ein TF verschickt und ein TF\_ACK empfangen. Somit ist die Frequentierung eines Teilnehmers sehr hoch. Bei einer größeren Gruppe sind die Teilnehmer nicht so oft Tokenhalter und müssen demnach weniger häufig das Token weitergeben.

<sup>16</sup>Die Bandbreite der Videokonferenz bezieht sich auf folgende Randbedingungen: Videostrom mit 25 Bildern pro Sekunde bei einer Paketlänge von 3 kByte; Audiostrom mit 64 kbit pro Sekunde. Die Werte sind mit einer SUNVideo Karte gemessen worden.

In diesem Zusammenhang ist es sinnvoll, über eine Anpassung der Tokenhaltezeit in Abhängigkeit der Umlaufzeit und der Anzahl der Teilnehmer zu diskutieren. Um die Bandbreite zu reduzieren, besonders bei einer kleinen Anzahl von Teilnehmern, kann von folgender Annahme ausgegangen werden. Eine Tokenumlaufzeit von etwa 120 ms bewirkt einen Bandbreitenoverhead von etwa 5 bis 6 Prozent. Dies ist ein normaler Wert, der auch von dem Steuerprotokoll RTCP des *Realtime Transport Protocol (RTP)* [137] eingehalten wird. Dieser Wert wird bei etwa 17 Teilnehmern erreicht. Um schon bei einer

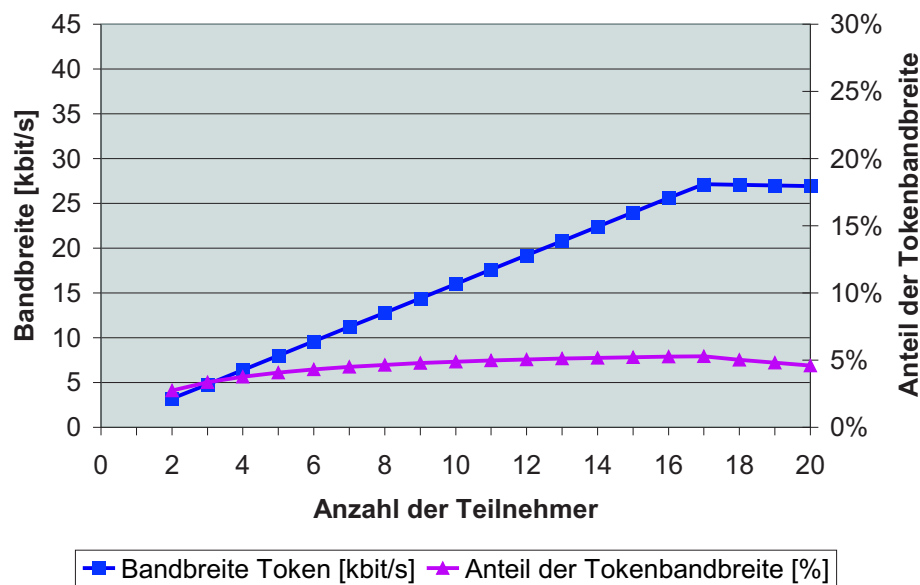


Abbildung 4.23.: Bandbreitenabschätzung für den Signalisierungs-overhead mit Anpassung

kleineren Gruppenstärke diesen Wert zu erreichen, muss die Tokenumlaufzeit vergrößert werden. Dies kann durch Verlängerung der Tokenhaltezeit  $t_{\text{Tokenhalt}}$  erreicht werden. Wenn die Tokenumlaufzeit kleiner als die diskutierten 120 ms ist, kann die Tokenhaltezeit folgendermaßen festgesetzt werden:

$$t_{\text{Tokenhalt, (neu)}} = \frac{120 \text{ ms} - t_{\text{RTT-Token}} + n \cdot t_{\text{Tokenhalt, (alt)}}}{n} \quad (4.24)$$

Das Ergebnis dieser Überlegungen ist in Abbildung 4.23 gezeigt. Es wird deutlich, dass der Overhead nicht über 6 Prozent ansteigt. Die dadurch entstehende

zusätzliche Verzögerung bei der Übertragung der Signalisierungsdaten ist akzeptabel, da die maximale Reaktionszeit noch immer unter 1000 ms liegt.

Zu dieser Anpassung ist zu bemerken, dass diese wahrscheinlich nur in lokalen Umgebungen zum Tragen kommt. In globalen Umgebungen wird die Tokenumlaufzeit durch die Leitungsverzögerungen automatisch steigen und somit den Overhead gering halten. Die Anpassung muss bei jeder Änderung der Gruppenzusammensetzung aufgehoben werden. Anschließend muss von neuem überprüft werden, ob eine erneute Anpassung notwendig ist oder nicht.

## 4.4. Signalisierung in hierarchischen Topologien

Die bisherigen Betrachtungen bezogen sich eher auf den Fall, dass sich die Teilnehmer einer Videokonferenz in einer logischen Ebene befinden, wie das beispielsweise bei BRAVIS der Fall ist. OCTOPUS verfolgt jedoch einen hierarchischen Ansatz, bei dem die Signalisierungsdaten über mehrere Ebenen geleitet werden müssen. Hierbei stellt sich das Problem, dass die Signalisierungsnachricht nicht nur durch die Leitung verzögert wird, sondern auch durch die zusätzlichen Schnittstellen zwischen den einzelnen Ebenen. Im Falle des OCTOPUS müssen das lokale Netz, der lokale Gruppenkoordinator, das globale Netz, der entfernte Gruppenkoordinator und das entfernte lokale Netz durchlaufen werden, bevor ein Paket einen Empfänger erreicht (siehe Abbildung 4.24). Der Gruppenkoordinator kann hierbei zu einem Engpass in der Kommunikation werden. Somit ist zu prüfen, ob die höhere Robustheit des hierarchischen Ansatzes den Nachteil der höheren Verzögerung gerechtfertigt. Dazu wird am Beispiel des OCTOPUS ein hierarchisches System in Bezug auf die Gesamtverzögerung der Signalisierungsdaten untersucht.

### 4.4.1. Signalisierungsprinzipien in OCTOPUS

Ein hierarchisches System, das wie OCTOPUS geschlossene Gruppen unterstützt, muss ebenfalls die Konsistenz der verteilten Managementdaten zwischen den entfernten Teilnehmern gewährleisten. Dazu ist es erforderlich, dass zunächst die oberste Ebene, also die globale Ebene die Daten zuverlässig, atomar und total geordnet verteilt. Erst dann dürfen die Daten der unteren Ebene

weitergeleitet werden. Das soll am Beispiel einer Floor-Anforderung erläutert werden (siehe Abbildung 4.24).

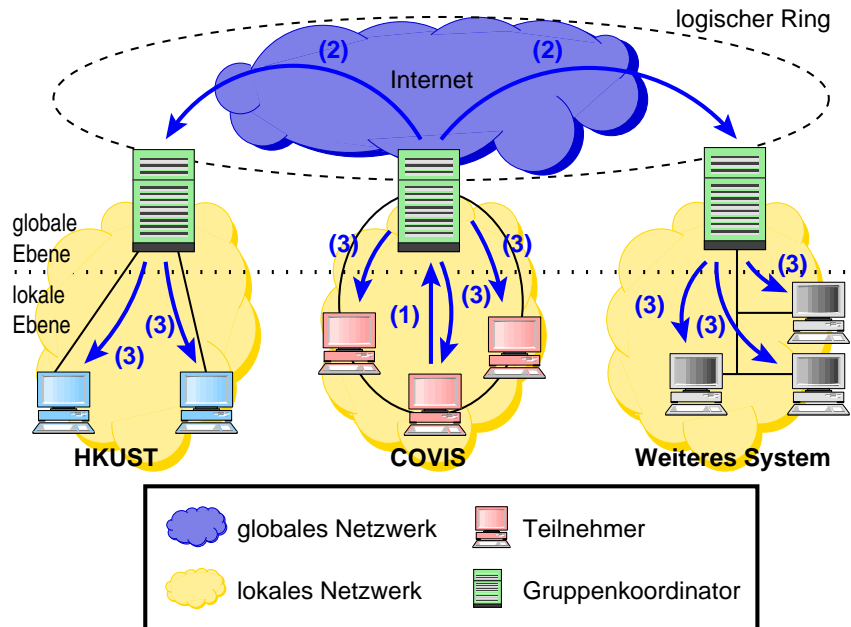


Abbildung 4.24.: Anforderung des Rederechts im OCTOPUS-System

Die Floor-Anfrage wird zunächst an den Gruppenkoordinator gesendet (siehe Abbildung 4.24, Schritt (1)). Im COVIS muss dazu die Floorkontrolle in der Weise geändert werden, dass der Request nicht an alle COVIS-Teilnehmer geschickt wird. In einem zentral organisierten System, wie beispielsweise dem Hongkonger System (HKUST), braucht hingegen nichts am Kommunikationsablauf geändert zu werden, wenn der Gruppenserver die Aufgaben des Gruppenkoordinators übernimmt. Nachdem der Gruppenkoordinator die Request-Nachricht empfangen hat, verteilt er diese an die entfernten Gruppenkoordinatoren (2). Danach kann die Nachricht an alle Teilnehmer der lokalen Gruppen übertragen werden (3).

Der ursprüngliche Kommunikationsablauf des COVIS musste wie beschrieben geändert werden, da bei einer gleichzeitigen Anforderung des Floors von Teilnehmern unterschiedlicher lokaler Gruppen zuerst die jeweilige lokale Gruppe die Floorkontrolle aktualisiert. Erst kurze Zeit später würde die Anforderung von der entfernten Gruppe eintreffen. Dies hätte eine unterschiedliche Reihenfolge der Daten in der Floorkontrolle in den beiden Gruppen zur Folge.



Die Konsequenz ist, dass zuerst die oberste Ebene eines hierarchischen Systems aktualisiert werden muss, bevor die tieferen Schichten aktualisiert werden können.

#### 4.4.2. Abschätzung der Signalisierungsverzögerung

In OCTOPUS muss eine Signalisierungsnachricht zunächst vom Sender an den Gruppenkoordinator gesandt werden, bevor sie global übertragen werden kann. Anschließend muss die Nachricht wieder vom Gruppenkoordinator an die eigene Gruppe übermittelt werden. Die lokale Verzögerung wird im Vergleich zu anderen lokalen Systemen voraussichtlich beim COVIS aufgrund der Ringtopologie am höchsten sein. Aus diesem Grund wird für die Abschätzung der Gesamtverzögerung die Anordnung benutzt, wie sie in Abbildung 4.25 dargestellt ist.

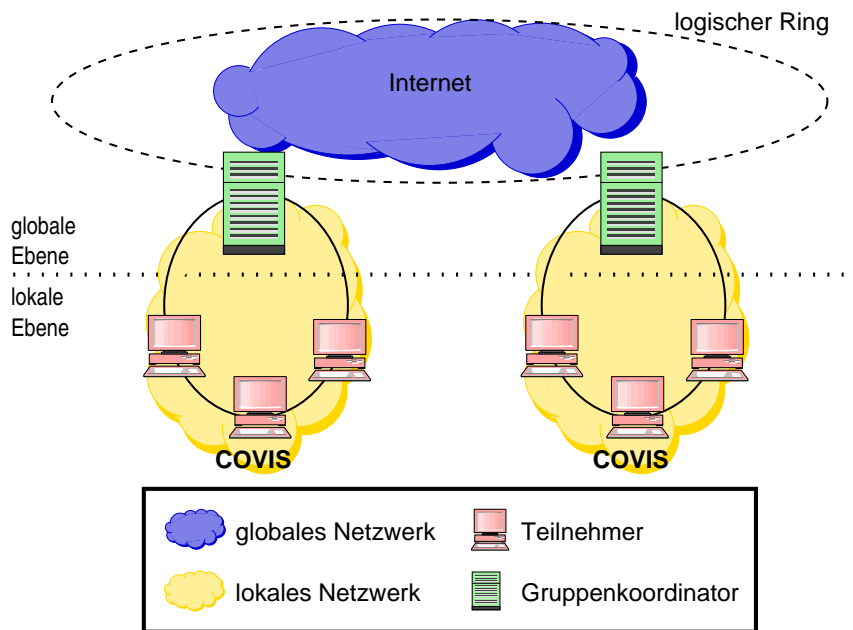


Abbildung 4.25.: Anordnung für die Abschätzung der Gesamtverzögerung der Signalisierung im OCTOPUS-System

Aus [10] sind die Tokenumlaufzeiten für das  $GCP_{ATM}$  bekannt, die in Tabelle 4.1 enthalten sind. Die Verzögerung des Token um jeweils 10 ms wurde hier eingeführt, um die Belastung des Netzwerks und der beteiligten Endsysteme zu reduzieren.

Anzahl der Teilnehmer	Tokenumlaufzeit in ms	
	ohne Verzögerung des Tokens	mit Verzögerung des Tokens
2	6,749	36,060
3	11,193	54,383
4	14,946	75,038
5	25,974	95,581

Tabelle 4.1.: Tokenumlaufzeiten des  $GCP_{ATM}$  nach [10]

Zunächst soll die maximale Verzögerung einer Nachricht für zwei Teilnehmer, die sich an unterschiedlichen Orten befinden, abgeschätzt werden. Im lokalen Bereich tauscht  $GCP_{ATM}$  sechs PDUs für eine Rotation des Token aus. Somit benötigt das Senden einer PDU unter Berücksichtigung der gemessenen Werte der Tabelle etwa 1,1 ms. Die maximale Verzögerung von einem Teilnehmer zu dem Gruppenkoordinator beträgt also etwa 8 ms<sup>17</sup>. Die globale Verzögerung zwischen den Gruppenkoordinatoren hängt, wie schon geschildert, von der Entfernung zwischen den Systemen ab. Unter der Annahme, dass sich die beiden Teilnehmer innerhalb von Europa befinden, kann die einfache Verzögerung zwischen den Standorten wie in Abschnitt 4.3.3 mit 20 ms veranschlagt werden. Mit der Gleichung 4.13 beträgt die globale Verzögerung also maximal 104 ms. Die Verzögerungen, die durch die Gruppenkoordinatoren auftreten, sind vergleichsweise gering. Messungen ergaben eine durchschnittliche Verzögerung von etwa 0,25 ms pro Gruppenkoordinator. Somit ergibt sich eine abgeschätzte Gesamtverzögerung von  $104 \text{ ms} + 2 \cdot 8 \text{ ms} = 120 \text{ ms}$  für eine Zweierkonferenz über OCTOPUS in Europa.

Vergleich mit  
dezentralem  
Ansatz

Unter denselben Voraussetzungen (lokale Verzögerung = 8 ms, globale Verzögerung = 20 ms) beträgt die Verzögerung in einem verteilten Ansatz wie dem BRAVIS-System 104 ms. Somit steigt die Verzögerung der Datenübertragung um nur 15 Prozent. Da der Hauptanteil der Verzögerung bei der Kommunikation zwischen den Gruppenkoordinatoren entsteht, hängt die Gesamtverzögerung im starken Maße von der Anzahl der beteiligten lokalen Videokonferenz-

<sup>17</sup>maximale Verzögerung ergibt sich wieder aus der Tokenumlaufzeit = 6,749 ms und der Datenübertragungszeit = 1,1 ms

systeme ab, die mittels OCTOPUS verbunden sind. Hingegen ist der Einfluss der Teilnehmerzahl, die sich in einer lokalen Konferenz befinden, eher gering.

Diese Abschätzung zeigt, dass der Ansatz einer hierarchischen Signalisierung akzeptabel und praktikabel ist. Trotzdem sollte aufgrund der doch geringfügig erhöhten Verzögerungen auf eine starke Strukturierung in Hierarchieebenen verzichtet werden. In OCTOPUS konnte dies jedoch nicht vermieden werden, da die lokalen Videokonferenzsysteme eine unterschiedliche Architektur besitzen und demzufolge mit unterschiedlichen Signalisierungsprotokollen kommunizieren. Somit ist der hierarchische Ansatz ein guter Kompromiss gegenüber einer Anpassung der lokalen Systeme an ein einheitliches Signalisierungsprotokoll.

## 4.5. Zusammenfassung

In diesem Kapitel wurde die Signalisierung für globale und geschlossene Videokonferenzsystemen untersucht. Es wurden die Anforderungen an das Transportsystem die Zuverlässigkeit, die totale Ordnungserhaltung und die Atomicität der Auslieferung der Daten für die Aufgaben die Signalisierung festgehalten. Vorhandene Protokolle wurden auf die Erfüllung dieser Anforderungen und den Einsatz in einer globalen, geschlossenen Konferenzumgebung geprüft und bewertet. Hierbei hat sich gezeigt, dass die vorhandenen Protokolle kaum die Anforderungen erfüllen können und zumeist auf einer offenen Multicaststruktur basieren. Sie sind für die Verwaltung geschlossener Gruppen ungeeignet. Aufgrund dieser Unzulänglichkeiten entstand der vorgestellte Ansatz des  $GCP_{IP}$ , der eine Kommunikation geschlossener Gruppen ermöglicht. Das auf einem logischen Ring und einem Tokenprinzip beruhende  $GCP_{IP}$ , wurde in diesem Kapitel durch eine Dienst- und eine Protokollspezifikation beschrieben. Hervorzuheben ist, dass  $GCP_{IP}$  nur auf Punkt-zu-Punkt-Verbindungen basiert. Es erfordert damit keine senderbasierte Multicaststruktur wie das  $GCP_{ATM}$  und kann so praktisch überall eingesetzt werden. Mit  $GCP_{IP}$  wurde somit eine Lösung geschaffen, geschlossene Gruppen im Internet dezentral zu organisieren. Eine fundierte Bewertung und Diskussion der Lösung belegen, dass das  $GCP_{IP}$  für eine Kommunikation von bis zu 30 Teilnehmern in BRAVIS beziehungsweise Gruppenkoordinatoren im OCTOPUS einsetzbar ist. Der hierarchische Signalisierungsansatz des OCTOPUS-Videokonferenzsystems mittels

#### *4. Sicherung der Geschlossenheit der Gruppe in globalen Videokonferenzsystemen*

---

GCP<sub>IP</sub> ist, wie Abschätzungen bestätigt haben, ebenso praktisch einsetzbar. Aufgrund von erhöhten Verzögerungen sollte jedoch bei der Entwicklungen von eigenständigen Videokonferenzlösungen eine hierarchische Signalisierung vermieden werden.

---

## 5. Übertragung multimedialer Daten in globalen Videokonferenzsystemen

---

Die Übertragung der Medien Audio und Video in Videokonferenzen stellt besondere Anforderungen an die Qualität der Verbindung zwischen den Kommunikationspartnern. Die Qualität, die der menschliche Nutzer wahrnimmt, wird vor allem durch die Verzögerungen, die Verzögerungsschwankungen, die Verluste und die Bandbreite des Übertragungskanals bestimmt. In lokalen Netzen, in ATM-Netzwerken und im ISDN können diese Qualitätsanforderungen weitgehend erfüllt werden. Problematisch ist dies jedoch in globalen öffentlichen Netzen wie dem Internet, in dem keine Übertragungsqualitäten garantiert werden können. Ein weiteres Problem stellt die Wahrung der Geschlossenheit einer Konferenzgruppe dar, da durch die Verteilung der Mediendaten nur eine minimale Belastung des Netzwerkes auftreten sollte. In diesem Kapitel werden diese Probleme für die Audio- und Videoübertragung in globalen, geschlossenen Videokonferenzsystemen für das Internet diskutiert und eigene Konzepte für die Lösung hergeleitet, die im OCTOPUS- und BRAVIS-System angewendet wurden. Eine Bewertung der Lösung bildet den Abschluss des Kapitels.

### 5.1. Medientransport in Videokonferenzsystemen

#### 5.1.1. Anforderungen

Die Qualität der Kommunikation in Videokonferenzsystemen wird im Wesentlichen von vier Größen beeinflusst: die verfügbare Bandbreite, die Verzögerung, die Verzögerungsschwankung (Jitter) und die Übertragungsverlustrate. Ein interaktiver Informationsaustausch zwischen den Kommunikationspartnern erfordert geringe Verzögerungen, die für die Sprachkommunikation nicht mehr

als 400 ms pro Richtung betragen sollten [77]. Für den Informationsaustausch werden Audio- und Videodaten zwischen den Teilnehmern übertragen, die, obwohl sie stark komprimiert werden, eine relativ hohe Bandbreite über einen längeren Zeitraum benötigen. Audiodaten haben typischerweise Bandbreiten bis zu 64 kbit/s, Videodaten sogar bis zu 2 Mbit/s. Audio- und Videodaten sind kontinuierliche Medien, die beim Sender in einem festen Zeitintervall aufgenommen werden und in einem festen Zeitintervall beim Empfänger abgespielt werden müssen. Verzögerungsschwankungen in Audio- und Videodaten wirken sich oft störend für den Nutzer aus [120, 145]. Übertragungsverluste können die Kommunikation zwischen den Konferenzteilnehmern stören, wenn sie bei einem Audiostrom über etwa 6% und bei einem Videostrom über etwa 51% ansteigen [114].

### ATM und ISDN

ATM- und ISDN-Netzwerke bieten qualitätssichernde Dienste, die eine maximale Verzögerung, einen geringen Jitter und eine konstante Bandbreite garantieren können. Übertragungsfehler treten hier kaum auf. Jedoch sind diese Netze für die globale Kommunikation mit Videokonferenzen zu teuer, und die Konnektivität erreicht bei weitem nicht die des Internets.

### Internet

Das Internet hingegen arbeitet bisher nach dem *Best Effort* Prinzip, ohne jegliche Gewährleistungen von Dienstgütegarantien. In lokalen IP-Netzen können jedoch die genannten Anforderung relativ leicht erfüllt werden, weil die räumliche Ausdehnung gering ist und kaum Zwischensysteme, wie Router, die Kommunikation beeinflussen. Somit halten sich Verzögerungen und Jitter in geringen Maßen. Im Allgemeinen werden auch hohe Bandbreiten zur Verfügung gestellt, da die Anzahl der Nutzer beschränkt ist.

In globalen öffentlichen IP-Netzen durchläuft ein Datenpaket in der Regel mehrere Netzwerkknoten (Routern), um von der Quelle zum Ziel zu gelangen. In den Routern werden die Datenpakete in einer Warteschlange nach dem FIFO-Prinzip zwischengespeichert, bevor sie weiter zum Ziel vermittelt werden. Ist die Warteschlange leer, wird ein eintreffendes Paket sofort und ohne Verzögerungen vermittelt. Da die Verarbeitungszeit des Routers zwar hoch, aber dennoch begrenzt ist, kann die Warteschlange zum Teil gefüllt sein, wodurch Verzögerungen im Netzknoten auftreten. Treffen in einen Router gleichzeitig so viele Pakete ein, dass die Warteschlange vollständig gefüllt ist, werden weiter eintreffende Pakete verworfen. Die Verluste machen sich beim Abspie-

len eines übertragenen Video- oder Audiostroms in Form von Rauschen des Bildes, im Stocken des Bildlaufs oder durch Knistern und Knacken im Ton störend bemerkbar. Eine weitere Folge von Staus sind starke Verzögerungen, die durch das Anwachsen der Paketwarteschlangen in den Routern hervorgerufen werden. Demnach ist mit diesem einfachen Prinzip der Paketvermittlung eine Sicherung der Dienstgüte im Internet kaum möglich. Mittels verschiedener Ansätze wie den Integrated Services [21], den Differentiated Services [15, 56, 80] oder dem MPLS [48] wird versucht, nachträglich eine Dienstgütesicherung in das IP-Protokoll zu integrieren. Wie jedoch schon erwähnt, ist die breite Anwendung solcher Dienste in globalen öffentlichen Netzwerken kurzfristig nicht zu erwarten. Um trotzdem Verzögerungen, Jitter und Verluste zu minimieren, müssen folglich Staus in den Netzwerkknoten vermieden werden. Mit einer Anpassung der Sendeleistung von Video- oder Audioströmen an die verfügbaren Ressourcen kann dies erreicht werden. Dadurch wird gesichert, dass die Übertragungsqualität bei den Empfängern so hoch ist, wie Verbindung maximal übertragen kann.

Stausituation können durch die Empfänger anhand von subjektiven Einschätzungen der Videoqualität festgestellt werden. Durch Messungen der empfangenen Bandbreite, der Verlustrate und der Verzögerung ist eine objektivere Beurteilung möglich. Das RTP-Protokoll [137] unterstützt solche Messungen durch Sequenznummern und Zeitstempel im Paketkopf. RTP enthält aber keine Bandbreitenzuweisungs- und Überlastkontrollmechanismen.

Stauerkennung
---------------

In Mehrteilnehmerkonferenzen, besonders in heterogenen Umgebungen wie dem Internet, stellt die Anpassung der Sendeleistung an die verfügbare Bandbreite ein zentrales Problem dar, weil *jedes* Gruppenmitglied eine andere Bandbreite als der Sender haben kann. Die empfangene Dienstqualität, die dem Sender mitteilt wird, kann von Empfänger zu Empfänger sehr unterschiedlich sein. Der Sender muss dann entscheiden, welche Teilnehmer in der Anpassung der Sendeleistung zu berücksichtigen sind. Zur Lösung dieses Problems können vier Möglichkeiten in Betracht gezogen werden:

1. Anpassung an den Teilnehmer mit der höchsten Bandbreite,
2. Anpassung an den Teilnehmer mit der geringsten Bandbreite,

3. Anpassung an den Teilnehmer mit einer durchschnittlichen Bandbreite oder
4. Nutzung von *Layered-Coding*-Techniken.

Die erste Möglichkeit hat den Nachteil, dass nur eine Minderheit der Teilnehmer einer Sitzung qualitativ hochwertige Datenströme empfängt. Die meisten Teilnehmer empfangen jedoch eine schlechtere Qualität, weil die geringere Bandbreite ihrer Verbindung zu Überlastungen führt. Das kann für einige Nutzer inakzeptabel sein, da die angezeigten Bilder und die abgespielten Audioströme viele Fehler enthalten, die sich in Rauschen, Knacken und Aussetzern bemerkbar machen. Des Weiteren können sich Stausituationen von leistungsschwachen Verbindungen auch auf leistungsstarke Verbindungen auswirken.

Der zweite Fall wurde im Anpassungsmechanismus von USMINT realisiert, bei dem sich nach dem Teilnehmer mit den geringsten Ressourcen gerichtet wird. Die restlichen Teilnehmer bleiben weitgehend unberücksichtigt [141]. Dies hat zur Folge, dass einzelne Teilnehmer mit einer schlechten Netzwerkverbindung die Qualität der gesamten Videokonferenz senken können.

Für das RENDEZVOUS-Vorgängersystem IVS wurde ein Regulierungsansatz entwickelt, der drei Verbindungszustände unterscheidet: Überlast ( $\geq 5\%$  Verlust), Ausgelastet (5 bis 2 % Verlust) und Unbelastet ( $\leq 2\%$  Verlust) [19]. Die Sendeleistung wird bei Überlast reduziert, bei der Auslastung konstant gehalten und bei keiner Belastung erhöht. Bei einer Mehrteilnehmerkonferenz wird die Sendeleistung verringert, wenn 10% der Gruppe überlastet sind. Sie wird jedoch auch dann angepasst, wenn 90% der Verbindungen zu den Teilnehmern ausgelastet beziehungsweise unbelastet sind. Damit empfängt die Mehrheit der Teilnehmer eine nicht zufrieden stellende Qualität. Andererseits wird die Senderate erhöht, wenn 10% der Gruppenteilnehmer keine Stausituationen (Unterlast) entdecken, was wiederum zu einer Überlastung der verbleibenden 90% der Teilnehmer nach der Anpassung führen kann. In heterogenen Netzwerkkumgebungen wie dem Internet ist die gegenseitige Abhängigkeit der Bandbreite der einzelnen Teilnehmer offensichtlicher als in homogenen Umgebungen. Dies wird deutlich, wenn 10% der Teilnehmer einer Gruppe eine schwache und 10% eine sehr gute Netzanbindung haben. Beide Teilgruppen versuchen nun die Anpassung der Sendeleistung in entgegengesetzte Richtungen, was in keinen stabilen Zustand konvergiert.



RENDEZVOUS nutzt weiterhin *Layered Coding* in Verbindung mit dem *Layered-Multicast*-Konzept [6, 131], um heterogene Systemumgebungen zu unterstützen. Vom Sender werden verschiedene Qualitätsstufen über jeweils eine Multicastadresse verteilt. Jedoch muss der Empfänger manuell einstellen, welche Ebenen er für eine gute Anzeigequalität empfangen will. Dies erfordert Erfahrungen und Geschick seitens des Nutzers. Mit *Receiver-Driven Layered Multicast (RLM)* [100] existiert ein automatischer Mechanismus auf der Basis von IP-Multicast, der zu Testzwecken in das Mbone-Werkzeug VIC [112] integriert wurde. SCUBA [5], eine Erweiterung von *RLM*, führt einen zusätzlichen Auswahlmechanismus ein, der verschiedene Qualitätskanäle anbietet. Alle Sender teilen sich eine Untermenge dieser Kanäle, abhängig von der Abstimmung der Empfänger. Die Empfänger können dem gewünschten Kanal beitreten. Jedoch verursachen *RLM* und SCUBA starke Überlastungen während der Anpassung an die verfügbaren Ressourcen. Deshalb wurden sie in RENDEZVOUS vermutlich nicht integriert.

In offenen Konferenzgruppen ist eine Anpassung der Sendeleistung besonders schwierig, da der Sender die Zusammensetzung der Gruppe nicht genau kennt. So kann die akzeptable Bandbreite nur aufgrund der Informationen der bekannten Teilnehmer abgeschätzt werden. Des Weiteren kann die Qualität der Konferenz durch das beabsichtigte oder unbeabsichtigte Senden falscher Messergebnisse an den Sender reduziert werden.

In einer geschlossenen Konferenzgruppe hat jeder Teilnehmer einen exakten Überblick über die Anforderungen der anderen Teilnehmer, da er den QoS-Zustand aller Teilnehmer kennt. Die Möglichkeit einer Störung durch Außenstehende ist hier kaum gegeben. Somit kann jedes Gruppenmitglied die Sendeleistung für jeden einzelnen Teilnehmer individuell anpassen, um Überlastsituationen zu vermeiden. Jedoch steigt damit der Verwaltungs- und Bearbeitungsaufwand für jeden einzelnen Teilnehmer, wodurch ein Kompromiss zwischen Individualität und Aufwand notwendig wird.

Somit muss die Medienübertragung in Videokonferenzsystemen derart gestaltet werden, dass Stausituationen in heterogenen Umgebungen vermieden werden. Im folgenden Abschnitt werden dazu existierende Mechanismen und Systeme analysiert.

Vergleich offene und geschlossene Gruppen
---

### 5.1.2. Überlastkontrolle in existierenden Systemen

manuelle Anpassung

Die Mbone-Tools VIC, VAT und RAT [112, 113, 157] sowie das Videokonferenzsystem DAVIKO [34] sind sehr einfach gestaltete Videokonferenzsysteme. Sie enthalten keine automatischen Mechanismen zur Vermeidung von Stausituationen. Stattdessen müssen die Nutzer ständig die Verlustrate und die Verzögerungen der Pakete überwachen. Treten zu hohe Verluste auf, muss die Sendeleistung reduziert werden. Diese manuelle Anpassung erfordert Kenntnisse der Nutzer über die gegenseitigen Abhängigkeiten zwischen Sendeleistung, Übertragungsqualität und Verhalten des Netzwerks. Die ständige Kontrolle der Übertragungsqualität lenkt außerdem vom eigentlichen Zweck einer Videokonferenz, dem Informationsaustausch, ab. In heterogenen Umgebungen muss sich der Teilnehmer in einer Mehrteilnehmerkonferenz selbst entscheiden, ob leistungsstarke Systeme eine geringe Dienstgüte erhalten oder leistungsschwache Systeme überlastet werden sollen.

H.323-Systeme enthalten kaum Mechanismen zur dynamischen Regulierung der Sendeleistung, um Stausituationen im Netzwerk vermeiden beziehungsweise erkennen und beheben zu können. Stattdessen muss der Nutzer während des Beitritts angeben, welche Datenrate verarbeitet werden kann. Dazu sind Kenntnisse des Nutzers über vergangene Sitzungen und deren Übertragungsqualität sowie über den aktuellen Zustand des Netzes erforderlich, was für einen gewöhnlichen Nutzer kaum möglich ist.

automatische Mechanismen

Die H.323-Systeme der Firma VCON enthalten mit *PacketAssist<sup>TM</sup>* [33, 166] eine Regulierungsmöglichkeit, mit deren Hilfe die Sendeleistung der Videoströme in Abhängigkeit von Engpässen im Netzwerk dynamisch anpasst werden kann. Ein ähnliches Konzept verfolgt der Ansatz *Dynamic Bandwidth Allocation (DBA)* der Firma POLYCOM [116, 122]. Automatische Überlastmechanismen sind außerdem in IVS [18, 19] und USMINT [141] enthalten. Sie arbeiten nach ähnlichen Prinzipien. Wird eine Überlastsituation aufgrund von Verlusten festgestellt, wird die Senderate reduziert. In RENDEZVOUS wird sie halbiert. USMINT reduziert die Senderate auf den Wert, bei dem die Verlustrate noch akzeptabel ist. Dieser Wert wird durch den Sender des Videostroms festgelegt, der wissen muss, bei welcher Verlustrate eine noch zufrieden stellende Qualität gewährleistet wird.

### 5.1.3. Fehlerbehandlung

Trotz einer Überlastkontrolle können beim Transport Audio- und Videodaten eines Videokonferenzsystems vom Sender zu den Empfängern verloren gehen, verspätet ankommen oder verfälscht werden. Die Folge sind Rauschen oder Flackern des Bildes beziehungsweise Rauschen und Knacken des Tons, wodurch die Kommunikation zwischen den Teilnehmern erheblich gestört werden kann. In IP-Netzen ist die Wahrscheinlichkeit eines Verlustes höher als in ISDN- oder ATM-Netzen, da IP nur einen Best Effort Dienst zur Verfügung stellt. Um Übertragungsfehler zu behandeln, wurden bisher drei Ansätze genutzt: die Neuübertragung, die *Forward Error Correction (FEC)* und die Approximation.

Zusätzlich zu den Standardprotokollen der Transportschicht des TCP/IP Protokollstapels TCP und UDP wurden Ansätze wie *Heterogenous Packet Flow (HPF)* [41] oder *Partially Reliable Transport Protocol using Explicit Congestion Notification (PRTP-ENC)* [49, 50] entwickelt, die die Transportschicht um einen *teilweise zuverlässigen* Transportdienst ergänzen. Diese Lösungen dienen zur relativ zuverlässigen Übertragung von kontinuierlichen Datenströmen in Echtzeit. Im Gegensatz zu TCP fordern die Empfänger relativ zeitig eine Neuübertragung von verloren oder fehlerhaften Datenpaketen vom Sender an, um zu gewährleisten, dass das erneut übertragenen Paket noch rechtzeitig eintrifft, um abgespielt zu werden. Trifft das Paket nicht rechtzeitig ein, muss der Audio- oder Videostrom fehlerhaft abgespielt werden. Diese Protokolle erzeugen aber durch die Neuübertragung eine höhere Netzbelastung und können folglich Stau- und Fehlersituationen noch verschlimmern. Zudem wurden sie bisher nur in lokalen Umgebungen getestet, in denen die Timeout-Intervalle für die Neuübertragung gut dimensioniert werden können. In Mehrteilnehmerkonferenzen wurden die genannten Ansätze noch nicht eingesetzt. Es lässt sich jedoch vermuten, dass auch hier die wiederholte Übertragung die Netzsituation verschlechtert.

Eine andere Möglichkeit, die Verlustrate gering zu halten und somit eine hohe Darstellungsqualität zu gewährleisten, ist die Integration einer *Forward Error Correction (FEC)* [14]. FEC wird beispielsweise im Videokonferenzwerkzeug *RAT* [157] eingesetzt, um Informationen aus verloren gegangenen Audiopaketen wieder herstellen zu können. Wie in Abbildung 5.1 dargestellt, werden dazu redundante Informationen in dem Paketstrom mitge-

Neuübertragung
----------------

FEC
-----

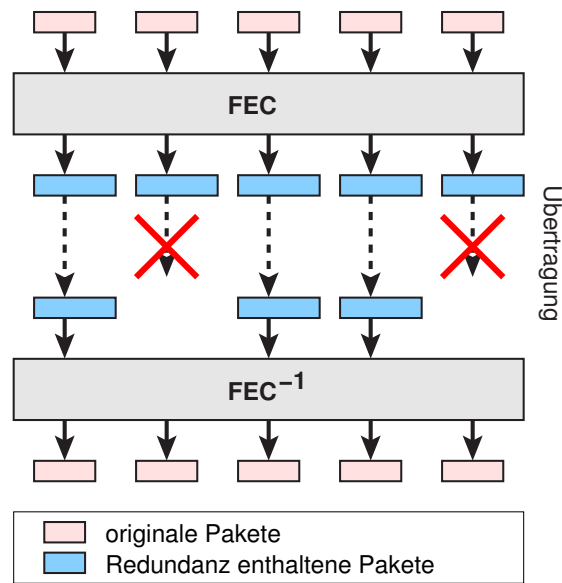


Abbildung 5.1.: Prinzip der Forward Error Correction

schickt. Der Vorteil liegt darin, dass Audioströme trotz Paketverlusten über fehleranfällige Leitungen mit einer gleich bleibend guten Qualität abgespielt werden können. Die Übertragung der FEC-Informationen benötigt jedoch bis zu 30% an zusätzlichen Ressourcen [115, 119]. Weiterhin treten Verzögerungen für die FEC-Kodierung und -Dekodierung auf. Bei einer ungünstigen Wahl der FEC-Parameter in sehr fehleranfälligen Umgebungen kann es dennoch passieren, dass beim Abspielen der Mediendaten Fehler durch nicht wiederherstellbare Informationen auftreten können.

*Approximation*

Die dritte Möglichkeit, Übertragungsfehler zu beheben, nutzt die Eigenschaft der Audio- und Videosignale aus, dass sie sich in einem kurzen Zeitraum oft nur geringfügig ändern. Geht ein Datenpaket eines Audio- oder Videostroms verloren, kann dieses durch das vorangegangene oder nachfolgende Paket ersetzt werden. Dem Nutzer fällt dieser Unterschied zum originalen Signal kaum auf. Etwas komplexer und aufwendiger gestaltet sich die Berechnung des fehlenden Pakets durch das vorangegangene und nachfolgende Paket. Dazu müssen jedoch die Daten gepuffert werden, das höhere Verzögerungen zur Folge hat.

Bis auf die FEC wurde bisher keine Fehlerbehandlung in einem der zur Zeit existierenden Videokonferenzsysteme integriert. Der Grund liegt wohl an dem Mehraufwand und den höheren Verzögerungen, den diese Mechanismen

bewirken. Da Nutzer bei der Videoübertragung in einem Videokonferenzsystem eine Ausfallrate von bis zu 51 % noch tolerieren, scheint das Nutzen–Kosten–Verhältnis sehr gering zu sein.

#### 5.1.4. Topologien

Die Topologie für die Übertragung der Audio– und Videodaten in einem Videokonferenzsystem für geschlossene Gruppen hat einen entscheidenden Einfluss auf dessen Qualität. Wie in Kapitel 2 diskutiert, existieren im Wesentlichen drei Ansätze: zentral, dezentral und hierarchisch [129].

Zentralisierte Systeme, wie beispielsweise im H.323–Standard vorgesehen, sind recht einfach zu installieren und zu warten. Die zentralen Gruppenserver oder MCUs, die diese Systeme enthalten, können jedoch zum Flaschenhals der Kommunikation sowie zum Single–Point–of–Failure werden. Des Weiteren sind die Nutzer eines zentralen Videokonferenzsystems von dem Anbieter des Servers und dessen Standort abhängig. Der Standort spielt unter Berücksichtigung der Dienstqualität eine sehr wichtige Rolle. Liegt der Server weit von den Nutzern entfernt, erhöhen sich die Verzögerungen, und es kann zu vermehrten Bandbreitenengpässen kommen. Ein Videokonferenzserver in einem stark belasteten Netzwerk oder mit einer schlechten Netzwerkanbindung kann weiterhin zu einer stark eingeschränkten Qualität für alle Teilnehmer führen. Ein wichtiger Aspekt ist auch die Kostenfrage. Der Server empfängt und verteilt alle Daten der Konferenz. Für die Institution, bei der der Videokonferenzserver steht, entstehen deshalb deutlich höhere Kosten, weil sie für die transportierte Datenmenge entsprechend beim Netzwerkprovider bezahlen muss. Die Kosten für die Teilnehmer, die eine Konferenz von außerhalb über diesen Server aufbauen, sind hingegen gering.

Zentrale Ansätze

Relativ einfach sind bei H.323–Systemen Mechanismen zur dynamischen Anpassung an die verfügbaren Ressourcen zu integrieren, da zwischen dem Server beziehungsweise der MCU eine Punkt–zu–Punkt–Verbindung zu den Teilnehmern besteht. Jedoch wurde dies bisher nur durch die Produkte der Firma VCON mit der Erweiterung *PacketAssist*<sup>TM</sup> umgesetzt. Eine Software–basierte Lösung einer MCU benötigt weiterhin eine hohe Rechenkapazität, um die individuell angepasste Kompression der Mediendaten zu bewerkstelligen.

Dezentralisierte Systeme versprechen minimale Verzögerungen zwischen den

Dezentrale Ansätze

Teilnehmern, da diese direkt miteinander verbunden werden. Dadurch ist es bei einer total vermaschten Topologie sogar möglich, die Übertragungsqualität für jeden Teilnehmer individuell anzupassen. Dies erfordert jedoch einen höheren Bandbreitenbedarf. Die Nutzung von Multicastdiensten, wie in den Mbone-Tools oder COVIS, schont zwar die benötigten Bandbreitenressourcen, jedoch sind Multicastdienste nicht überwiegend vorhanden. Weiterhin ist eine individuelle Anpassung der Übertragungsqualität nicht möglich. In dezentral organisierten Konferenzen wirken sich Fehler einzelner Teilnehmer nicht auf die gesamte Konferenz aus. Flaschenhals der Kommunikation kann hier nur das Netzwerk werden, da andere Komponenten das Konferenzgeschehen nicht beeinflussen. Die Kostenfrage ist ebenfalls geklärt, da alle Teilnehmer die Kosten für die von ihnen gesendeten und empfangenen Daten tragen. Voraussetzung für den verteilten Ansatz ist das Vorhandensein eines Signalisierungsmechanismus, der sichert, dass die Medienströme nur zu den Teilnehmern der Konferenz gelangen. In IP-Multicast ist das beispielsweise nicht gegeben.

### Hierarchische Ansätze

Hierarchische Systeme können geschlossene Gruppen unterstützen und bieten die Vorteile von verteilt und von zentral organisierten Ansätzen. Teilnehmer, die etwa gleiche Qualitätsanforderungen besitzen, können in eine Untergruppe beziehungsweise Hierarchieebene eingeordnet werden. Dadurch kann die Übertragungsqualität an die Bedürfnisse dieser Untergruppe angepasst werden. Gleichzeitig wird auch Bandbreite eingespart, da der Audio- oder Videostrom zwischen den Untergruppen nur einmal übertragen werden muss. Die Verzögerung der Datenübertragung kann zwar leicht ansteigen, da die Verteilung stufenweise über die einzelnen Hierarchieebenen erfolgt (siehe auch Abbildung 2.19). Durch die heute verfügbare leistungsstarke Rechentechnik ist die Verzögerung aber kaum von anderen Ansätzen unterscheidbar. Hierarchische Ansätze zur Medienverteilung sind jedoch kaum auf dem Markt zu finden. Einzig die H.32x Standards lassen eine Kaskadierung der MCUs zu, wodurch mehrere Hierarchieebenen eingerichtet werden können. Jedoch stellt die MCU der obersten Ebene wiederum einen zentralen Punkt der Kommunikation mit den bereits diskutierten Nachteilen dar.

### 5.1.5. Fazit

Die Betrachtung der gegenwärtigen Möglichkeiten der Übertragung von Audio- und Videodaten in einem Mehrteilnehmer-Videokonferenzsystem für geschlossene Gruppen zeigt, dass heterogene Systemumgebungen wie das Internet nur unzureichend unterstützt werden. Dadurch kann nicht für alle Teilnehmer einer Konferenz eine zufrieden stellende Empfangsqualität gewährleistet werden. Hierarchische Ansätze bilden einen guten Kompromiss zwischen einer individuellen Anpassung an die verfügbaren Ressourcen jedes Teilnehmers einer Konferenz und der Vermeidung eines Engpasses in der Kommunikation. Jedoch wurden diese Ansätze bisher kaum umgesetzt. Anhand der Videokonferenzsysteme OCTOPUS und BRAVIS werden in den folgenden Abschnitten Möglichkeiten zur hierarchischen Verteilung der Medienströme untersucht.

## 5.2. Medienübertragung in OCTOPUS und BRAVIS

### Die OCTOPUS-Topologie

Die Idee von OCTOPUS was es, lokale Systeme ohne tief greifende Modifikationen an der Systemarchitektur über globale Entfernungen zu verbinden (siehe Kapitel 3.1). Dazu wurde eine Schnittstelle für jedes lokale Videokonferenzsystem definiert, die den globalen Datenaustausch realisiert. Diese Schnittstelle ist der Gruppenkoordinator. Mit diesem Ansatz wird quasi auch die hierarchische Topologie der Medienübertragung festgelegt. Da sich in der untersten lokalen Ebene verschiedenartige Videokonferenzsysteme befinden können, ist OCTOPUS ein spezieller Fall einer hierarchischen Medienverteilung, mit dem auch das Verhalten unterschiedlicher lokaler Verteilungsstrategien untersucht werden kann.

Die Audio- und Videodaten werden über den Gruppenkoordinator an die entfernten Konferenzgruppen verteilt. Da senderbasierte Multicastdienste nur unzureichend im Internet unterstützt werden, nutzt OCTOPUS Unicastverbindungen, um die Geschlossenheit der Gruppe zu gewährleisten. Im Gegensatz zur Signalisierung werden die Gruppenkoordinatoren nicht in einer Ringtopologie, sondern in einer total vermaschten Topologie mit Punkt-zu-Punkt-

*globale Medienverteilung*

Verbindungen angeordnet. Damit können die Verzögerungen zwischen den Gruppenkoordinatoren so gering wie möglich gehalten werden und die Sendeleistung der Gruppenkoordinatoren an die verfügbaren Ressourcen individuell angepasst werden (siehe Abschnitt 5.4.2). Diese Topologie benötigt hohe Bandbreiten. Da mit OCTOPUS nur maximal fünf bis sechs Gruppen gleichzeitig verbunden werden sollen und ein Bandbreitenmanagement vorhanden ist, wird jedoch eine übermäßige Beanspruchung der Bandbreitenressourcen vermieden.

<i>Overlay Multicast</i>
------------------------------

Die Entwicklungen im Bereich der *Overlay-Multicast*-Protokolle seit dem Jahr 2000 wurden auch für OCTOPUS interessant. Das Protokoll NICE [9] verfolgt ebenfalls eine hierarchische Struktur zur Verteilung der Daten an mehrere Teilnehmer. NICE setzt jedoch in der untersten Hierarchieebene eine gleiche Architektur der Endsysteme voraus. Die Ansätze NICE, SCRIBE, CAN, HMTP oder BAYEUX [9, 24, 130, 179, 182] unterstützen allerdings von sich aus keine heterogenen Netzwerkumgebungen. Somit muss die Sendeleistung an das leistungsschwächste System angepasst werden. Außerdem unterstützen die genannten Protokolle keine geschlossenen Gruppen, da Teilnehmer, ähnlich wie bei IP-Multicast, ohne eine Einladung einer laufenden Konferenz beitreten können. Dies entspricht jedoch nicht dem Konzept von OCTOPUS, wodurch auf eine nachträgliche Integration eines vorhandenen *Overlay-Multicast*-Ansatzes verzichtet wurde.

### Die BRAVIS-Topologie

BRAVIS wurde von Anfang an als eigenständige Lösung für ein dezentral organisiertes Videokonferenzsystem für geschlossene Gruppen entworfen. Dabei wurde anfänglich davon ausgegangen, dass sich die Gruppe einer Konferenz aus Untergruppen von verschiedenen Standorten zusammensetzt, ähnlich dem OCTOPUS-Ansatz (siehe Abbildung 3.1). Deshalb war eine Idee von BRAVIS, einen funktional reduzierten Gruppenkoordinator als Gateway zwischen den Standorten einzusetzen. Das Gateway wäre dann für die globale Verteilung der Audio- und Videodaten zwischen den global verteilten Standorten verantwortlich gewesen. Dies hätte jedoch den Nachteil, dass dieses Gateway an jedem Standort verfügbar sein muss und weiterhin angesteuert werden muss. Zudem könnte solch ein Gateway zu einem Flaschenhals in der Kommunikation werden. Deshalb wurde auf den Entwurf des Gateways verzichtet und stattdessen die Integration der Gatewayfunktionen in die Endsysteme vorgeschlagen.



Dieser Ansatz hat nun den entscheidenden Vorteil, dass das BRAVIS-Konzept unabhängig von der Netzinfrastruktur ist und somit nahezu überall eingesetzt werden kann.

Wie in Abbildung 3.9 schon am Beispiel eines Teilnehmers dargestellt, sendet jedes Endsystem die eigenen Audio- und Videodaten zu einem *Ausgewählten Empfänger* (AE) einer Untergruppe vom entfernten Standort. Jede Untergruppe enthält einen AE, der die eintreffenden Medienströme an die lokalen Teilnehmer verteilt. Der AE arbeitet dann bis auf Konvertierungsfunktionen quasi wie der Gruppenkoordinator des OCTOPUS-Systems. Jedes Endsystem enthält gemäß dem Peer-to-Peer Konzept die Funktionen eines AEs, die zur Verteilung der Medienströme von entfernten Teilnehmern an die lokalen Teilnehmer benötigt werden. Die Funktionen eines AEs werden jedoch nur bei einem Teilnehmer einer Untergruppe aktiviert, der durch das in BRAVIS integrierte QoS-Management ausgewählt wurde. Die Medienverteilung ist mit diesem Ansatz nach einem zweistufigen hierarchischen Schema aufgebaut, wobei die AE als Schnittstelle zwischen den zwei Ebenen dienen.

globale Medien-  
verteilung

Der Einsatz von *Overlay-Multicast*-Ansätzen wäre in BRAVIS möglich. Jedoch müssten existierende Protokolle wie BAYEUX, SCRIBE, NICE, HMTTP oder CAN [182, 24, 9, 179, 130] dahingehend geändert werden, dass sie geschlossene Gruppen unterstützen. Dies würde jedoch grundlegende Änderungen der Protokolle bedeuten, da auf einen *Rendezvous Point* verzichtet und stattdessen eine Einladungsprozedur integriert werden muss. Des Weiteren kann der Aufwand für das Erstellen der Multicastbäume auf der Anwendungsebene recht hoch werden. Aus diesen Gründen wurde auf die Nutzung von vorhandenen *Overlay-Multicast*-Lösungen verzichtet und die eigene, relativ einfache, zweistufige Verteilung der Audio- und Videostreams genutzt.

Overlay Multi-  
cast

## 5.3. Ein Ansatz zur QoS-Sicherung der Audiodaten

Die Audiodaten haben in Videokonferenzen einen sehr hohen Informationsgehalt und müssen deshalb mit einer sehr guten Qualität übertragen werden. Dazu gehören eine geringe Verzögerung sowie eine geringe Verlustrate. Dies kann nur durch die Anpassung der Sendeleistung an die verfügbaren Ressourcen er-

reicht werden. In den meisten Videokonferenzsystemen wird eine solche Anpassung jedoch nicht unterstützt. Es kommen zwar Mechanismen zur Vermeidung eines hohen Jitters und zur Behebung von Paketverlusten zum Einsatz, jedoch wurde bisher die Vermeidung von Stausituationen nicht betrachtet. Im Folgenden soll dieses Problem anhand der vorgeschlagenen hierarchischen Topologie in OCTOPUS und BRAVIS untersucht werden.

### 5.3.1. Einflussgrößen auf die Sendeleistung von Audioströmen

Aufgrund des hohen Informationsgehalts von Audiodaten kann die Senderate der Audioströme nicht durch Weglassen einzelner Datenteile des Stroms reduziert werden, wie das bei Videoströmen möglich wäre. Die Folge wären deutliches Knacken und Rauschen, wodurch die Sprache nicht mehr verständlich wiedergegeben würde. Aus diesem Grund kann die Sendeleistung nur durch die Parameter *Samplingrate*, *Quantisierungstiefe*, *Silence Detection* und *Kompression* reduziert werden.

*Samplingrate*

Die Samplingrate stellt die Abtasthäufigkeit dar. Je höher die Samplingrate ist, desto höher ist die Sendeleistung. Andererseits führt die Verringerung der Abtastfrequenz aufgrund des Nyquist-Theorems zu einer Einengung des Sprachsignalspektrums und damit zu einer Verringerung der Sprachqualität. Ein für Videokonferenzen und die Telefonie üblicher Wert ist eine Samplingrate von 8 kHz. Nach dem Nyquist-Theorem kann hiermit eine Bandbreite des Sprachsignals von maximal 4 kHz abgetastet werden. Wird die Abtastrate auf die Hälfte verringert, verringert sich die Bandbreite des Sprachsignals ebenfalls um die Hälfte auf 2 kHz. Dies ist jedoch nur etwa ein Achtel des für den Menschen hörbaren Sprachspektrums. Die Verständlichkeit der Sprache nach dieser Abtastung wäre nur noch sehr gering. Aus diesem Grund ist eine Anpassung der Sendeleistung über die Samplingrate nur bedingt möglich.

*Quantisierungstiefe*

Die Genauigkeit der Reproduktion eines Audiosignals hängt von der Anzahl der verwendeten Quantisierungsbits pro Abtastung ab. Je mehr Bits verwendet werden, desto ähnlicher ist das abgespielte Audiosignal dem Original. Die Sendeleistung steigt linear mit der Quantisierungstiefe. Für die digitale Sprachkodierung ist ein Wert von 8 bit üblich. Weil die Quantisierung jedoch hauptsächlich in Hardware durchgeführt wird, ist eine Anpassung der Sendeleistung

leistung auch hier kaum möglich. Es werden zumeist nur Quantisierungstiefen von 8 und 16 bit angeboten.

Silence Detection erkennt längere Sprechpausen eines Teilnehmers. In diesen Pausen werden keine Daten übermittelt, wodurch Bandbreite eingespart werden kann. Für kontinuierliche Vorträge mit kurzen Sprechpausen kann die Silence Detection jedoch kaum die Sendeleistung effektiv reduzieren.

*Silence Detection*

Durch die Wahl geeigneter Kompressionstechniken kann die Sendeleistung von Audiosignalen am wirkungsvollsten und wenn erforderlich auch drastisch reduziert werden. In Abschnitt 2.5.2 wurde bereits auf die Möglichkeiten der Audiokomprimierung eingegangen. Die Kompression und die Dekompression beim Sender beziehungsweise beim Empfänger benötigen jedoch auch Rechenkapazitäten und Rechenzeit, die Verzögerungen bei der Übertragung des Audiosignals vom Sender zum Empfänger zur Folge haben.

*Kompressionsmethoden*

Für die meisten Videokonferenzsysteme haben sich die Werte der ISDN-Telefonie von 8 kHz Samplingrate und der 8 bit Quantisierungstiefe als ausreichend erwiesen. Dies resultiert in einer Übertragungsrate von 64 kbit/s. In der ersten Entwicklungsphase des OCTOPUS- und des BRAVIS-Systems wurde diese Übertragungsqualität ebenfalls genutzt, da sowohl COVIS als auch HKUST dieses Audioformat genutzt haben [12, 25]. Tests zwischen der BTU Cottbus und der Hongkonger Universität haben gezeigt, dass die Übertragungsqualität über diese große Entfernungen meistens zufrieden stellend ist. Jedoch ist es für zukünftige Videokonferenzanwendungen wichtig, so wenig wie möglich Bandbreite für die Audioübertragung zu nutzen. Ansonsten würde es bei der schnell wachsenden Nutzung von Videokonferenzsystemen zu einer gewaltigen Belastung der Netzwerke kommen, auch wenn die Audiobandbreite nur sehr gering im Vergleich zur Videobandbreite ist.

#### 5.3.2. Anpassung der Sendeleistung eines Audiostroms

Im OCTOPUS-Verbundsystem legen die Endsysteme in den lokalen Gruppen die Samplingrate, die Quantisierungstiefe und eventuell eine Silence Detection des Audiostroms fest. Da in lokalen Netzwerkkumgebungen, insbesondere wenn wie bei COVIS und HKUST ATM genutzt wird, kaum Bandbreitenengpässe auftreten, bedarf es hier keiner Komprimierung. Somit muss für die globale Über-

OCTOPUS

tragung nur der Gruppenkoordinator die Fähigkeit besitzen, einen Audiostrom zu komprimieren und entsprechend der verfügbaren Ressourcen anzupassen.

BRAVIS

Die Samplingrate, die Quantisierungstiefe und die Silence Detection wird in BRAVIS einheitlich definiert. Aufgrund der Topologie der Audioverteilung muss jedes Endsystem die Audioströme komprimieren und dekomprimieren können sowie die Datenrate anpassen können. Da BRAVIS im günstigsten Fall davon ausgeht, dass sich eine Untergruppe in einem lokalen Netzwerk befindet, bei dem kaum Bandbreitenengpässe auftreten, kann der *Ausgewählte Empfänger (AE)* den empfangenen komprimierten Audiostrom dekomprimieren und dann an die anderen „lokalen“ Teilnehmer verteilen. Somit können deren Ressourcen geschont werden. Dieser Mechanismus funktioniert jedoch nicht ausreichend gut, wenn sich alle Teilnehmer einer Untergruppe an einem anderen Standort befinden. In diesem Fall ist es günstiger, dass der AE die komprimierten Daten einfach weiterleitet. Durch das dynamische QoS-Management in BRAVIS können diese Fälle erkannt und die AEs entsprechend eingestellt werden.

Kompressions-  
methode

Die Reduktion der Datenrate eines Audiostroms wird aufgrund der oben genannten Diskussion durch Kompression im Gruppenkoordinator von OCTOPUS beziehungsweise in den Endsystemen von BRAVIS erreicht. Ein einzelner Audiocodec der G.7xx-Familie ist jedoch für eine dynamische Anpassung der Sendeleistung ungeeignet, da nur eine fest spezifizierte Senderate pro Codec möglich ist. Deshalb müssen verschiedene Kompressionsstandards unterstützt werden. Abhängig von der verfügbaren Bandbreite kann der entsprechende Codec ausgewählt werden. Die INRIA-Forschungsgruppe hat diesen Ansatz im IVS umgesetzt [17]. Dabei wurde festgestellt, dass aufgrund der relativ großen Unterschiede der Senderate verschiedener Codecs eine genaue Anpassung an die vorhandenen Ressourcen kaum möglich ist. Diese Sprunghaftigkeit kann sich bei einer Erhöhung der Sendequalität negativ auf die Empfangsqualität auswirken, da durch das plötzliche Ansteigen der Datenrate Stausituationen provoziert werden können.

Günstiger ist die Nutzung der MPEG-Layer-3-Standards (MP3), obwohl nicht unbedingt für die Kompression von Sprachsignalen konzipiert wurde<sup>1</sup>. MP3 ist weit verbreitet und unterstützt eine dynamisch an die vorhande-

---

<sup>1</sup>MP3 unterstützt ein breites Spektrum von Audiofrequenzen für HiFi-Signale (etwa von 50 Hz bis 20 kHz). Sprache hingegen nutzt nur einen Bruchteil dessen (etwa von 300 Hz bis 4 kHz).

nen Netzwerkressourcen angepasste Bitrate. Durch eine kontinuierliche Änderung der Sendeleistung werden kaum drastische Überlastsituationen provoziert. MP3 benötigt jedoch hohe Ressourcen für die Kompression. Somit ist in zukünftigen Arbeiten noch zu prüfen, wie viele Gruppen ein OCTOPUS-Gruppenkoordinator beziehungsweise wie viele Untergruppen ein BRAVIS-Endsystem mit unterschiedlicher Audio-Qualität unterstützen kann.

In globalen öffentlichen Netzwerken können trotz angepasster Sendeleistung geringfügige Verluste auftreten. Deshalb ist es für die Videokonferenzsysteme für globale Netzwerke sinnvoll, eine Fehlerkorrektur mittels FEC zu integrieren. Abbildung 5.2 zeigt das Schema dieses Konzepts. Die digitalen Audiosignale

Fehlerkorrektur

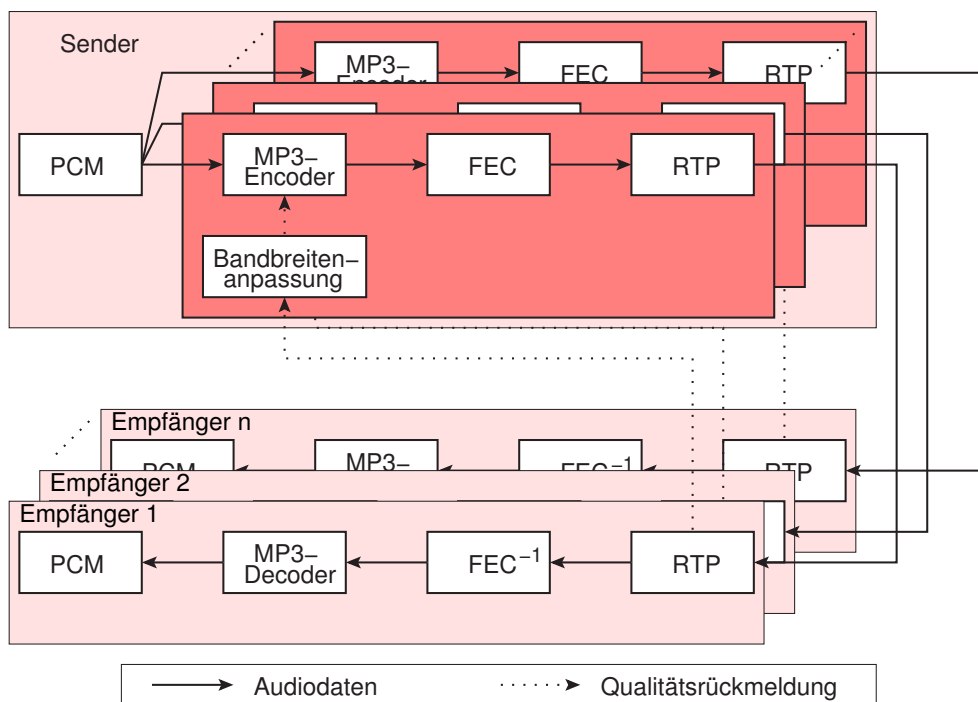


Abbildung 5.2.: Das Konzept der Audiokommunikation

werden vorerst durch einen Kodierer, zum Beispiel MP3, auf eine vorgegebene Bitrate komprimiert. Um Paketverluste auszugleichen, wird der komprimierte Audiostrom mit Redundanzen wieder vergrößert und übertragen. Um Heterogenität zu unterstützen, muss der Audiostrom individuell an die Vorgaben der Empfänger angepasst werden. Der Empfänger, also der Gruppenkoordinator im OCTOPUS und der AE im BRAVIS, teilt dem Sender die erkannten Verluste

beziehungsweise die empfangene Bitrate mit. Unter Berücksichtigung der hinzugefügten Redundanzen wird daraufhin die Bitrate bestimmt, mit der kodiert werden muss.

Für den Fall, dass sich BRAVIS-Teilnehmer einer Untergruppe nicht in einem lokalen Netz befinden, muss der AE die Datenströme weiterleiten. Die Empfänger bestimmen dann die empfangene Qualität und teilen diese dem Sender mit.

Ein ähnliches Konzept wurde bereits vor einigen Jahren von der INRIA-Gruppe im System IVS umgesetzt [17]. Es wurde eine feste Menge an „leichtgewichtigen“ Kompressionsmethoden wie dem heutigen G.722-Standard genutzt, um eine stufenweise Anpassung der Sendeleistung zu erreichen. Zeitgleich wurde der Audiostrom mit einem „schwergewichtigen“ Kodierer wie beispielsweise G.729 komprimiert, um als FEC zu dienen (siehe Abbildung 5.3). Der Nutzen-Kosten-Faktor dieses Konzeptes ist aufgrund der doppelten Kompression und der damit verbundenen notwendigen Rechenkapazitäten eher gering, da die gesendeten Datenraten nur unwesentlich verringert wurden. Des Weiteren ist der Ansatz der INRIA nicht für heterogene Netzwerkumgebungen tauglich, da der Sender nur *eine* Senderate unterstützt.

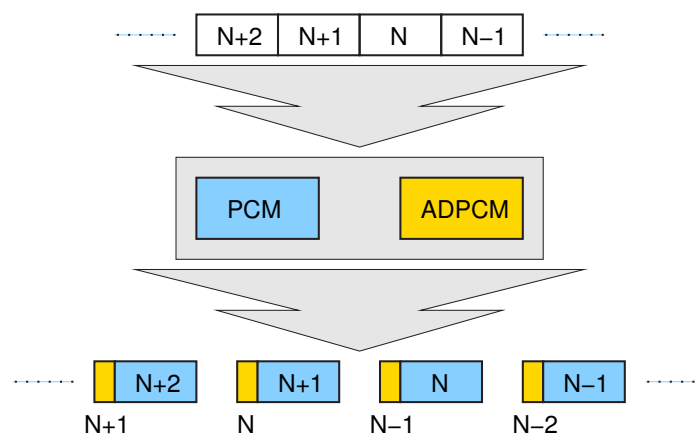


Abbildung 5.3.: Konzept der Audioübertragung in IVS

**Erfahrungen**

Das in Abbildung 5.2 gezeigte Konzept für OCTOPUS und BRAVIS ist zur Zeit noch nicht vollständig in beiden Systemen integriert. Jedoch verspricht eine erste Prototypimplementierung eine sehr robuste und hochqualitative Audioübertragung, da kaum Verluste der reinen Audiodaten auftreten können.

Selbst wenn durch Erhöhung der Datenrate eine Überlastsituation mit Paketverlusten auftreten sollte, werden diese Verluste durch das FEC wieder aufgehoben. Somit kann rechtzeitig die Sendeleistung reduziert werden, bevor die Verluste so hoch werden, dass sie nicht mehr durch das FEC ausgeglichen werden können.

Der MP3-Codec als auch FEC benötigen erhöhte Rechenkapazitäten auf den ausführenden Hosts und erhöhen die Übertragungszeiten. Aufgrund der Nutzung einer Floorkontrolle, die nur einem Teilnehmer das Rederecht überträgt, muss jedoch keine hundertprozentige Echtzeitkommunikation unterstützt werden. Deshalb spielt eine geringfügige Erhöhung der Übertragungszeit keine große Rolle.

## 5.4. Ein Ansatz zur QoS-Sicherung der Videodaten

Die Übertragung von Videodaten ist die eigentliche Herausforderung von Videokonferenzsystemen, da die Sendeleistung der Videoströme im Allgemeinen deutlich höher ist als die der Audiodaten. Auch bei der Videoübertragung ist es wichtig, eine für die Nutzer annehmbare Dienstqualität zur Verfügung zu stellen. In diesem Abschnitt wird die Sicherung der Dienstqualität in geschlossenen Videokonferenzsystemen mit einer hierarchischen Verteilungsstruktur untersucht. Die Vermeidung von Verzögerungen spielt dabei eine bedeutendere Rolle als die Vermeidung von Verlusten, da im Allgemeinen Verluste bis zu 51 % von Nutzern noch akzeptiert werden [114].

### 5.4.1. Einflussgrößen auf die Sendeleistung von Videoströmen

Die für den Menschen wahrnehmbare Qualität einer Videoübertragung hängt von der Größe des Bildes, der Farbechtheit, der Verzögerung des Stromes zwischen Sender und Empfänger sowie der Bildwiederholungsrate ab. Eine gute Darstellungsqualität benötigt jedoch für eine Übertragung eine sehr hohe Bandbreite. Trotz der Kompression der Videoströme kann die Bandbreite in globalen öffentlichen Netzen knapp werden. Um Stausituationen im Netz-

werk zu vermeiden, muss deshalb die Datenrate eines Senders an die verfügbaren Ressourcen angepasst werden. Eine Anpassung der Sendeleistung des Videostroms kann im Vorfeld der Kompression hauptsächlich durch die Parameter *Bildgröße*, *Farbtiefe* und *Bildwiederholfrequenz* erfolgen.

### Bildgröße

Die Bildgröße legt die Anzahl der Pixel in der Höhe und Breite fest. Bei Verdopplung der Breite und der Höhe eines Bildes erhöht sich die Sendeleistung um das Vierfache. Eine dynamische Anpassung der Bildgröße in einem Videokonferenzsystem ist ungünstig, da dies beim Empfänger eine ständig, in Abhängigkeit der verfügbaren Bandbreite, schwankende Bildgröße zur Folge hätte. Jedoch können weniger wichtige Datenströme, zum Beispiel die Zuhörer, mit einer geringeren Bildgröße übertragen werden als wichtige, zum Beispiel der des Sprechers, um somit Bandbreite zu sparen. Beispielsweise realisiert das Videokonferenzsystem COVIS diese Unterscheidung mit Hilfe der integrierten Floorkontrolle. Typische Größen sind CIF und QCIF mit 382x288 beziehungsweise 196x144 Pixeln.

### Farbtiefe

Die Farbtiefe für jedes Pixel eines Bildes legt die Farbtreue fest. Ein typischer Wert für ein Videokonferenzsystem ist 8 bit/Pixel. Wie die 8 bit auf die Farben aufgeteilt werden, hängt hauptsächlich von dem verwendeten Farbmodell ab. Das gesendete Datenvolumen ist proportional zur Farbtiefe. Eine dynamische Anpassung ist hierbei kaum möglich, da ständig das Farbmodell geändert werden müsste. Die dynamische Variation der Farbinformation ist außerdem für viele Benutzer ungewöhnlich und deshalb nicht akzeptabel.

### Bildwiederhol- frequenz

Die Bildwiederholfrequenz, auch Bildrate oder Framerate genannt, bestimmt die Auflösung des Videostroms in der Zeitachse, indem die Anzahl der Videobilder in einer Zeiteinheit festgelegt wird. Die Sendeleistung ist proportional zur Framerate. Dies stellt das günstigste Mittel dar, um vor der Kompression die Datenrate zu reduzieren. 25 Bilder pro Sekunde (*Frames per second, fps*) liefern die meisten Videokarten für Workstations und PCs. Dies ist für das flüssige Abspielen von Videosequenzen im Spielfilmbereich notwendig. Für Videokonferenzen reichen in der Regel 15 fps aus, um die für Videokonferenzen typischen Kopf-Schulter-Aufnahmen zu übertragen.

Mit diesen Parametern lässt sich die Datenrate des Videostroms jedoch nicht so reduzieren, dass sie günstig über ein Netzwerk übertragen werden kann. Zum Beispiel benötigt ein Videostrom der Größe QCIF mit 8 bit/Pixel bei 15



fps eine Bandbreite von ca. 3,3 Mbit/s. Die Sendeleistung wird deshalb durch die Komprimierung entscheidend reduziert, wobei die Kompression durch die *Quantisierung*, die *Redundanzerkennung* und die *Bewegungserkennung* beeinflusst wird.

Die Quantisierung steht im Zusammenhang mit der Bildkompression mittels *Discrete Cosine Transformation (DCT)*. Je gröber die Quantisierung gewählt wird, desto mehr Details des Bildes gehen verloren. Die Kompression des Bildes ist aber höher und die Sendeleistung geringer.

Quantisierung

Bei der Videoübertragung eines relativ bewegungslosen Objekts, wie einem Teilnehmer einer Videokonferenz in der typischen Kopf-Schulter-Aufnahme, gleichen sich aufeinander folgende Bilder. Je ähnlicher die Bilder sind, desto geringer ist die Sendeleistung bei Erkennung der Redundanzen, da nur die Differenzen der Bilder übertragen werden müssen. Der Nachteil ist jedoch, dass sich der Verlust eines Pakets aus dem Videostrom auf dessen gesamte nachfolgende Darstellung negativ auswirken kann, da wichtige Referenzpunkte fehlen können.

Redundanz-  
erkennung

Im Gegensatz zur Erkennung von Redundanzen werden hier sich bewegende Objekte in einer ruhigen Umgebung erkannt. Die Bewegung eines Objektes kann dann durch einen Bewegungsvektor beschrieben werden, wodurch sich die Sendeleistung ebenfalls verringern kann. Dies erfordert jedoch eine hohe Rechenkapazität und bewirkt höhere Verzögerungen. Für Videokonferenzen ist das Verfahren außerdem kaum geeignet, weil sich die Sprecher kaum bewegen.

Bewegungs-  
erkennung

In Videokonferenzsystemen lässt sich also die Senderate des Videostroms am besten durch die Bildwiederholungsrate und die Quantisierung während der Kompression dynamisch beeinflussen. Unter diesem Gesichtspunkt wird der im Folgenden beschriebene Anpassungsmechanismus betrachtet.

### 5.4.2. Ein Ansatz zur Bandbreitenanpassung durch Abstimmung

In geschlossenen Videokonferenzsystemen kennen sich alle Teilnehmer untereinander. Außerdem sendet jeder Teilnehmer an alle anderen sein eigenes Videobild. Bei größeren Konferenzen, insbesondere in einem Konferenzverbund wie dem OCTOPUS, müssen somit eine Vielzahl von Videoströmen angezeigt werden. Das führt zu dem so genannten *Sichtbarkeitsproblem*. Die Bildschirm-

Sichtbarkeitspro-  
blem

größe und –auflösung begrenzt die Anzahl der gleichzeitig nebeneinander darstellbaren Teilnehmer. Tabelle 5.1 gibt Beispiele für die Anzahl der anzeigbaren Teilnehmer für unterschiedliche Bildschirm– und Videoauflösungen an. Zusätzlich reduzieren in der Praxis oft ein Whiteboard oder andere *Joint Viewing*–Anwendungen die Darstellungsfläche des Bildschirms. Die Leistungsfähigkeit

Bildschirm– auflösung	Videogröße	
	QPAL (384x288)	1/16 PAL (192x144)
VGA (640x460)	1	9
SVGA (800x600)	4	16
XGA (1024x786)	4	25
1280x1024	9	42

Tabelle 5.1.: Anzahl der Teilnehmer, die abhängig von der Bildschirmauflösung und der Videobildgröße auf dem Bildschirm Platz finden

des darstellenden Endsystems begrenzt ebenfalls die Anzahl der angezeigten Videoströme, weil diese zur Anzeige von beispielsweise mehr als 10 Teilnehmern nicht ausreicht. Deshalb ist es sinnvoll, nur einige Teilnehmer der Gruppe darzustellen. Neben dem Sprecher können dies Teilnehmer sein, die besonders relevant für die Diskussion in der Gruppe sind. Beispielsweise sind in Podiumsdiskussionen nur die Gruppe der Diskutierenden und deren Reaktionen interessant. In geschäftlichen Verhandlungen, bei denen Mitarbeiter die Geschäftsführer konsultieren, sind ebenfalls nur die Geschäftsführer und deren Reaktionen beachtenswert. Weitere Beispiele lassen sich beim Teleteaching, insbesondere bei einem Seminarszenario finden, bei dem meistens der Tutor und aktive Studenten im Mittelpunkt stehen. In medizinischen Beratungen ist es wichtig, neben dem aktuell sprechenden Arzt eine Dokumentenkamera mit den Befund oder den OP–Tisch zu übertragen.

Basierend auf diesen Beobachtungen wurde im OCTOPUS–System eine abstimmungs-basierte Bandbreitenverwaltung auf den Punkt–zu–Punkt Verbindungen zwischen den Gruppenkoordinatoren entworfen [184]. Der Ansatz ist so allgemeingültig, dass er auch für BRAVIS und andere geschlossene Videokonferenzsysteme genutzt werden kann.

Um eine solche Auswahl zu ermöglichen, wurde im Rahmen dieser Arbeit folgender Abstimmungsmechanismus in OCTOPUS integriert. Jeder Teilnehmer verwaltet eine *Auswahltabelle (Voting Table)* auf seinem Bildschirm, die die

Prinzip des Algori-  
thmus

Namen aller Teilnehmer der Konferenz enthält. Er kann dynamisch auswählen, welche Teilnehmer er sehen möchte oder nicht. Der aktuelle Sprecher wird immer angezeigt, egal, ob er ausgewählt ist. Bild 5.4 zeigt ein Beispiel für diese Auswahl. Konferenzgruppe *A* hat die drei Teilnehmer: *r*, *s* und *t*. Die Konferenzgruppe *B* besteht aus den Teilnehmern *m*, *n*, und *o*, während Gruppe *C* nur die zwei Teilnehmer *p* und *q* enthält. Es wird die Auswahl der Gruppe *A* dargestellt. Jeder Teilnehmer aus *A* legt fest, welche Teilnehmer aus *B* und *C* er sehen möchte.

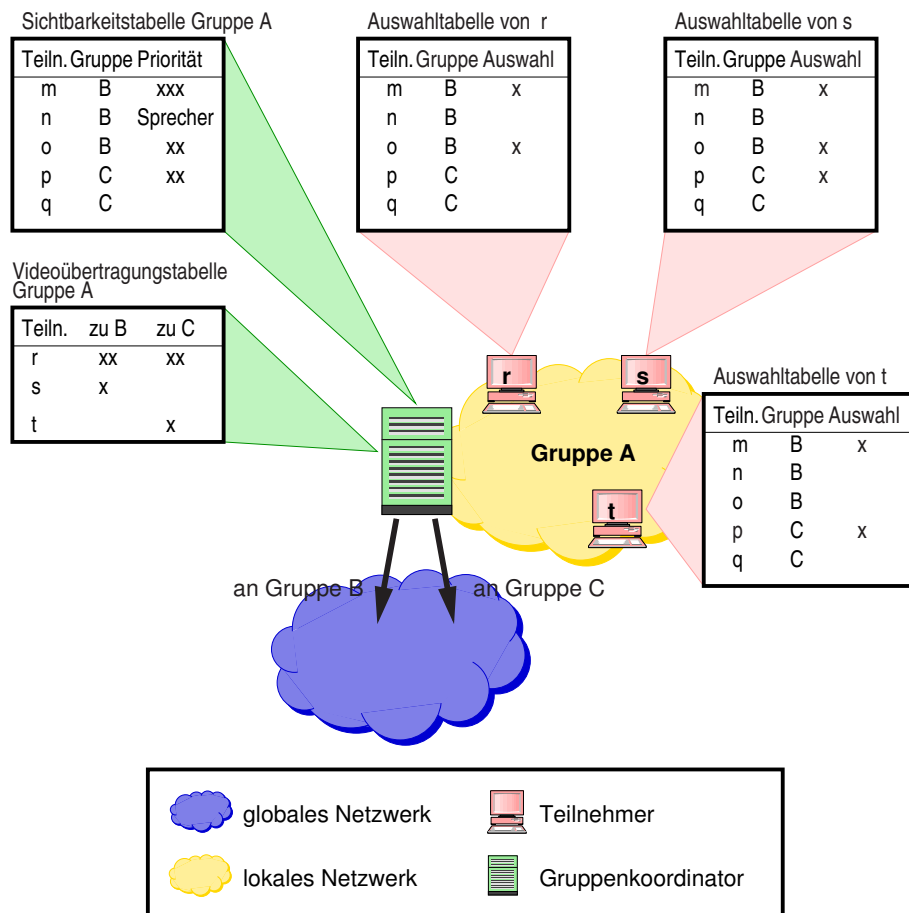


Abbildung 5.4.: Beispielszene einer Auswahl in einer OCTOPUS-Sitzung

Die Gruppenkoordinatoren werten periodisch die Auswahlstabellen der eigenen lokalen Teilnehmer aus und bestimmen nach der Mehrheit der abgegebenen Stimmen die Teilnehmer, die in der lokalen Konferenz dargestellt werden sollen. Das Abstimmungsergebnis wird in der *Sichtbarkeitstabelle* (*Visibility Table*)

auf jedem Gruppenkoordinator gespeichert. Die Videoströme der Teilnehmer mit den meisten Stimmen und bis zu der Anzeigegrenze werden abgespielt. Die Auswahl kann zu jeder Zeit durch die Teilnehmer geändert werden. Die Sichtbarkeitstabelle wird bei einem Sprecherwechsel und periodisch aktualisiert. Nach jeder Aktualisierung werden die Sichtbarkeitstabellen zwischen den Gruppenkoordinatoren ausgetauscht. Die ausgewählten Teilnehmer werden in der *Videoübertragungstabelle (Video Transmission Table)* markiert und gespeichert. Nur deren Videoströme werden zu den entfernten Gruppen gesendet, um somit Platz auf den Bildschirmen der Empfänger und Bandbreite einzusparen.

Bandbreiten-  
messung und  
-anpassung

Die Abstimmungsergebnisse werden weiterhin genutzt, um die zu übertragenden Videoströme auszuwählen, wenn die verfügbare Bandbreite auf den globalen Verbindungen nicht für alle Teilnehmer ausreicht. Zu diesem Zweck wird die Bandbreite auf diesen Verbindungen durch die Gruppenkoordinatoren auf Empfängerseite periodisch gemessen. Die Messergebnisse werden über das GCP<sub>IP</sub> an die Quelle zurückgeschickt. Die RTCP **ReceiverReports (RR)** werden an dieser Stelle nicht genutzt, da das GCP<sub>IP</sub> einen 50% geringeren Signalisierungsoverhead benötigt [183]. Des Weiteren beziehen sich RR auf jeden einzelnen RTP-Strom, während GCP<sub>IP</sub> in der Lage ist, die Messergebnisse aller Ströme in einem Paket auszutauschen. Dieses Paket wird durch die Gruppenkoordinatoren ausgewertet, so dass jeder Gruppenkoordinator die verfügbare Bandbreite auf jeder Verbindung kennt. Für die Anpassung der Sendeleistung wird ein Algorithmus wie der von Wakeman [19, 171] angewendet:

```
(1)  if(Überlast)
(2)      Senderate = max(Senderate/GAIN, MIN_RATE);
(3)  else if(KeineÜberlast)
(4)      Senderate = min(Senderate + INC, MAX_RATE);
```

Im vorgeschlagenen Ansatz wird anstelle der gesendeten Bitrate die gesendete Bildwiederholungsrate genutzt<sup>2</sup>, um die Videoströme einfacher skalieren zu können. In Abschnitt 5.4.5 wird dieser Aspekt detaillierter diskutiert. Weiterhin berücksichtigt der Algorithmus die gesamte Bandbreite der Verbindung

---

<sup>2</sup>Die Möglichkeit, die Datenrate durch die Skalierung der Bildqualität zu reduzieren, konnte aufgrund eingeschränkter Funktionalitäten der verwendeten Kompressionssoftware (X-Image-Library) für SUN Solaris nicht untersucht werden.

und nicht nur die Bandbreite einzelner Ströme, wodurch Einflüsse zwischen den Strömen reduziert werden.

Wenn der Paketverlust einen Schwellwert von 5% übersteigt, wird eine **Überlast**-Situation angenommen. Die **Senderate** wird daraufhin reduziert. Nach Wakeman wird die aktuelle Senderate durch den Wert **GAIN** geteilt oder eine definierte minimale Senderate mit der geringsten akzeptablen Qualität (**MIN\_RATE**) festgelegt. Der Wert **GAIN** wird gewöhnlich auf 2 gesetzt. Im vorgeschlagenen Ansatz für OCTOPUS wird die Zeile (2) folgendermaßen ersetzt:

Überlast

$$(2) \quad \text{Senderate} = \max(\text{Empfangsrate}, \text{MIN\_RATE});$$

wobei der Wert der **Empfangsrate** der gemessenen Framerate entspricht. Dadurch wirkt sich ein auftretender Bandbreitenengpass im Vergleich zu [171] nicht so stark aus.

Liegen die Paketverluste unter 2%, wird **KeineÜberlast** angenommen, und die Senderate wird periodisch um den Summanden  $\text{INC} = 1$  fps erhöht, bis eine definierte maximale Übertragungsrate (**MAX\_RATE**) erreicht wird. Diese maximale Rate kann nach drei Kriterien festgelegt werden:

keine Überlast

**TCP-freundlich:** **MAX\_RATE** wird so gewählt, dass sich die Medienströme die verfügbare Bandbreite einer physischen Verbindung mit parallelen TCP-Strömen teilen. Ansätze für eine dynamische Bestimmung dieses Wertes sind zum Teil sehr komplex, da der gesamte Datenverkehr über das Netz berücksichtigt werden muss [92, 111].

**Empfängerkapazität:** Ist es in einer Videokonferenzgruppe bekannt, dass kein Teilnehmer mehr als eine bestimmte Datenrate verarbeiten kann, so braucht die **MAX\_RATE** nicht über diesem Wert zu liegen.

**Senderkapazität:** Die meisten PCs können maximal 25 bis 30 fps von der Videokarte aufnehmen und verarbeiten. Somit macht es keinen Sinn, die **MAX\_RATE** über diese Werte zu legen.

Aufgrund der geringen Komplexität wird die maximale Senderate in dem entwickelten Anpassungsmechanismus vorerst auf 25 fps gesetzt. Es bleibt jedoch die Option offen, **MAX\_RATE** nach der Empfängerkapazität oder nach der TCP-Freundlichkeit zu setzen.

Die resultierende **Senderate** des Anpassungsprozesses bildet die Grundlage **FR** für den Algorithmus der auswahlbasierten Bandbreitenverwaltung, der in Abbildung 5.5 dargestellt ist. Dieser Algorithmus wurde im Rahmen dieser Arbeit für das OCTOPUS-System entworfen.

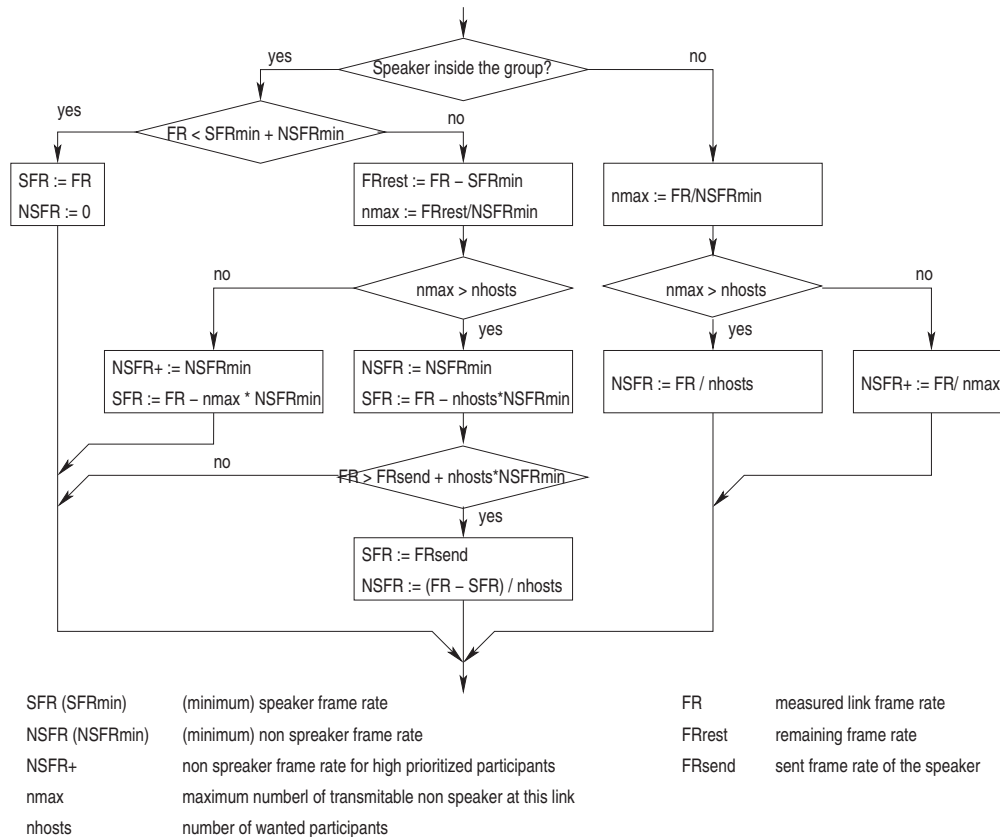


Abbildung 5.5.: Der auswahlbasierte Bandbreitenzuweisungsalgorithmus

**Bandbreitenzuweisung**

Der Algorithmus prüft zunächst, ob sich der aktuelle Sprecher in der lokalen Gruppe befindet. Ist der Sprecher in der eigenen Gruppe, wird ihm die höchste Priorität gegeben. Die *minimale Sprecherframerate* (*minimum speaker frame rate*) (**SFR<sub>min</sub>**) wird auf 10 fps festgelegt. Falls die verfügbare Bandbreite **FR** kleiner als dieser Schwellwert ist, wird die verfügbare Bandbreite nur von dem Sprecher genutzt. Die verbleibenden ausgewählten Nichtsprecher werden nicht übertragen. Um die ausgewählten Nichtsprecher zu übertragen, muss die verfügbare Bandbreite **FR** so groß sein, dass der Sprecher mit **SFR<sub>min</sub>** und die Nichtsprecher mit **NSFR<sub>min</sub>** (*minimale Nichtsprecherframerate, minimum non speaker frame rate*) übertragen werden können. Ist die verfügbare

Bandbreite zu klein, um alle gewählten Nichtsprecher zu übertragen, werden nur die Ströme der Nichtsprecher mit den meisten Stimmen übertragen. Bandbreite, die die Nichtsprecher nicht belegen, wird wie in dem Beispiel im folgenden Absatz zu der Sprecherbandbreite (SFR) hinzu addiert. Die Sendeleistung des Sprechers kann weiter steigen, wenn alle ausgewählten Nichtsprecher mit  $NSFR_{min}$  übertragen werden. Die Bandbreite der ausgewählten Nichtsprecher bleibt konstant bei ( $NSFR$ ). Erst wenn die zugewiesene Sprecherframerate den Wert der im Gruppenkoordinator empfangenen Framerate des Sprechers ist ( $SFR == FR_{sent}$ ), wird die Sendeleistung der gewählten Nichtsprecher ( $NSFR$ ) erhöht. Befindet sich der Sprecher nicht in der eigenen Gruppe, wird die Bandbreite  $FR$  auf alle ausgewählten Teilnehmer gleichmäßig aufgeteilt. Die Ströme der anderen Teilnehmer werden nicht übertragen. Diese Prozedur stellt somit sicher, dass der Sprecher mit der höchsten Priorität und mit mindestens 10 fps übertragen wird. Die ausgewählten Teilnehmer werden, solange die Bandbreite ausreicht, mit mindestens 3 fps übertragen.

Die Prozedur bezieht sich auf je eine Verbindung zu einer entfernten Gruppe. Dadurch ist es möglich, dass ein Strom eines Teilnehmers mit von einander abweichender Sendeleistung zu verschiedenen Gruppen gesendet wird.

Abbildung 5.6 demonstriert beispielhaft eine Bandbreitenzuweisung für die Abstimmung aus Abbildung 5.4. Es wird angenommen, dass die Bandbreite zwischen  $A$  und  $B$  19 fps sowie zwischen  $A$  und  $C$  12 fps beträgt. Die Gruppenkoordinatoren  $B$  und  $C$  bestimmen die Bandbreitenzuweisung ihrer Gruppenmitglieder aufgrund der Sichtbarkeitstabelle der Gruppe  $A$ . Zuerst soll die Gruppe  $B$  mit den drei Teilnehmern  $m$ ,  $n$ , und  $o$  betrachtet werden. Der aktuelle Sprecher soll  $n$  sein. Dieser wird mit mindestens 10 fps gesendet. Da die verbleibende Framerate 9 fps beträgt, können die Videoströme der Teilnehmer  $m$  und  $o$  mit jeweils 3 fps übertragen werden. Die verbleibenden 3 fps können dem Sprecher zugewiesen werden. Von der Gruppe  $C$  wurde nur Teilnehmer  $p$  durch die Gruppe  $A$  ausgewählt. So können diesem Teilnehmer alle 12 fps zugewiesen werden. Teilnehmer  $q$  wird nicht angezeigt. Falls die Bandbreite zwischen  $A$  und  $B$  auf 14 fps sinkt, wird nur der Videostrom des Teilnehmers  $m$  (meiste Stimmen) mit 3 fps gesendet. Der Sprecher  $n$  bekommt in diesem Fall 11 fps. Teilnehmer  $o$  kann nicht länger zur Gruppe  $A$  übertragen werden,

<i>Beispiel</i>
-----------------

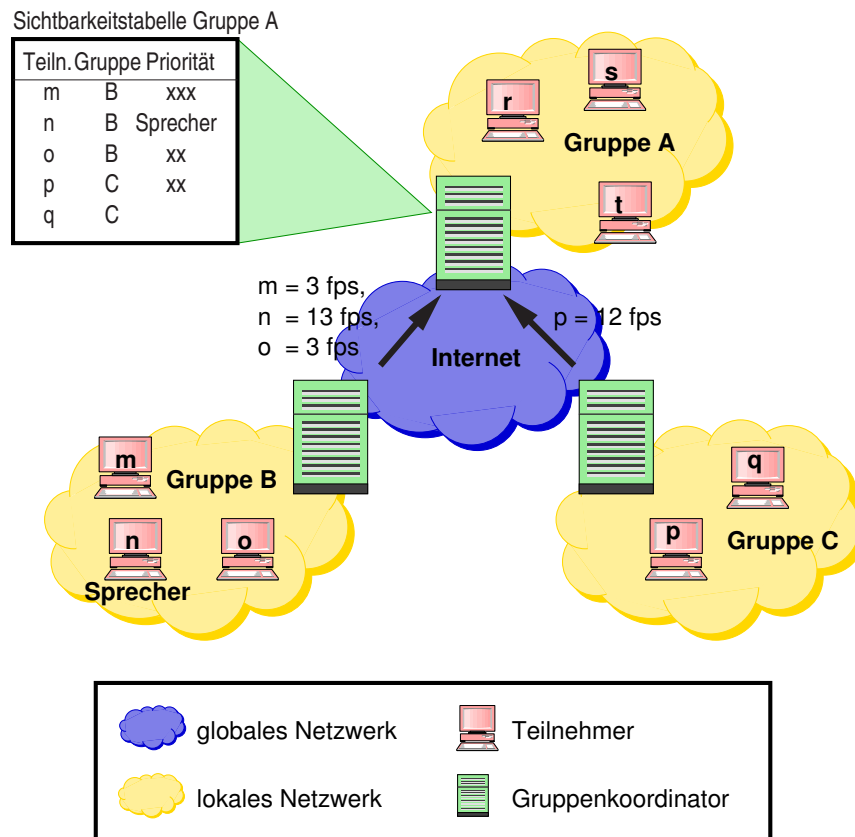


Abbildung 5.6.: Bandbreitenzuweisung basierend auf der Abstimmung

da die verfügbare Bandbreite keine zwei Videostreams mit mindestens 3 fps zulässt.

*lokale Skalierung*

Um eine Überlastung der Host der Empfänger zu vermeiden, wird die folgende Regel für das Empfangen der Videostreams angewendet. Im Allgemeinen werden die entfernten Videostreams mit höherer Priorität eingestuft als lokale Streams. Kann beispielsweise ein Host in Gruppe A 40 fps verarbeiten, dann kann dieser alle Streams der Gruppe B (19 fps) und C (12 fps) empfangen und verarbeiten. Die verbleibende Bandbreite von 9 fps wird durch die Teilnehmer der Gruppe A zu gleichen Teilen genutzt. Die Auswahltabelle kann zur Vermeidung des Sichtbarkeitsproblems auch auf die lokalen Teilnehmer ausgedehnt werden. Diese Auswahl hat jedoch keine Einfluss auf die globale Abstimmung.



### 5.4.3. Bewertung des Ansatzes der Bandbreitenanpassung

In diesem Abschnitt soll der vorgestellte Ansatz durch Messungen auf seine praktische Anwendbarkeit bewertet werden. Die Messungen nehmen Bezug auf die Übertragungsverzögerung und die Bandbreitenverfügbarkeit auf globalen Verbindungen, das Anpassungsverhalten sowie die Verzögerung, die durch Gateways beziehungsweise durch die Gruppenkoordinatoren auftreten.

#### Verzögerungen auf globalen Verbindungen

Die Übertragungsverzögerung ist ein bedeutender Faktor einer Videokonferenz. Das betrifft nicht nur die Audio- und Videoübertragung, sondern auch die Signalisierung. Um zu prüfen, für welche geographischen Entfernungen ein Videokonferenzsystem wie das OCTOPUS geeignet ist, wurde zunächst die Verzögerung für typische Paketlängen<sup>3</sup> gemessen.

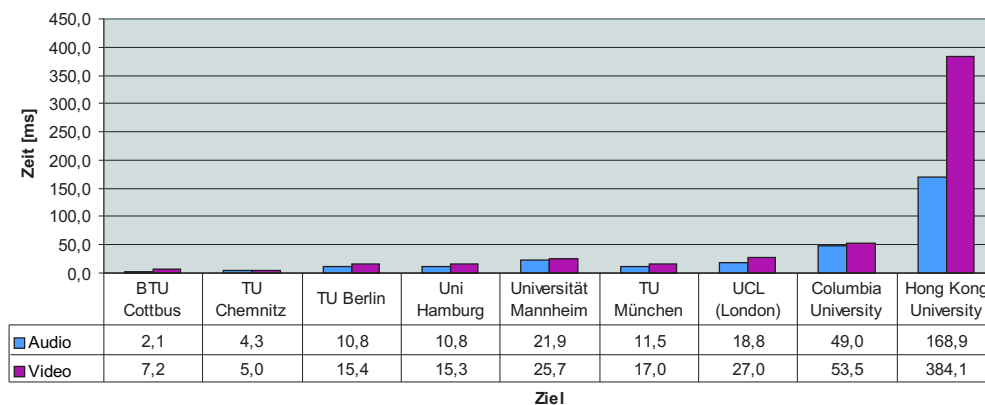


Abbildung 5.7.: Übertragungsverzögerungen für verschiedene Paketgrößen

Die Messungen wurden parallel zu den Messungen der Signalisierung im Vorfeld der Implementierung vom Standort Cottbus zu verschiedenen Standorten innerhalb Deutschlands, Europas und der Welt durchgeführt (siehe auch Abschnitt 4.3.3). Die Ergebnisse der Messungen, die für die Verzögerungen der Audio- und Videodaten relevant sind, sind in Abbildung 5.7 dargestellt. Sie zeigen, dass die Verzögerungen insbesondere der Audiodaten nach Nordamerika und selbst nach Hong Kong unter 400 ms liegen. Dadurch wird insbesondere

<sup>3</sup>Aus dem COVIS wurden Paketgrößen von 512 Byte für Audiodaten und 3000 Byte für Videodaten für die Messungen im Vorfeld der Implementierung übernommen

der Grenzwert von 400 ms für eine akzeptable Audioverzögerung nicht überschritten. Somit werden qualitativ hochwertige Audio- und Videokonferenzen weltweit möglich. Die geringere Bandbreite und besonders die Verzögerungen durch eine Satellitenverbindung, wie beispielsweise nach Hong Kong, bilden schlechte Bedingungen für eine Videokonferenz. Wie die Messungen belegen, ist dennoch ein interaktiver Datenaustausch möglich. Eine Kopplung eines Videokonferenzsystems über zwei Satellitenverbindungen bewirken jedoch inakzeptable Verzögerungen.

### **Bandbreite**

Neben der Verzögerung ist die Bandbreite der wichtigste Faktor für die Qualität einer Videokonferenz. Im deutschen Forschungsnetz G-WiN [167] wird zur Zeit eine Bandbreite von 2,5 Gbit/s angeboten, wodurch Engpässe in der Bandbreite kaum auftreten. Auch die Anbindung beispielsweise der BTU Cottbus mit 34 Mbit/s stellt keinen Flaschenhals mehr dar, da diese Bandbreite im normalen täglichen Verkehr kaum über 10% ausgelastet wird. Messungen zwischen der BTU Cottbus und der Universität Mannheim über einen Zeitraum von 24 Stunden haben dies bestätigt. Jedoch ist zu beachten, dass in Zukunft bei einer stark verbreiteten Nutzung von Videokonferenzsystemen die Bandbreitenressourcen doch knapp werden können.

Die durchgeführten Messungen konzentrierten sich deshalb auf globale Verbindungen mit hohen Verzögerungen und begrenzten Bandbreitenressourcen wie beispielsweise bei einer Satellitenverbindung. Beispielhaft wurde die Bandbreite nach Hong Kong gemessen.

Die Messungen wurden folgendermaßen durchgeführt. Ein Sender steigert seine Sendeleistung kontinuierlich alle 30 Sekunden um 1 fps, wobei mit 3 fps begonnen wird. Auf der Empfängerseite wird die empfangene Bandbreite in *fps* und *kbit/s* gemessen. Des Weiteren wurde die Verzögerung abgeschätzt, nachdem die Zeitdifferenz zwischen Sender und Empfänger festgestellt wurde. Die Messungen wurden mehrmals in der Hauptgeschäftszeit zwischen 8 Uhr und 16 Uhr wiederholt. Die Ergebnisse dieser Messung in der Richtung von Hong Kong nach Cottbus sind in Abbildung 5.8 dargestellt. Sie zeigen, dass die Verbindung nach etwa 5 Minuten überlastet wurde. Die Framerate oszillierte danach um einen Wert von 17 fps. Die Verzögerung der Übertragung stieg sichtbar an, nachdem die Stausituation eingetreten ist. Die Verbindung

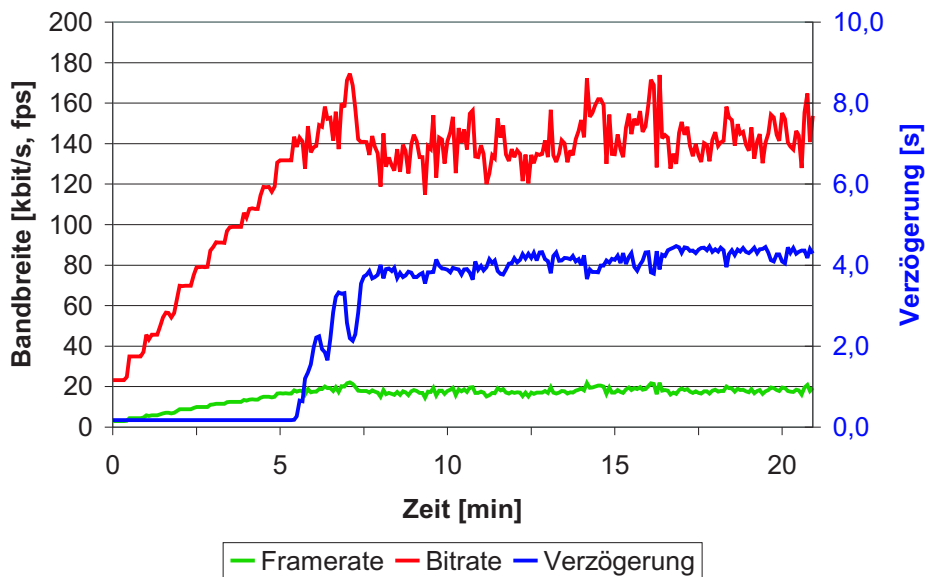


Abbildung 5.8.: Verzögerung, Framerate und Bandbreite auf der Verbindung von Hong Kong nach Cottbus

in der entgegengesetzten Richtung war weniger überlastet. Eine durchschnittliche Framerate von 100 fps (800 kbit/s) mit einer Verzögerung von etwa 400 ms wurde gemessen (siehe Abbildung 5.9).

Die Messungen belegen, dass zur Zeit Videokonferenzen über solche Entfernungen nur mit einer eingeschränkten Bandbreite und somit mit einer begrenzten Qualität durchgeführt werden können. Weiterhin ist die Qualität nicht unbedingt nur von der Entfernung, sondern auch von dem durch das Routing eingestellten Weg im Netz abhängig, wie das Beispiel Cottbus – Hong Kong zeigt. Die begrenzten Ressourcen reichen aber aus, um das vorgeschlagene Bandbreitenanpassungsschema anwenden zu können, da eine Bandbreite von über 10 fps gemessen wurden.

### Jitter

Aus den gemessenen Verzögerungen wurde außerdem der Jitter, die Verzögerungsschwankung, abgeleitet. Der durchschnittliche Jitter von Cottbus nach Hongkong betrug 36 ms, von Hong Kong nach Cottbus etwa 34 ms. Diese Werte wurden bei der Pufferung der Daten beim empfangenden Gruppenordinator berücksichtigt, um den Jitter für die lokalen Empfänger wieder zu

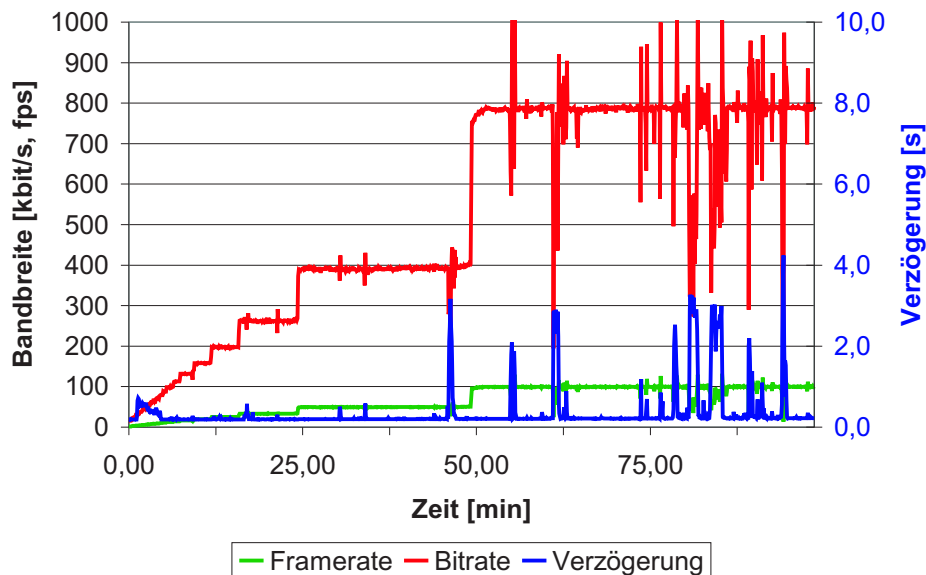


Abbildung 5.9.: Verzögerung, Framerate und Bandbreite auf der Verbindung von Cottbus nach Hong Kong

reduzieren. Die durch die Pufferung entstehenden Gesamtverzögerungen stellen kein Problem dar, da sie trotzdem unter 400 ms liegen.

### Bandbreitenanpassung

Die Bandbreitenanpassung dient dazu, auf Stausituationen im Netzwerk zu reagieren. Um die Wirkung, Stabilität und Geschwindigkeit des implementierten Algorithmus zu testen und zu bewerten, wurden verschiedene Situationen einer Konferenzgruppe simuliert. Es wurde jeweils eine lokale Gruppe von vier Teilnehmern angenommen, die nach dem Auswahlverfahren unterschiedliche Prioritäten von anderen entfernten Gruppen erhalten hatten, das heißt unterschiedliche Stimmanteile. Im Fall (a) gehört der aktuelle Sprecher zu der Gruppe, im Fall (b) nicht. Angelehnt an die Bandbreitenmessung im vorhergehenden Abschnitt liegt die anfängliche Bandbreite bei etwa 19 fps. Weiterhin wurde angenommen, dass nach etwas 35 Sekunden die Bandbreite kurzzeitig auf etwa 5 fps absinkt. Dies kann beispielsweise durch einen parallelen TCP-Strom ausgelöst worden sein. Danach wird der alte Wert von 19 fps wieder erreicht. Zu beachten ist, dass sich die Simulation nur auf eine Verbindung zu *einem* entfernten Gruppenkoordinator bezieht.

Im Fall (a) behandelt die Anpassungsprozedur alle ausgewählten Nichtsprecher gleichberechtigt, wenn die Bandbreite ausreichend hoch ist, um sie mit mindestens 3 fps zu übertragen. Reicht die Bandbreite nicht aus, um alle Teilnehmer zu übertragen, dürfen nur die Teilnehmer mit der höchsten Priorität übertragen werden. In dieser Simulation steigt die Priorität von Teilnehmer  $P_1$  zu  $P_4$  an. Die anfängliche Bandbreite von 19 fps ist ausreichend, um al-

Kein Sprecher in der Gruppe

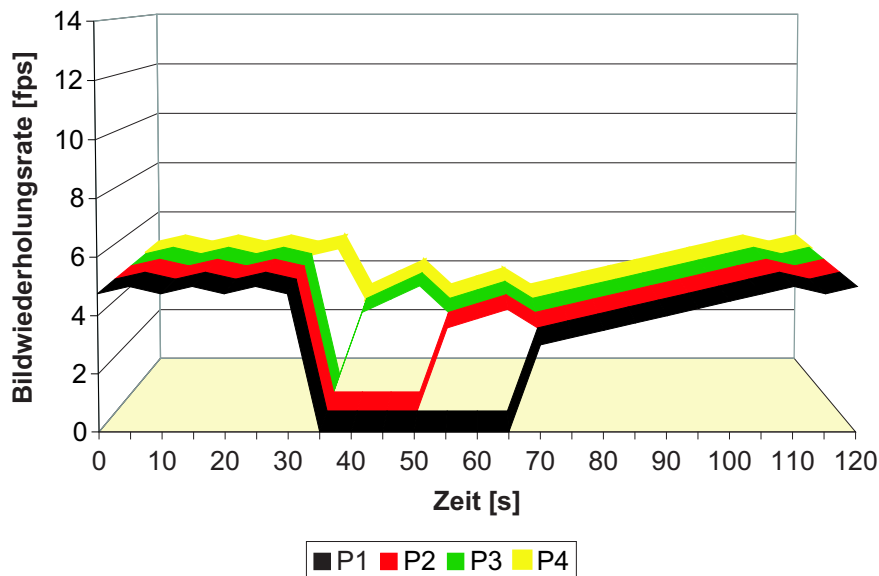


Abbildung 5.10.: Bandbreitenanpassung ohne einen Sprecher in der Gruppe

le Teilnehmer  $P_1$  bis  $P_4$  zu übertragen (siehe Abbildung 5.10). Somit teilen sich alle vier Teilnehmer die verfügbare Bandbreite gleichmäßig auf. Nachdem die Bandbreite nach etwa 35 Sekunden absinkt, kann nur noch Teilnehmer  $P_4$  übertragen werden, da die verfügbare Bandbreite nur noch für den Videostrom eines Nichtsprechers ausreicht. Die Anpassungsprozedur erhöht die Sendeleistung wieder Schritt für Schritt, nachdem die simulierte TCP-Übertragung beendet ist. Nach einer Weile werden alle ausgewählten Teilnehmer wieder gesendet:  $P_3$  nach 40 Sekunden,  $P_2$  nach 55 Sekunden und  $P_1$  nach 70 Sekunden.

Im Fall (b) wird der Sprecher  $Sp$  bevorzugt behandelt (siehe Abbildung 5.11). Wenn die Bandbreite nach 35 Sekunden auf 5 fps sinkt, kann nur noch der Sprecher mit der bestmöglichen Bildrate übertragen werden. Nach dem

Sprecher in der Gruppe

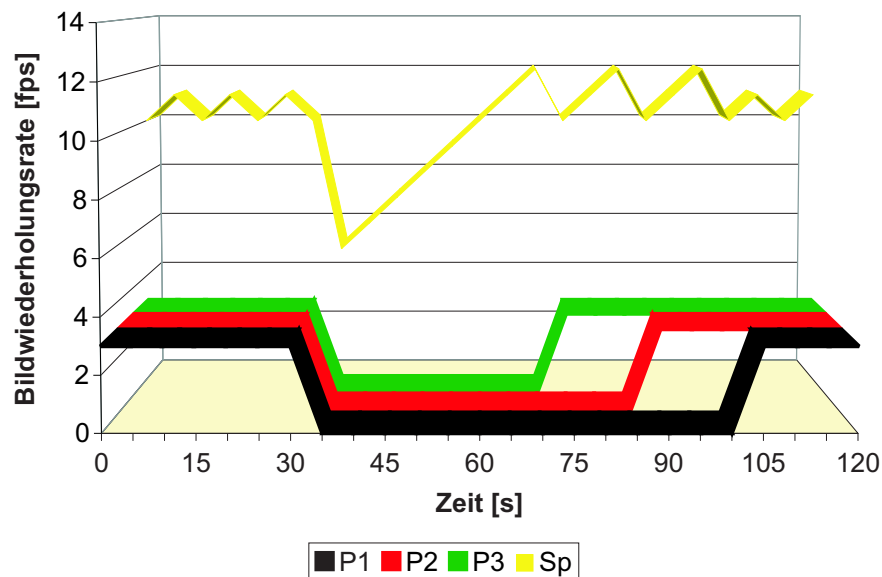


Abbildung 5.11.: Bandbreitenanpassung mit einem Sprecher in der Gruppe

Bandbreitenengpass steigt die Sendeleistung wieder kontinuierlich aufgrund der Bandbreitenmessung an. Die verfügbare Bandbreite ist nach 75 Sekunden wieder ausreichend, um den Sprecher und zumindest einen Nichtsprecher, im Beispiel *P1*, zu übertragen (13 fps). Somit wird die Framerate des Sprechers auf 10 fps und die des Nichtsprechers mit der höchsten Priorität auf 3 fps gesetzt. Die anderen Nichtsprecher *P2* und *P3* werden erst wieder übertragen, wenn die verfügbare Bandbreite die entsprechenden Werte von 16 fps und 19 fps erreicht.

Prioritäten- und Sprecherwechsel

Abbildung 5.12 zeigt die Bandbreitenanpassung für eine angenommene Änderung des Auswahlresultates zwischen zwei Teilnehmern und für einen Wechsel des Sprechers. Die verfügbare Bandbreite liegt bei einem etwa konstanten Wert von 16 fps für die gesamte Zeit. Nach 35 Sekunden wird die Sichtbarkeitstabelle geändert, wodurch sich die Priorität von Teilnehmer *P2* zu Teilnehmer *P1* hin verschiebt. Somit wird *P1* anstelle von *P2* übertragen. Nach 85 Sekunden wechselt das Rederecht vom Sprecher *Sp* zum Teilnehmer *P3*. *Sp* hat als Nichtsprecher eine geringere Priorität als die Teilnehmer *P1* und *P2*. Somit wird der ehemalige Sprecher nicht weiter übertragen.

Bewertung

Die Simulationen haben gezeigt, dass der Anpassungsmechanismus stabil

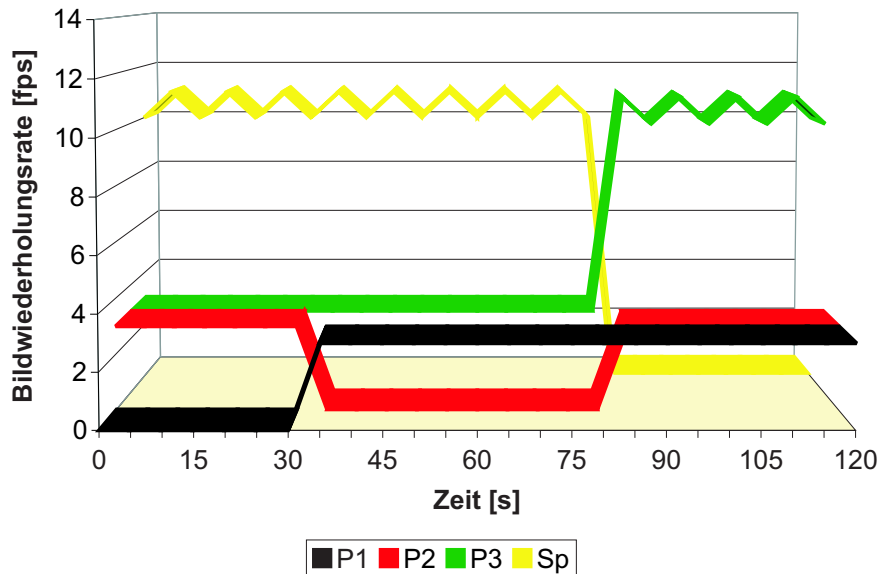


Abbildung 5.12.: Bandbreitenanpassung mit einer Änderung der Prioritäten zwischen den Teilnehmern  $P1$  und  $P2$  sowie einer Änderung des Sprechers von  $Sp$  zu  $P3$

arbeitet und den Zweck der Überlasterkennung und -behebung erfüllt. Die Reaktion auf eine extreme Stausituation innerhalb von 5 Sekunden ist hinreichend gut. Um sich aus einer Stausituation wieder zu erholen, vergehen etwa 60 Sekunden (siehe Abbildung 5.11). Die Geschwindigkeit für die Erhöhung der Sendeleistung scheint demnach langsam zu sein. Die Anpassungsgeschwindigkeit hängt jedoch maßgeblich von der Häufigkeit der Bandbreitenmessung ab. Je öfter die Bandbreite beim Empfänger gemessen wird, desto schneller kann eine Reaktion erfolgen. Dies bewirkt allerdings auch einen höheren Aufwand in der Signalisierung, da die Messergebnisse häufiger an den Sender geschickt und von diesem ausgewertet werden müssen. Aus diesem Grund ist das eingestellte Messintervall von 5 Sekunden ein guter Kompromiss zwischen hoher Anpassungsgeschwindigkeit und hohem Signalisierungsaufwand.

#### 5.4.4. Bandbreitenanpassung in BRAVIS

Im BRAVIS-System existiert kein Gruppenkoordinator, der die Videoströme den Bedingungen der globalen Verbindung anpasst. Diese Aufgabe überneh-

men die BRAVIS-Endsysteme selbst. Es liegt jedoch derselbe Ansatz zugrunde, wie in OCTOPUS: ein oder mehrere Teilnehmer aus einem lokalen Bereich, beispielsweise einem Campus-Netzwerk, bilden eine Untergruppe, die mit anderen Untergruppen kommuniziert (siehe Abbildung 3.9). Die Zusammensetzung der Untergruppen und die dazugehörigen *Ausgewählten Empfänger* sind allen Teilnehmern bekannt.

Das Problem dieses Ansatzes sind zwei spezielle Sonderfälle, bei denen die Unterteilung der Gruppe anhand der Netzwerkzugehörigkeit scheitert:

1. Die Gruppe besteht aus vielen kleinen Untergruppen, im Extremfall aus einem Teilnehmer, weil sich alle Teilnehmer an verschiedenen Standorten und nicht im selben regionalen Netz befinden.
2. In einer Untergruppe befinden sich sehr viele Teilnehmer. Im schlechtesten Fall befinden sich alle Teilnehmer in einer Untergruppe, weil sie sich alle am selben Standort befinden.

Im ersten Fall muss der Sender einen hohen Aufwand für die Skalierung der Videodaten, das Duplizieren der Audiodaten sowie das Verschicken beider Ströme betreiben. Weiterhin wird die Netzanbindung des Senders stark belastet. Im zweiten Fall müssen die AEs die Video- und Audiodaten sehr oft duplizieren und verschicken. Dadurch werden diese und das jeweilige lokale Netzwerk sehr stark belastet.

Die Lösung dieses Problems liegt in einer *künstlichen* Unterteilung in Untergruppen (siehe Abbildung 5.13). Eine Untergruppe kann dabei beispielsweise eine Gruppe von Teilnehmern sein, die regional nah beieinander liegen. Die Größe einer Untergruppe sollte dabei aber nicht willkürlich festgelegt werden. In ersten Überlegungen wird eine Anzahl von vier Teilnehmern pro Untergruppe angenommen, um einerseits den Aufwand für Sender und AEs möglichst gering zu halten, andererseits aber auch die Verzögerung der Datenübertragung zu minimieren. Genauere Untersuchungen werden dazu gegenwärtig im Rahmen einer Diplomarbeit durchgeführt.

Die Bandbreite, die Verzögerung und die Auslastung der Endsysteme müssen für den Aufbau einer optimalen Verteilungsstruktur berücksichtigt werden. Im Ansatz von Chu [28] werden durch periodisch durchgeführte aktive Messungen die Bandbreite und die Verzögerungen zwischen den Teilnehmern bestimmt,



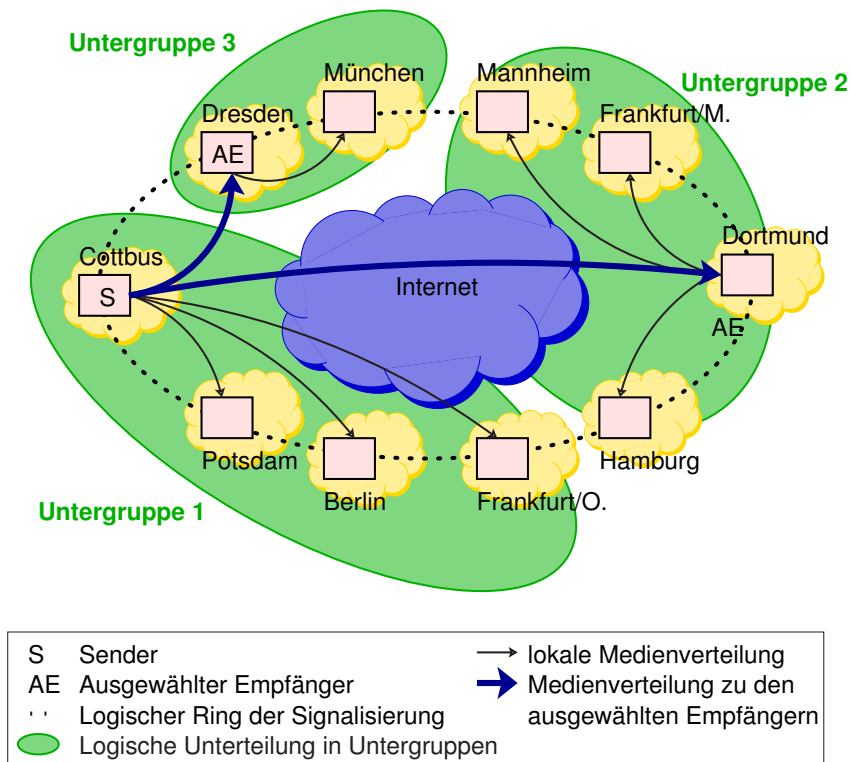


Abbildung 5.13.: Bildung von künstlichen Untergruppen in BRAVIS

durch den ein erheblicher Netzwerkverkehr verursacht wird. Die Last an in den Endsystemen wird nicht betrachtet.

BRAVIS berücksichtigt die Bandbreite des Netzwerkanschlusses eines Teilnehmers, wie zum Beispiel Modem, ISDN, DSL oder direkter Anschluss an das Internet. Die Einteilung wird während der Installation getroffen und kann später auch geändert werden. Somit werden aufwendige und netzbelastende Messungen vermieden. Aufgrund der Erfahrung, dass viele Netzbetreiber, wie beispielsweise auch der DFN-Verein, ihre Netzwerke in Bezug auf die Bandbreite überdimensionieren, ist dieser Ansatz in BRAVIS ausreichend. Eine minimale Verzögerung bei der Übertragung der Mediendaten wird in BRAVIS durch die Nutzung des logischen Ringes, auf dem sich die Teilnehmer befinden, erreicht. Durch den Mechanismus zur Realisierung der minimalen Tokenumlaufzeit im GCP<sub>IP</sub> (siehe Abschnitt 4.3.2) befinden sich benachbarte Teilnehmer in einem minimalen Abstand.

Eine Besonderheit im BRAVIS-Ansatz stellt die Berücksichtigung der Auslastung und der Leistungsfähigkeit der Endsysteme beim Aufbau einer optimierten Verteilungsstruktur dar. Leistungsschwache Endsysteme können einen Flaschenhals darstellen, wenn sie als ausgewählte Empfänger in der Konferenz arbeiten. Aus diesem Grund bestimmt das QoS-Management anhand der Leistungsparameter der Teilnehmerendsysteme unter Berücksichtigung der aktuellen Auslastung die ausgewählten Empfänger, die die Verteilung an die Teilnehmer der jeweiligen Untergruppe übernehmen.

Der Verteilungsmechanismus in BRAVIS wurde zeitlich parallel zu den Entwicklungen des *Overlay Multicast* [8, 138, 182] erarbeitet (siehe auch Abschnitt 2.6). Bisherige Overlay Multicast Ansätze orientieren sich jedoch wie IP-Multicast auf eine empfängerbezogene Verteilung der Daten, bei der der Sender die Empfänger nicht kennen muss. Diese Ansätze können deshalb nicht direkt in den mit BRAVIS verfolgten Ansatz integriert werden. Der BRAVIS-Ansatz geht grundsätzlich von einer Verteilung der Audio- und Videodaten in geschlossenen Konferenzgruppen aus.

Bandbreitenzuweisung

Wie der Gruppenkoordinator im OCTOPUS messen die ausgewählten Empfänger die Bandbreite, die sie von den anderen Teilnehmern empfangen haben. Aus den einzelnen Messergebnissen können die ausgewählten Empfänger die Bandbreite berechnen, die zwischen den Untergruppen zur Verfügung steht. Die gemessene Gesamtbandbreite wird allen Teilnehmern der sendenden Untergruppe über  $GCP_{IP}$  mitgeteilt. Aufgrund dieser Information können die Teilnehmer einer Untergruppe berechnen, wie viel Bandbreite jedem einzelnen zur Verfügung steht. Diese Berechnung berücksichtigt eine eventuelle Abstimmung der Teilnehmer einer Untergruppe und gegebenenfalls einen Sprecher in der Untergruppe. Der Bandbreitenzuweisungsalgorithmus aus dem OCTOPUS wird hier übernommen, mit dem Unterschied, dass in BRAVIS eine dezentrale Berechnung in den Endsystemen erfolgt. Auch hierzu werden gegenwärtig Untersuchungen durchgeführt, die die Leistungsfähigkeit des entwickelten Ansatzes untermauern.

### 5.4.5. Skalierung von Videoströmen

Die Skalierung von Videoströmen dient vor allem dazu, den die Datenmenge des Videostroms so zu reduzieren, dass die Empfänger und das Netzwerk nicht

überlastet werden. Videoströme lassen sich zum einen empfängerbezogen und zum anderen senderbezogen skalieren. Bei der empfängerbezogenen Skalierung wird der Videostrom beim Empfänger skaliert, so dass er von diesem verarbeitet werden kann. Überlastungen des Netzes werden hier nicht berücksichtigt, wodurch ein Einsatz in globalen öffentlichen Netzen ungünstig ist. Bei einer senderbasierten Skalierung passt der Sender den Videostrom an die Bedürfnisse der Empfänger und des Netzes an. Die Empfänger benötigen keine Skalierungsfunktionen, wodurch dort Rechenkapazitäten gespart werden können. Dafür wird jedoch der Sender stark belastet. Eine Kombination aus beiden Skalierungsmechanismen kann die Belastung auf Sender und Empfänger gleichmäßig verteilen.

In einem geschlossenen Videokonferenzsystem ist jeder Teilnehmer gleichzeitig Sender und Empfänger, wobei die Möglichkeit der individuellen Anpassung der Dienstgüte an die Bedürfnisse und Ressourcen der einzelnen Teilnehmer besteht. Eine senderbezogene Skalierung kann hier sinnvoll eingesetzt werden. Dadurch erhält jeder Nutzer die für ihn bestmögliche Qualität in einer Videokonferenz. Dies erfordert jedoch einen geeigneten Video-Codec, der eine entsprechende Skalierung zulässt. Bei einer hierarchischen Verteilung der Videodaten können die Videoströme außerdem von einer Ebene zur nächsten Ebene skaliert werden. Somit darf der Codec keine hohen Ansprüche an die Leistung der Endsysteme stellen, um Kommunikationsengpässe zu vermeiden.

Layered-Coding-Verfahren sind eine Möglichkeit der senderbezogenen Skalierung, die die zu erwartende Heterogenität in globalen öffentlichen Netzwerken unterstützen. Jedoch können Codecs mit räumlicher Unterteilung der Kompressionsebenen zur Zeit nicht zufrieden stellend eingesetzt werden, da keine Hardwarerealisierungen existieren und die Softwarecodecs eine zu hohe Leistung der Endsysteme erfordern. Aus diesem Grund ist es günstiger, eine temporale Unterteilung zu nutzen, um die Videoströme zu skalieren.

*Layered Coding*

Videoströme können sowohl als unkomprimierte Daten oder als komprimierte Daten temporal skaliert werden. Die Skalierung des unkomprimierten Stromes erfordert aber hohe Rechenkapazitäten, da jeder einzelne Strom separat kodiert werden muss (siehe Abbildung 5.14a). Günstiger ist die Skalierung des komprimierten Videostroms mittels einem Frame-Dropping-Filter, bei dem einzelne Bilder aus dem fortlaufenden Strom entfernt werden, um

*Skalierungsschemen*

so die gewünschte Sendeleistung zu erreichen (Abbildung 5.14b) ). Der Nach-

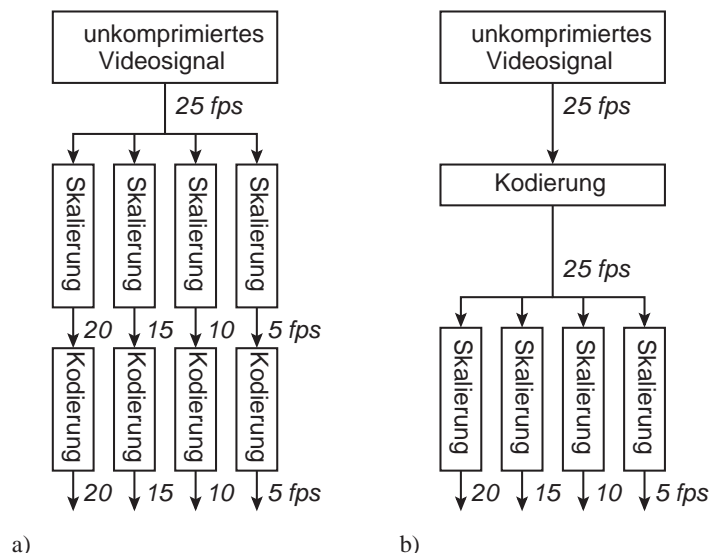


Abbildung 5.14.: Skalierungsschemen für Videostreams

teil des Einsatzes eines Frame-Dropping-Filters ist die beschränkte Auswahl der Kompressionsmethoden. Da komprimierte Bilder aus dem Videostream entfernt werden, können *Interframe* kodierende Codecs nur bedingt genutzt werden. Das Entfernen beliebiger Frames hat die Bildung eines Schweißs hinter sich bewegenden Objekten, Rauschen und Flackern des Bildes bis hin zur völligen Zerstörung des Stroms zur Folge, da auch Referenzen auf nachfolgende Bilder verworfen werden könnten. Aus diesem Grund müssen anfangs nur die P-Frames, dann die B-Frames und zum Schluss die I-Frames aus einem Interframe-kodierten Strom, wie zum Beispiel MPEG oder H.26x, entfernt werden. Durch dieses Verfahren wäre jedoch eine Feinabstimmung der Sendeleistung schwierig. Zudem gestaltet sich die Suche nach den entsprechenden Frames in einem kontinuierlichen Strom als sehr aufwendig. Kompressionsverfahren, die wie M-JPEG nur I-Frames produzieren, sind besser für die Skalierung mittels Frame-Dropping-Filter geeignet, da keine Abhängigkeiten zwischen nachfolgenden Bildern bestehen. Die beiden Skalierungsvarianten mit komprimierten und unkomprimierten Videodaten wurden im OCTOPUS-System getestet und bewertet.

Unkomprimierte  
Skalierung

Die Gruppenkoordinatoren des OCTOPUS übernehmen hierbei die Aufgabe der dynamischen Skalierung der Videostreams in Abhängigkeit ihrer Priorität,

der globalen Gruppe und der verfügbaren Bandbreite. Da in den lokalen Gruppen komprimierte Videoströme übertragen werden, müssen diese zunächst dekomprimiert werden. Erst danach wird das Videosignal skaliert und neu komprimiert. Dieser Ansatz entspricht dem eines intelligenten Videogateways, das erstmals in [144] realisiert wurde. Nach der Integration dieses Ansatzes in den OCTOPUS-Gruppenkoordinator wurden Tests auf einer SUN ultra 10 Workstation (433MHz) mit der X Image Library (XIL) und einem M-JPEG Codec ohne Hardwareunterstützung für Videoströme mit unterschiedlichen Eingangs- und Ausgangsframeraten durchgeführt. Es wurde das Kompressionsverfahren M-JPEG für die Messung gewählt, weil dieser Codec weniger komplex als MPEG-1, 2, 4 oder H.261 ist und ohne Bewegungserkennung arbeitet. Die Last, die bei dieser Messung durch die Dekodierung eines Videostroms mit 25 fps erzeugt wurde, beträgt weniger als 1% der Prozessorkapazität. Die Auslastung des Prozessors für die Kodierung eines Videostroms in Prozent entspricht etwa der Sendeleistung in fps. Beispielsweise erzeugt ein Gruppenkoordinator eine Last von 35% der Prozessorleistung, wenn zwei Videoströme mit je 15 fps und 20 fps übertragen werden sollen. Unter der Annahme, dass die Bandbreite zwischen den Gruppen etwa 17 fps beträgt, kann somit ein Gruppenkoordinator nur bis zu sechs entfernte Gruppen unterstützen, ohne selbst zum Flaschenhals der Kommunikation zu werden. Diese Anzahl gilt, wenn sich nur ein Teilnehmer in der lokalen Gruppe befindet. Senden zwei Teilnehmer aus einer lokalen Gruppe heraus, dann kann der Gruppenkoordinator nur maximal drei entfernte Gruppen bedienen. Dies ist nicht zufriedenstellend, besonders wenn die Bandbreite zu den entfernten Gruppen höher als 17 fps ist. Unter Nutzung von komplexeren Codecs, wie MPEG-1 bis 4 oder H.26x, würde die Anzahl der unterstützten Gruppen weiter sinken.

Der Einsatz einer Frame-Dropping-Komponente für Videoströme im Gruppenkoordinator vermeidet hohe Prozessorauslastungen. Messungen mit einem M-JPEG komprimierten Videostrom haben gezeigt, dass weniger als 0,5% der Prozessorleistung benötigt wurden, um diesen zu vier verschiedenen entfernten Gruppen mit 25 fps zu übertragen. Somit können 100-mal mehr Gruppen mit einem Gruppenkoordinator verbunden werden als mit der in [144] vorgeschlagenen Methode. Aus diesem Grund wurde für OCTOPUS M-JPEG als Videostandard gewählt.

<i>Komprimierte Skalierung</i>
------------------------------------

Das Übertragungskonzept von OCTOPUS wurde in BRAVIS übernommen und in die Endsysteme integriert. Somit besteht auch hier das Problem der Skalierung der Videoströme. Aufgrund der genannten Messungen des Gruppenkoordinators ist es günstig, beim Sender eine Skalierung des komprimierten Videostroms durchzuführen. M-JPEG kommt deshalb auch in BRAVIS zum Einsatz.

Bandbreiten-  
problematik

Durch die fehlende Interframekodierung in M-JPEG besteht jedoch ein höherer Bandbreitenbedarf. Messungen haben ergeben, dass M-JPEG bei 25 fps für ein CIF-Bild etwa 1 Mbit/s und für QCIF etwa 500 kbit/s benötigt<sup>4</sup>. Dies bedeutet unter der Annahme von einem Sprecher mit einem CIF-Bild und 20 weiteren Teilnehmern mit einem QCIF-Bild ein Gesamtvolumen von etwa 11 Mbit/s. Reduzieren lässt sich dieses Volumen durch eine Verringerung der Bildwiederholungsfrequenz und durch den in diesem Abschnitt vorgeschlagenen Auswahlmechanismus. Des Weiteren können in der Zukunft bessere Kodierungstechniken eingesetzt werden. Messungen mit einem Wavelet-basierten JPEG2000-Kodierer [81, 39] für Standbilder haben ergeben, dass sich die Bitrate auf etwa 10 % der Bitrate von M-JPEG reduzieren lassen würde. Jedoch ist die Rechenleistung der heutigen PCs und Workstations noch nicht ausreichend, um damit Videoströme in Echtzeit zu komprimieren<sup>5</sup>. Mit der rasanten Steigerung der Leistungsfähigkeit der Endsysteme soll es jedoch in Zukunft möglich sein, Skalierungsmechanismen für unkomprimierte Videoströme und Layered-Coding-Techniken zu verwenden.

### 5.5. Zusammenfassung

In diesem Kapitel wurde die Audio- und Videoübertragung in geschlossenen Videokonferenzsystemen untersucht. Es wurde gezeigt, dass hierarchische Topologien zur Verteilung der Mediendaten günstig in Bezug auf die Skalierbarkeit, der Verzögerung, der Bandbreitenausnutzung und vor allem der Unterstützung heterogener Umgebungen ist. Jedoch werden sie bisher kaum eingesetzt. Für das Videokonferenzverbundsystem OCTOPUS und das Videokonferenzsystem BRAVIS wurden hierarchische Verteilungsstrukturen zur Übertra-

---

<sup>4</sup>Im Vergleich dazu benötigen CellB und H.261 etwa halb so viel Bandbreite.

<sup>5</sup>Auf einem Pentium 4 mit 1,6 GHz konnten maximal vier Bilder pro Sekunde komprimiert werden.

gung der Video- und Audiodaten entwickelt. Darauf aufbauend wurden Konzepte für ein QoS-Management erprobt, die die Qualität der Video- und Audioübertragung über globale Strecken auf einem konstant hohen Level halten. Der Ansatz für die Qualitätssicherung der Videodaten basiert hierbei auf einem die Sendeleistung anpassenden Mechanismus, der durch eine abstimmungsbasierte Zuweisung der verfügbaren Ressourcen ergänzt wird. Messungen haben die Anwendbarkeit dieses Ansatzes untermauert. Abschließend wurden Möglichkeiten zur Skalierbarkeit von Videoströmen analysiert. Es hat sich gezeigt, dass die heutigen Endsysteme noch nicht in der Lage sind, zeitgleich mehrere Videoströme mit unterschiedlichen Leistungsparametern zu komprimieren. Somit ist eine Skalierung von komprimierten Videoströmen erforderlich, die nur mit Intraframekodierung wie M-JPEG möglich ist. Mit den Komponenten des QoS-Managements und der Skalierung, die in OCTOPUS und BRAVIS integriert wurden, ist erstmals ein Videokonferenzsystem in der Lage, geschlossene Gruppen innerhalb einer heterogenen Systemumgebung zu unterstützen.





---

## 6. Konzepte der Implementierung

---

Nachdem in den vorangegangenen Kapiteln die Konzepte für die Gestaltung von geschlossenen Mehrteilnehmer-Videokonferenzen für globale Netze beschrieben wurden, folgen in diesem Kapitel die wichtigsten Implementierungskonzepte für Umsetzung der beiden Lösungen OCTOPUS und BRAVIS. Es wird die Wahl der Betriebssystem- und der Programmierplattform erläutert, auf der die hierarchischen Ansätze beruhen. Programmiertechnische Änderungen an bereits bestehenden lokalen Systemen, die für den Betrieb im globalen Verbund notwendig sind, werden hier ebenfalls diskutiert.

### 6.1. Auswahl der Systemplattform

Die Idee von OCTOPUS basierte auf der Idee, die Videokonferenzsysteme COVIS und HKUST, die am Lehrstuhl Rechnernetze der BTU Cottbus und an der Hong Kong University of Science and Technology entwickelt wurden, über ein globales Netz zu verbinden. Beide Systeme wurden unter dem UNIX-Betriebssystem SUNSolaris 2.6 implementiert. Um einzelne Module der vorhandenen Systeme auch im OCTOPUS-Gruppenkoordinator nutzen zu können, wurde SUNSolaris für die Implementierung des Videokonferenzsystems OCTOPUS gewählt. Komponenten wie die Floorkontrolle, das Gruppenmanagement sowie die Audio- und Videoverarbeitung konnten für OCTOPUS übernommen werden.

Da die Funktionalitäten von BRAVIS sehr ähnlich mit COVIS sind, konnten einige Module wiederverwendet werden, wie zum Beispiel die Floorkontrolle, das Gruppen- und das QoS-Management. Einige Teile der Video- und Audioübertragung konnten vom OCTOPUS-System übernommen werden. Gleichzeitig bestand die Motivation, BRAVIS einer größeren Nutzergemeinde, dem PC-Markt, zur Verfügung zu stellen. Da das Betriebssystem Linux und SUNSolaris viele Gemeinsamkeiten haben, wurde vorerst eine Linux-Variante von

BRAVIS entwickelt. Speziell die Audio- und Videoverarbeitung wurde so erweitert, dass es auch unter beiden Betriebssystemen lauffähig ist. Eine Windows-Portierung wäre jedoch sehr aufwendig gewesen, da alle Module komplett neu entwickelt werden müssten.

POSIX-C

Als Programmiersprache wurde *C* aufgrund dessen Leistungsfähigkeit und der bereits vorhandenen Implementierungen von COVIS gewählt. Die modulare Implementierung von COVIS wurde in OCTOPUS und BRAVIS übernommen. Es werden die Möglichkeit der Interprozesskommunikation der UNIX-Betriebssysteme genutzt. *Message Queues* dienen dem synchronen Datenaustausch zwischen einzelnen Prozessen. Ein *Shared Memory (ShM)* wird für den Zugriff auf gemeinsame Datenbestände genutzt, wobei der gleichzeitige Zugriff durch *Semaphore* geregelt wird. Zur Parallelisierung von Abläufen dienen POSIX-Threads [125, 151, 152]. Der Zugriff auf gemeinsame Speicherbereiche innerhalb eines Prozesses wurde durch *Mutexe* geregelt.

Im Weiteren wird auf Besonderheiten der Implementierung der Ansätze OCTOPUS und BRAVIS eingegangen.

## 6.2. Gruppenkoordinator im OCTOPUS

Der Gruppenkoordinator im OCTOPUS-Ansatz stellt, wie in Abschnitt 3.1 erläutert, Gatewayfunktionen zwischen einem lokalen Videokonferenzsystem und einer global einheitlichen Schnittstelle zur Verfügung. Die in Abbildung 3.3 gezeigte Architektur des Gruppenkoordinator wurde für das lokale ATM-basierte Videokonferenzsystem COVIS<sup>1</sup> implementiert und getestet. Sowohl COVIS als auch der dazugehörige Gruppenkoordinator nutzen das lokale ATM-Netzwerk, das drei SUN Ultra Sparc 10 Workstations und eine SUN Sparc 20 Workstation über einen ASX200 Fore Switch verbindet. Für die globale Vernetzung der Gruppenkoordinatoren wurde das am Lehrstuhl vorhandene Fast-Ethernet-LAN mit 100 Mbit/s verwendet.

### 6.2.1. Signalisierung

GCP

GCP<sub>IP</sub> wurde wie GCP<sub>ATM</sub> nach dem Server-Modell-Prinzip implementiert,

---

<sup>1</sup>Ein Gruppenkoordinator für das HKUST-System wurde während der Projektphase aufgrund mangelnder Kapazitäten nicht implementiert.

das eine einfache und geradlinige Umsetzung gestattet. Da  $GCP_{IP}$  wie  $GCP_{ATM}$  auch nur geringe Datenmengen verarbeiten muss, ist das Server-Modell ausreichend.

Die Teilnehmer im  $GCP_{ATM}$  wurden in einem Shared Memory gespeichert. Shared Memories sind jedoch nicht dynamisch erweiterbar. Aus diesem Grund ist die Anzahl der Teilnehmer im  $GCP_{ATM}$  statisch vorgegeben. Im  $GCP_{IP}$  wurde auf die Verwendung eines Shared Memory verzichtet. Stattdessen wurde eine dynamische Liste implementiert, die nicht in ihrer Größe beschränkt ist.

Für die Kommunikation mit den über GCP liegenden Modulen werden Message Queues eingerichtet. Die `gcpip_queue` und die `gcpatm_queue` dienen als Eingangswarteschlange des  $GCP_{IP}$  beziehungsweise  $GCP_{ATM}$ . Über diese erfolgt der Austausch der GCP-Dienstprimitive. Die GCP-PDUs werden über IP- beziehungsweise ATM-Sockets als Netzwerkschnittstelle ausgetauscht. Um zu vermeiden, dass der GCP-Protokollautomat auf zwei unterschiedliche Eingangswarteschlangen, *Message Queue* und *Sockets*, reagieren muss, wurde die Schnittstelle zum IP- beziehungsweise ATM-Netz als eigenständiger Prozess implementiert. Dieser Prozess empfängt Pakete aus dem Netz und übergibt sie an die GCP *Message Queue*. Somit braucht nur eine Eingangswarteschlange im GCP berücksichtigt werden (siehe Abbildung 6.1).

Message Queues

Die Nutzung von Message Queues ist dahingehend nachteilig, da sie nur in *System V* Betriebssystemen angeboten werden, wie beispielsweise SUNSolaris und Linux. Diese Mechanismen müssen bei einer Überführung in andere Betriebssysteme wie beispielsweise Windows ersetzt werden.

$GCP_{ATM}$  benutzt nur das Signal `SIGALRM`, um Timeouts bekannt zu machen. Somit können aber nicht unterschiedliche Ereignisse unterschieden werden, die zu dem Timeout geführt haben. Dies wurde jedoch in  $GCP_{IP}$  notwendig, um einerseits den Ausfall eines Tokenhalters zu ermitteln und andererseits das Ausbleiben einer Bestätigung der Tokenweitergabe zu erkennen. Deshalb beruhen die Timer im  $GCP_{IP}$  auf unterschiedlichen Signalen, die durch die vom Betriebssystem angebotene Funktion `timer_create()` angeboten werden. Somit können bei bestimmten Timeouts spezielle Routinen (*signal handler*) gestartet werden.

Timer

Im Gruppenkoordinator übergibt das  $GCP_{ATM}$  die Daten an das *Local Group Communication Module (LGCM)* und das  $GCP_{IP}$  an das *Global Group Com-*

LGCM und GG-CM

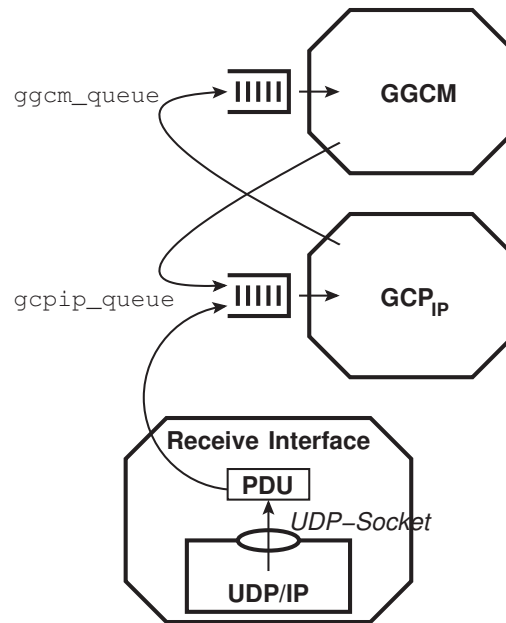


Abbildung 6.1.: Warteschlangenkonzept des GCP<sub>IP</sub>

*munication Module* (**GGCM**) (siehe Abbildung 3.3 in Abschnitt 3.1.3). **LGCM** und **GGCM** können abhängig von der zu sendenden Nachricht sowohl an das GCP<sub>ATM</sub> als auch an das GCP<sub>IP</sub> Nachrichten übergeben, wodurch die Kopplung zwischen der lokalen und der globalen Signalisierung realisiert wird. Ursprünglich war vorgesehen, dass der Verbund durch das *Group Coordinator Information Module* (**GCIM**) implementiert wird. Jedoch würden sich dadurch unnötige Wartezeiten durch das Schreiben und Lesen in die entsprechenden Warteschlangen ergeben. Deshalb wurde die direkte Kommunikation zwischen **LGCM** und **GGCM** implementiert.

**LGCM** und **GGCM** fungieren als Gateway der Signalisierungsdaten. Dabei werden im **LGCM** Datenpakete der lokalen Konferenz unter Berücksichtigung der in Tabelle 6.1 aufgestellten Regeln in das global gültige OCTOPUS-Format konvertiert.

Das **GGCM** enthält im Falle eines lokalen COVIS-Systems nur eine geringe Konvertierungsfunktionalität (siehe Tabelle 6.2). Gewöhnliche Datenpakete werden direkt an das lokale GCP<sub>ATM</sub> übergeben. Im Falle des HKUST leitet der Gruppenkoordinator die Datenpakete nicht an die Teilnehmer weiter, sondern muss selbst auf diese Nachrichten entsprechend reagieren.

Aktion	COVIS	OCTOPUS
Beitritt eines neuen lokalen Teilnehmers	JOINntf (Host)	DATreq (ADD_LOCAL_PARTICIPANT, Host)
Verlassen eines lokalen Teilnehmers	LEAVEntf (Group)	DATreq (REM_LOCAL_PARTICIPANT, Group)
Einladen einer neuen Gruppe	DATind (ADD_GLOBAL_GROUP, Host)	JOINreq (Host)
Austreten aus der globalen Konferenz	DATind (REM_GLOBAL_GROUP)	LEAVEreq
Signalisierungsdaten (z. B. Floorkontrolle oder QoS-Management)	DATind (Floor QoS)	DATreq (Floor QoS)

Tabelle 6.1.: Schnittstellenbeschreibung der OCTOPUS-Signalisierung (gesendete Daten)

Aktion	OCTOPUS	COVIS
Beitritt einer neuen Gruppe	JOINntf (Group)	DATreq (SET_REMOTE_PARTICIPANTS)
Verlassen eines lokalen Teilnehmers	LEAVEntf (Group)	DATreq (REMOVE_GLOBAL_GROUP_SUCCESS, Group)
Signalisierungsdaten	DATind (data)	DATreq (data)

Tabelle 6.2.: Schnittstellenbeschreibung der OCTOPUS-Signalisierung (empfangene Daten)

### 6.2.2. Medienübertragung

Im lokalen Bereich wurde unter Anwendung des COVIS ein ATM-Netzwerk mit dem AAL-5 genutzt. AAL-5 bietet einen senderbasierten Multicast-Dienst an. Für Echtzeitkommunikation wird jedoch AAL-2 bevorzugt und empfohlen, das aber nicht von den verwendeten ATM-Netzwerkkarten und ATM-Programmierbibliotheken unterstützt wird. Für den Transport der Echtzeitdaten wurde das proprietäre Protokoll MCM-TP [89] über AAL-5 genutzt, welches als lokale Schnittstelle für den Gruppenkoordinator dient. Es enthält keine Jitter-Reduzierung, da in einer lokalen ATM-Umgebung kaum ein Jitter auftritt.

Die globale Übertragung der Echtzeitdaten basiert auf dem Standardprotokoll RTP. Hierbei wurde zunächst versucht, den frei verfügbaren Quellcode

Wahl der RTP-Bibliothek

der Mbone-Tools VIC und VAT aus dem MASH Projekt [156] für die Zwecke des Gruppenkoordinators zu modifizieren, um gleichzeitig auch die grafische Schnittstelle zu nutzen und zu erweitern. Dies erwies sich jedoch als zu komplex, wodurch unnötig hohe Prozessorkapazitäten auf dem Gruppenkoordinator verbraucht wurden. Eine später angebotene eigenständige RTP-Softwarebibliothek des *University College London (UCL)* [164] wurde getestet. Jedoch wurden die Testdaten stark verzögert. Deshalb wurde für die Implementierung die RTP-Bibliothek von Lucent Technologies [164] genutzt. Die Verzögerungen der Echtzeitdaten war hier spürbar geringer als bei der UCL-Bibliothek und die Dokumentation der Funktionen der Bibliothek war ausführlicher.

Frame Dropping  
Filter

Die Komponente für die Umsetzung der Videoströme vom lokalen ATM-Netz zum Internet besteht aus einem Hauptprozess, der nur auf Verbindungsanfragen von anderen lokalen ATM-Teilnehmern wartet. Wird eine Verbindung von einem lokalen Teilnehmer zum Gruppenkoordinator aufgebaut, so wird dort ein Kindprozess erzeugt. Es existieren also so viele Kindprozesse, wie es Teilnehmer in der lokalen Konferenz gibt (siehe Abbildung 6.2). Jeder

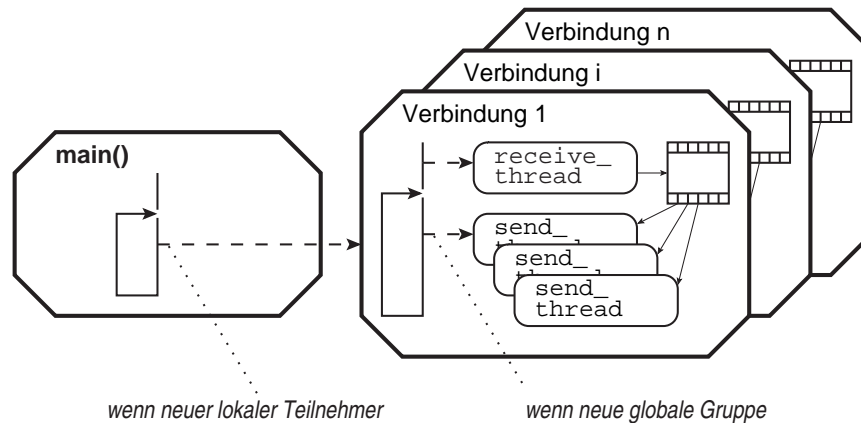


Abbildung 6.2.: Softwarearchitektur des ATM-RTP-Gateway für den Gruppenkoordinator

dieser Kindprozesse erzeugt einen `receive_thread`, der die Videodaten von der aufgebauten ATM-Verbindung empfängt und das komprimierte Bild zwischengespeichert. Alte Bilder werden dabei überschrieben. Nach jedem Beitritt einer neuen entfernte Gruppe in die Konferenz wird ein Thread `rtp_thread` für das Senden der zwischengespeicherten Videobilder zu dieser Gruppe über RTP

gestartet. Der Zugriff auf diesen gemeinsamen Speicher wird ebenfalls durch ein *Mutex*, eine Art Semaphor für Threads, gesichert. Das Senden der Daten über RTP kann langsamer und unabhängig vom Empfangen der Daten über ATM erfolgen, wodurch ein einfacher Frame Dropping Filter realisiert wurde. Es wird immer das letzte vom Empfängerthread empfangene komprimierte Bild über RTP verschickt. Da jedem Sendethread eine entfernte Gruppe zugeordnet wird, unterstützt dieses Konzept eine individuelle Anpassung der Sendeleistung.

### 6.3. Endsysteme im OCTOPUS

Ziel des OCTOPUS-Ansatzes war, dass die lokalen Endsysteme *nicht* oder nur unwesentlich verändert werden. Dies ist die grundlegende Voraussetzung für eine problemlose Integration des Ansatzes in andere lokale Systeme.

Die Module der Medienübertragung in COVIS wurden für den Einsatz von OCTOPUS nicht angepasst, da der Gruppenkoordinator als ein Teilnehmer der lokalen Gruppe interpretiert wird. Dass dieser Teilnehmer mehrere Video- und Audioströme verschickt, spielt aus Sicht der Implementierung der Endsysteme keine Rolle.

Eine tief greifende Änderung des Videoempfangsmoduls hätte die Umsetzung der lokalen Skalierung mit einer Priorisierung der globalen Ströme bewirkt. Zum einen hätte das Transportprotokoll MCM-TP derart geändert werden müssen, dass die Priorität des Videostroms mit übertragen wird. Zum anderen hätte das Modul für die Videobildanzeige die übertragene Priorität auswerten und die Systemressourcen entsprechend aufteilen müssen. Dies würde sich dann auch auf das in COVIS integrierte QoS-Management auswirken. Somit wurde vorerst auf die Umsetzung der lokalen Skalierung unter Berücksichtigung einer höheren Priorität globaler Videoströme verzichtet, da die Änderungen des lokalen Videokonferenzsystems COVIS zu weitreichend sind. Das in COVIS integrierte QoS-Management bleibt somit erhalten, so dass alle Videoströme, bis auf den des Sprechers, mit der gleichen Priorität angezeigt werden.

Die Analyse der COVIS-Implementierung hat gezeigt, dass das Gruppenmanagement durch das GCP<sub>ATM</sub> übernommen wurde, da dort hauptsächlich die

Medienübertragung
-------------------

Konferenzsteuerung
--------------------

Gruppenmitglieder in einem Shared Memory verwaltet werden. Andere Module greifen nur lesend auf diesen Speicher zu. Das Gruppenmanagement-Modul **GMM** dient lediglich als Schnittstelle zum Multiplexen der vom  $GCP_{ATM}$  übergebenen SDUs zum Floorkontroll- und dem QoS-Modul. Die Erweiterung des  $GCP_{ATM}$  für die Verwaltung der globalen Gruppenmitglieder im OCTOPUS hätte gravierende Änderungen des Protokollablaufs des  $GCP_{ATM}$  zur Folge. Aus diesem Grund wurde die Architektur des COVIS um das Global Group Management Module (**GGMM**) erweitert. Das ursprüngliche **GMM** wurde so modifiziert, dass es entsprechende Datenpakete vom  $GCP_{ATM}$  nun auch an das **GGMM** übergeben kann. Die nur für das OCTOPUS gültigen Funktionen von COVIS, wie beispielsweise das Voting und dessen graphische Nutzerschnittstelle, werden vom **GGMM** gesteuert.

Die Floorkontrolle arbeitete im COVIS in der Weise, dass eine Änderung der Warteschlange oder eine Anfrage an die Warteschlange an alle Teilnehmer gleichzeitig übertragen wurde. Im Videokonferenzverbund OCTOPUS muss jedoch eine solche Anfrage zuerst nur an den Gruppenkoordinator geschickt werden, damit dieser die Anfrage global verteilt (siehe Abschnitt 3.1.4). Aus diesem Grund musste die Floorkontrolle sehr geringfügig modifiziert werden. Das Empfänger-Flag des  $GCP_{ATM}$  darf im Betrieb mit OCTOPUS nicht auf `all_hosts` stehen sondern muss auf den Hostnamen des Gruppenkoordinator gesetzt werden. Dieser wird beim Umschalten in den globalen Modus allen Teilnehmern explizit bekannt gegeben.

Das QoS-Management von COVIS musste, wie schon erwähnt, nicht modifiziert werden, um mit OCTOPUS zusammen arbeiten zu können.

### 6.4. Konzepte des BRAVIS

BRAVIS wurde als direkter Nachfolger von COVIS für IP-Netzwerke konzipiert. Aus diesem Grund ähneln sich das Architekturbild und die Implementierung beider Systeme sehr. Im Wesentlichen wurde die Transportschnittstelle von AAL 5/ATM auf UDP/IP modifiziert.

Medienübertragung

Die Medienübertragung wurde basierend auf den Erfahrungen des OCTOPUS mit der RTP-Bibliothek von Lucent Technologies [94] implementiert. Damit ergaben sich auf Empfängerseite wesentliche Änderungen im Programmab-



lauf, da mit RTP kein direkter Verbindungsaufbau stattfindet und somit ein Kindprozess für den Empfang eines Videostroms nicht im direkten Zusammenhang mit einem Verbindungsaufbau erzeugt werden konnte. Dazu musste eine zusätzliche Abfrage der Gruppenzusammensetzung implementiert werden.

Da BRAVIS sowohl für Linux als auch SUNSolaris implementiert werden sollte, ergaben sich einige Änderungen der Video- und Audioverarbeitung in der Implementierung. Während unter SUNSolaris die *X-Image-Library (XIL)* [150] für das Videograbbing und die Videokompression zwingend notwendig ist, wurde unter Linux die weitgehend standardisierte *Video-for-Linux (V4L)*-Schnittstelle für das Abgreifen der Videodaten von der Videokarte genutzt. Für die Kompression stehen unter Linux diverse Bibliotheken zur Verfügung, die unterschiedlich stabil arbeiten und gut dokumentiert sind. Für BRAVIS wurde die JPEG-Lib [57] für die Kompression von M-JPEG genutzt. Des Weiteren wurden verschiedene Bibliotheken für die MPEG-, DivX- und H.26x-Kompression getestet, die jedoch eine zu hohe Leistung der Endsysteme erforderten. Es wurde in diesem Zusammenhang die Bibliothek `livico` für BRAVIS entwickelt, die eine einheitliche Programmierschnittstelle für die Videoverarbeitung sowohl unter SUNSolaris als auch unter Linux bereitstellt [128].

Die Audioverarbeitung in SUNSolaris und Linux unterscheidet sich ebenfalls grundsätzlich. SUNSolaris stellt eine Schnittstelle über IO-Controls zur Verfügung. Unter Linux existieren zwei nutzbare Audio-Bibliotheken, das *Advanced Linux Sound Architecture (ALSA)*-Projekt [4] und das ältere *Open Sound System (OSS)* [1]. Aufgrund der ausführlicheren Dokumentation wurde vorerst OSS als Grundlage für die Linux-Implementierung der Audiokomponenten gewählt.

Das COVIS-Konzept wurde in BRAVIS um eine Whiteboard- und Chatkomponente erweitert. Bei der Implementierung wurde auf Kompatibilität des Quellcodes bezüglich der Betriebssysteme SUNSolaris und Linux geachtet. Da im Gegensatz zur Video- und Audioverarbeitung keine Hardwaregeräte gesteuert werden müssen, konnte eine Lösung unabhängig vom verwendeten UNIX-Betriebssystem implementiert werden. Für die zuverlässige Übertragung der Whiteboard- und Chatdaten sollten zwei Ansätze untersucht werden: die Übertragung mit TCP und mit UDP. TCP hat den Vorteil der zuverlässigen Punkt-zu-Punkt-Übertragung mit einer integrierten Staukontrol-

le. Jedoch treten durch die Bytestrom-orientierte Übertragung Verzögerungen auf. Zusätzlicher Mehraufwand entsteht durch die Verwaltung der TCP-Verbindungen, die zu jedem Teilnehmer aufgebaut werden müssen. Hingegen ist die Übertragung mittels UDP schneller und weniger komplex, da keine Verbindung initialisiert werden muss. Jedoch muss bei Verwendung von UDP die Zuverlässigkeit der Übertragung durch zusätzliche Protokollmechanismen in der Anwendungsschicht gesichert werden. Der erste Prototyp enthält die Implementierung mittels TCP. Das konkrete Leistungsverhalten beider Ansätze wird derzeit in einer Studienarbeit untersucht. Zur Zeit liegen noch keine Messergebnisse vor, so dass keine Aussage gegeben werden kann, welcher Ansatz in BRAVIS letztendlich integriert wird.

Konferenzsteuerung

Die Implementierung des Konferenzmanagements in COVIS ist hardwareunabhängig und so weit wie möglich in Standard-C implementiert. Aus diesem Grund konnten die Module für die Floorkontrolle und das QoS-Management für BRAVIS übernommen werden. Das  $GCP_{ATM}$  wurde vorerst in ein  $GCP_B$  (*Group Communication Protocol for BRAVIS*) überführt, bei dem Multicast-Übertragungen auf mehrfache Punkt-zu-Punkt Verbindungen abgebildet wurden. Dadurch wurde eine Reaktionszeit unabhängig von der Teilnehmerzahl erwartet, was bei  $GCP_{IP}$  nicht der Fall ist. Jedoch besteht wie beim  $GCP_{ATM}$  auch beim  $GCP_B$  das Problem der Quittungsimplosion, das durch künstliche Verzögerungen zwischen dem Absenden zweier Datenpakete vermieden werden soll. Tests mit dem  $GCP_B$  haben jedoch gezeigt, dass die optimale Einstellung dieser Verzögerungen schwierig ist und infolge dessen die Gesamtverzögerung der Datenpakete nicht deterministisch war (siehe auch Anhang C).

Zusätzlich zu COVIS wurde SIP [55] in BRAVIS integriert, um den aktuellen Standort eines Teilnehmers zu ermitteln. Somit können anstelle von hostbezogenen nun personenbezogene Einladungen durchgeführt werden. Nach einer Analyse einiger Implementierungen von SIP wurde die Bibliothek *oSIP* [117] für BRAVIS genutzt, da *oSIP* einerseits frei verfügbar und andererseits gut strukturiert und dokumentiert ist.

Nutzerschnittstelle

Ein weiterer Unterschied der BRAVIS-Implementierung zu der des COVIS besteht in der Implementierung der graphischen Nutzerschnittstelle. Während COVIS noch an die Programmierung von X11 angelehnt ist, wurde für BRAVIS ein Konzept gesucht, das interoperabel mit SUNSolaris, Linux und auch

Windows ist. Die Software-Bibliothek *GTK+* [52] wird dieser Anforderung gerecht. *GTK+* stellt ein Rahmenwerk mit einer breiten Sammlung von Funktionen für die Implementierung von graphischen Oberflächen zur Verfügung. Aus diesem Grund wurde *GTK+* in BRAVIS genutzt. *TCL/TK* stellt ein weiteres Rahmenwerk für die interoperable Gestaltung von Oberflächen dar. Es benötigt jedoch aufgrund seiner Script-basierten Programmierung hohe Ressourcen des Endsystems. Des Weiteren wurde *TCL/TK* seit einigen Jahren nicht mehr weiterentwickelt und gepflegt.

## 6.5. Zusammenfassung und Einschätzung der Implementierung

Die Implementierung mit anschließenden Testläufen von OCTOPUS hat gezeigt, dass globale Konferenzen über eine hierarchische Struktur mit einer relativ guten Qualität möglich sind. Die Idee, lokale Videokonferenzsysteme transparent über globale Netze zu verbinden, konnte erfolgreich durchgesetzt werden. Jedoch konnte der Grundsatz, das lokale Videokonferenzsystem in seiner Funktionsweise nicht zu verändern, nicht vollständig beibehalten werden. So musste der Ablauf der Signalisierung für die Floorkontrolle geändert und ein zusätzliches Modul für das abstimmungs-basierte Bandbreitenmanagement integriert werden.

Während OCTOPUS nur als Prototyp implementiert wurde, um die Umsetzbarkeit des Ansatzes zu beweisen, geht der Stand der Implementierung von BRAVIS über den eines Prototypen hinaus. BRAVIS konnte erfolgreich auf der CeBIT 2003 ausgestellt werden und steht demnächst als Download zur Verfügung. Einige Komponenten wie das QoS-Management, die Audio- und Videoübertragung sowie das Whiteboard werden zur Zeit hinsichtlich Stabilität und Funktionsumfang verbessert. Aufgrund der späten Fertigstellung und Verbreitung einer stabilen Version konnten im Rahmen dieser Arbeit keine Akzeptanz- und Performanzanalysen von BRAVIS durchgeführt werden.

Ein weiteres Ziel der Implementierung von BRAVIS ist die Überführung in eine plattformunabhängige Version. JAVA mit dem *JAVA Media Framework (JMF)* bietet dazu eine gute Ausgangsbasis. Prototypische Implementierungen der Module für die Video- und Audioübertragung haben gezeigt, dass die

## *6. Konzepte der Implementierung*

---

Verarbeitung der Daten in Echtzeit möglich ist. Die komplette Umsetzung des Projektes in JAVA im Rahmen dieser Arbeit wäre jedoch zu aufwendig gewesen.

---

## 7. Zusammenfassung und Ausblick

---

Ziel der vorliegenden Arbeit war es, die Gestaltungsmöglichkeiten von geschlossenen Mehrteilnehmer-Videokonferenzsystemen über globale öffentliche Netzwerke zu diskutieren. Zunächst wurden dazu die am Markt verfügbaren Lösungen in Bezug auf ihre Unterstützung von geschlossenen Teilnehmergruppen im Internet untersucht. Weitere Kriterien für die Analyse waren die Unterstützung globaler, heterogener Netzwerke, der Spontaneität einer Konferenzdurchführung sowie automatischer Mechanismen zur einfachen Bedienung und zur Optimierung der Übertragungsqualität. Die Untersuchungen haben gezeigt, dass nach fast 70-jähriger Forschung und Entwicklung immer noch kein Videokonferenzsystem existiert, das alle genannten Anforderungen zufrieden stellend unterstützt.

In der vorliegenden Arbeit wurde eine Reihe von Ansätzen entwickelt, die eine breite Nutzung von Videokonferenzdiensten ermöglichen soll. Sie wurden prototypisch in den Videokonferenzsystemen OCTOPUS und BRAVIS umgesetzt. Der Schwerpunkt der Untersuchungen lag dabei auf verteilt organisierten geschlossenen Gruppen für die Durchführung von Diskussionen und Beratungen im Internet. Durch den verteilten Ansatz soll die Abhängigkeit der Nutzer von teuren, spezialisierten Servern vermieden werden und ihre Mobilität unterstützt werden. Bei den Endsystemen wird nicht von spezialisierten Geräten sondern Desktop-Systemen ausgegangen, um einen leichten Zugang zum Videokonferenzdienst zu ermöglichen.

Kernproblem des Ansatzes ist die Sicherung der Geschlossenheit der Konferenzgruppe im Internet. Hierfür wurden bisher zwei Ansätze angewendet: die Verschlüsselung der Nutzerdaten und signalisierungsbasierte Ansätze. Die Verschlüsselung der Nutzerdaten ist sehr aufwendig und prinzipiell für eine geschlossene Gruppen nicht zwingend notwendig. Sie erschwert die Nutzung eines Videokonferenzdienstes. Signalisierungsbasierte Ansätze für dezentral organi-

sierte geschlossene Gruppen gab es bisher nur für ATM-Netze, die ein senderbasiertes Multicast unterstützen. Für die Sicherung der Geschlossenheit von Konferenzgruppen über IP-Netze wurde im Rahmen dieser Arbeit das Signalisierungsprotokoll  $GCP_{IP}$  entwickelt. Es basiert auf einem logischen Ring, auf dem sich alle Teilnehmer befinden, unter Nutzung von Unicast-Verbindungen zwischen den Teilnehmern.

$GCP_{IP}$  kann im Gegensatz zu bisherigen Lösungen quasi überall eingesetzt werden, da es nicht auf Multicastdienste angewiesen ist.  $GCP_{IP}$  erfüllt die Forderungen an eine zuverlässige, total geordnete und atomare Datenübertragung. Das Protokoll ist deshalb geeignet, um verteilte Verwaltungsaufgaben, wie beispielsweise eine Floorkontrolle, zu unterstützen. Aber auch im Bereich der Peer-to-Peer-Kommunikation und -anwendungen kann  $GCP_{IP}$  zur Verteilung von Signalisierungsnachrichten verwendet werden.  $GCP_{IP}$  wurde in den Videokonferenzsystemen OCTOPUS und BRAVIS integriert und erprobt. Eine Vielzahl von Messungen und Abschätzungen sind in dieser Arbeit durchgeführt und dargestellt worden. Sie haben bewiesen, dass das  $GCP_{IP}$  bis zu etwa 20 in Europa verteilte Teilnehmer in einer geschlossenen Gruppe, wie in einer OCTOPUS- beziehungsweise BRAVIS-Konferenz, unterstützen kann. Somit ist auch das entwickelte Gesamtkonzept eines verteilten Gruppen- und QoS-Managements für Videokonferenzen anwendbar.

Gegenstand der Untersuchung war auch die Signalisierung in einer hierarchischen Topologie, wie sie im OCTOPUS zu finden ist. Messungen haben gezeigt, dass auch dieser Ansatz praktikabel ist, obwohl geringfügige Verzögerungen in den Gruppenkoordinatoren auftreten.

Einen weiteren Schwerpunkt der Arbeit bildeten Untersuchungen der Verteilung der Audio- und Videodaten in geschlossenen Konferenzgruppen. Bisherige Ansätze für verteilt organisierte Konferenzgruppen basieren auf einem senderbasierten Multicastdienst, der bislang nur durch ATM-Netze angeboten wird. Die Einführung senderbasierter Multicastdienste in IP-Netze erweist sich nach wie vor als schwierig. Eine rasche Nutzung dieser Technologie in Videokonferenzdiensten für geschlossene Gruppen ist deshalb nicht zu erwarten. Um den Nutzern heute und in der Zukunft einen einfachen und von speziellen Netzwerkdiensten unabhängigen Zugang zu einem Videokonferenzdienst zu ermöglichen, wurden in dieser Arbeit Ansätze zur Verteilung der Mediendaten durch die

---

Endsysteme untersucht, bei denen Punkt-zu-Punkt-Verbindungen über das Internet genutzt werden. Im Gegensatz zur Signalisierung wurde hierbei ein hierarchisches Konzept verfolgt. OCTOPUS als System für den Videokonferenzverbund führt dazu ein *anpassungsfähiges* Gateway, den Gruppenkoordinator, ein, das zur Kommunikation zwischen lokalen Videokonferenzsystemen über ein globales Netzwerk dient. Eine hierarchische Verteilung der Mediendaten wurde auch in BRAVIS umgesetzt. Hierbei wurde jedoch auf ein Gateway verzichtet. Statt dessen wurde das auf der Peer-to-Peer-Technologie beruhende Konzept eines *ausgewählten Empfängers* realisiert, bei dem bestimmte Teilnehmer der Konferenzgruppe die Weiterleitung der Mediendaten an andere Teilnehmer übernehmen. Der entscheidende Vorteil der untersuchten hierarchischen Verteilungsstrategie liegt in der optimalen Ausnutzung vorhandener Ressourcen, wie zum Beispiel der Bandbreite zwischen den Teilnehmern und der Rechenkapazität der Endsysteme. Des Weiteren werden Single-Point-of-Failures und Kommunikationsengpässe mit dem gewählten Ansatz vermieden.

Ein Kernthema der Arbeit war außerdem die Untersuchung der Sicherung der Dienstgüte und der optimalen Ausnutzung vorhandenen Bandbreitenressourcen im Internet. Hierbei wurde ein neuartiger Ansatz entwickelt, der die verfügbare Bandbreite einer Verbindung durch ein Abstimmungsverfahren dynamisch ausgewählten Teilnehmern zuweist. Wichtige Videodaten einer Videokonferenz können daraufhin mit einer höheren Qualität übertragen werden. Die Entwicklung dieses Schemas beruht auf Beobachtungen, dass nicht alle Teilnehmer einen gleichen Anteil an einer Videokonferenz haben. So sind nur der aktuelle Sprecher und vielleicht zwei bis drei weitere Konferenzpartner von Interesse, die aktiv die Diskussion gestalten. Außerdem löst dieser neu entwickelte Ansatz das bisher kaum beachtete Sichtbarkeitsproblem, das auftritt, wenn alle Teilnehmer einer Videokonferenz und beispielsweise eine Joint-Viewing-Anwendung auf dem Bildschirm dargestellt werden sollen. Oft überlappen sich dann die Videobilder der Teilnehmer, wodurch der Nutzer gezwungen ist, die Aufteilung seines Bildschirms dynamisch anzupassen.

Das Abstimmungsverfahren hat sich auch als zweckmäßig für die Unterstützung *heterogener Systemumgebungen* erwiesen, wie sie im globalen öffentlichen Internet vorkommen. Dadurch konnte ein dediziertes QoS-Management für geschlossene Gruppen entwickelt werden, dass die Dienstqualität individuell an

die Bedürfnisse der Teilnehmer beziehungsweise der Untergruppen anpasst. Zwischen den einzelnen Ebenen der untersuchten hierarchischen Verteilungsschemen für die Mediendaten berücksichtigt das entworfene Auswahlverfahren heterogene Netzwerkbedingungen. Dazu wird periodisch die Gesamtbandbreite auf den Verbindungen zwischen den Ebenen gemessen und daraufhin die Sendeleistung der Teilnehmer dynamisch angepasst. Um heterogene Endsysteme zu unterstützen, wurde in OCTOPUS die Skalierung der komprimierten Videoströme durch die Empfänger beibehalten. In BRAVIS hingegen wird die Kapazität der Endsysteme bei der dynamischen Anpassung der Sendeleistung berücksichtigt. Um jedoch heterogene Endsysteme innerhalb einer Hierarchieebene zu unterstützen, bleibt die empfängerbezogene Skalierung in den Endsystemen erhalten.

Messungen, Abschätzungen und Simulationen haben die technische Anwendbarkeit dieses abstimmungs-basierten Bandbreitenzuweisungsalgorithmus für Videokonferenzen über das Internet bewiesen. Eine Studie über die Akzeptanz durch die Nutzer muss jedoch noch durchgeführt werden.

Die in der Arbeit entwickelten Ansätze für geschlossene Konferenzgruppen sichern eine „Vertraulichkeit“, wie sie vom Telefondienst her bekannt ist. Dies reicht jedoch bei manchen Firmen- und Projekttreffen nicht aus. Deshalb sind weiterführende Arbeiten zur Integration einer Verschlüsselungskomponente erforderlich. Dazu bietet das GCP<sub>IP</sub> eine gute Grundlage für ein zu entwickelndes verteiltes Schlüsselmanagement.

GCP<sub>IP</sub> ist bisher als Protokoll in der Anwendungsebene implementiert. Zukünftig könnte dieses Protokoll auch in der Transportschicht neben TCP und UDP agieren. Dafür muss das Protokoll jedoch in den Betriebssystemkern hinein kompiliert werden und um Funktionalitäten, wie das Multiplexen auf verschiedene Transportschichtadressen, erweitert werden. Eine Implementierung einer Programmierschnittstelle (*API*) würde zudem eine einfachere Verbreitung und Nutzung des Protokolls bewirken.

Auf der anderen Seite ist das *Session Initiation Protocol (SIP)* schon ein sehr verbreitetes Signalisierungsprotokoll. Jedoch unterstützt es keine Sicherung der Übertragung, der Übertragungsreihenfolge und Atomarität der versendeten Daten. SIP ist jedoch sehr flexibel und erweiterbar, so dass es um Protokollfunktionen des GCP ergänzt werden kann. Gegenwärtig werden da-



---

zu Untersuchungen zur Umsetzbarkeit durchgeführt [124], die zukünftig durch eine Leistungsbewertung untermauert werden müssen.

Immer mehr Firmen und Universitäten setzen Applikationsserver ein, um den Verwaltungs- und Installationsaufwand zu reduzieren. *Thin-Clients* mit reduzierter Ausstattung werden an hochperformante Server angeschlossen, auf denen die Büroanwendungen abgearbeitet werden. Ein Problem stellt dabei die Realisierung von Videokonferenzen dar, die auf eigene Ressourcen wie Video- und Soundkarten der Endsysteme, also den Thin-Clients, zurückgreifen. Des Weiteren ist unklar, ob sich damit ein verteiltes Signalisierungskonzept, wie mit GCP<sub>IP</sub> durchsetzen lässt. Diese und weitere Probleme sind in diesem Zusammenhang zu analysieren und zu lösen.

OCTOPUS und BRAVIS benötigen relativ hohe Ressourcen. Dies ist vor allem in der Realisierung als reine Software-Lösung begründet. In der Zukunft wird zwar die Leistungsfähigkeit der Endsysteme ansteigen, aber die Performanz von Hardwarelösungen wird wohl kaum erreicht werden. Deshalb ist es günstig, zumindest die Video- und Audiokompression in Hardwarekomponenten zu realisieren. Hierbei muss das Problem untersucht werden, ob zukünftige Kompressionshardware die Ansprüchen an eine flexible, heterogen anpassungsfähige Kompression erfüllen kann. Ebenfalls interessant ist die Frage nach einer hardwareseitigen Dekompression, die unabhängig von der Anzahl der empfangenen Videoströme frei skalierbar ist.

Ein sehr wichtiger Punkt in der zukünftigen Entwicklung ist die Betrachtung von mobilen Teilnehmern, die über kabellose Netzwerke an Videokonferenzen teilnehmen. Hierbei müssen besondere Bedingungen berücksichtigt werden. Zum einen ist die verfügbare Bandbreite geringer. Zum anderen können Übertragungsfehler- und -verluste nicht nur auf Stausituationen im Netz zurückgeführt werden. Hier muss also auch die Qualität der Übertragungsstrecke bei der Sicherung der Dienstgüte eines Videokonferenzsystems betrachtet werden.

Die in dieser Arbeit untersuchten und verwirklichten Ansätze bilden eine sehr gute Grundlage für die Entwicklung zukünftiger Videokonferenzsysteme sowie Peer-to-Peer-Anwendungen. Besonders im privaten Bereich aber auch in Wirtschaft, Wissenschaft und Forschung bieten die erarbeiteten Lösungen einen entscheidenden finanziellen Vorteil als auch qualitativen Vorteil.

## *7. Zusammenfassung und Ausblick*

---

# A. Automatengraph des GCP<sub>IP</sub>

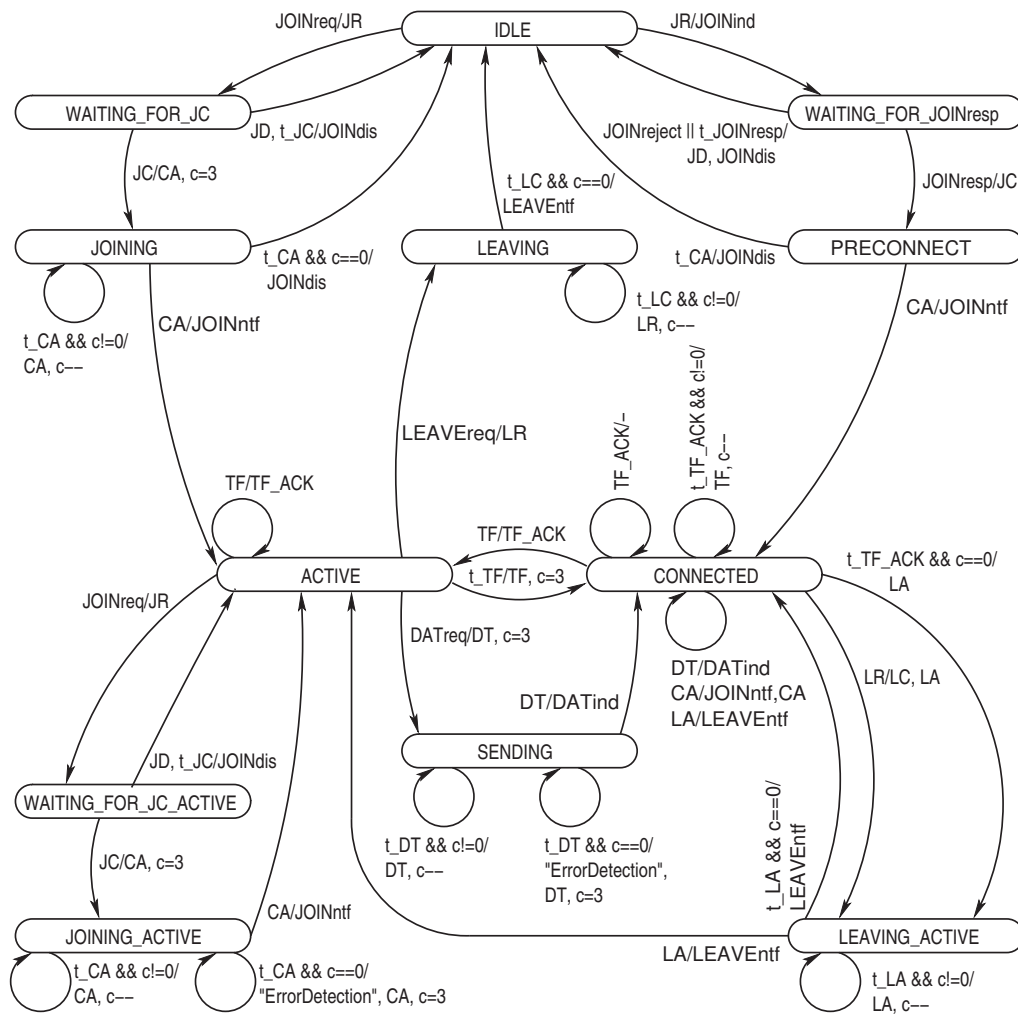


Abbildung A.1.: Automatengraph des GCP<sub>IP</sub>

Der in Abbildung A.1 dargestellte Automatengraph enthält nicht die Signalisierung für die Messung der kürzesten Verzögerungszeit während des Verbindungsaufbaus sowie für die Fehlerbehandlung „*ErrorDetection*“, da dadurch

### A. Automatengraph des $GCP_{IP}$

---

die Übersicht des Graphen gewährleistet bleibt. Ferner wurde ein Zähler  $c$  in den Graphen integriert, der die Anzahl der Wiederholungen einer missglückten Übertragung registriert.

---

## B. Fehlerfälle des GCP<sub>IP</sub>

---

### B.1. Konferenzinitialisierung

Die Konferenz kann nicht initialisiert werden, wenn eines der Datenpakete JR, JC oder CA verloren gehen. Des Weiteren schlägt die Initialisierung fehl, wenn die Einladung nicht bestätigt wird.

#### Verlust des JR

Geht das JR verloren, so wird dies vom Initiator der Konferenz durch einen Timeout bemerkt. Daraufhin wird ein JOINdis angezeigt (siehe Abbildung B.1).

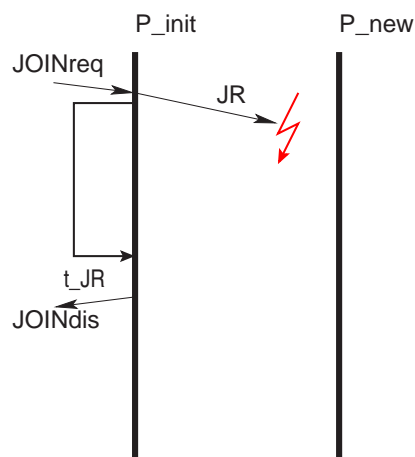


Abbildung B.1.: Verlust des JR

#### Verlust des JC

Geht das JC verloren, so wird dies vom Initiator der Konferenz durch einen Timeout bemerkt (siehe Abbildung B.1). Daraufhin wird ein JOINdis angezeigt.

Der Eingeladene bemerkt diese Situation ebenfalls durch einen Timeout. Auch hier wird ein JOINdis angezeigt.

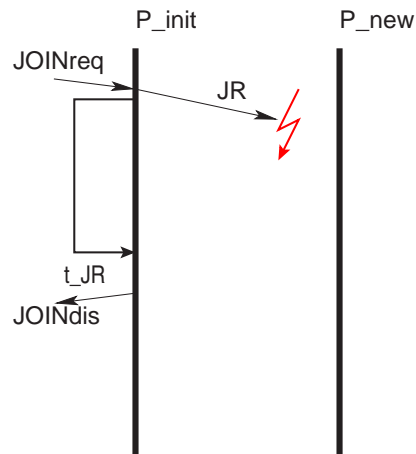


Abbildung B.2.: Verlust des JC

### Verlust des CA

Geht das CA verloren, so wird dies vom Initiator der Konferenz durch einen Timeout bemerkt (siehe Abbildung B.3). Da der Eingeladene eine Bestätigung geschickt hat, geht der Initiator davon aus, dass ein Übertragungsfehler aufgetreten ist. Daraufhin wird das CA bis zu dreimal wiederholt. Ist bis dahin keine Bestätigung angekommen, ist die Konferenzinitialisierung gescheitert und es wird ein JOINdis angezeigt. 4.3.2).

Der Eingeladene erwartet nach dem Senden der Bestätigung JC auf die CA Nachricht. Trifft diese nicht in einer bestimmten Zeit ein, gilt die Initialisierung als gescheitert und es wird ein JOINdis übergeben. Geht das vom eingeladenen Teilnehmer gesendete CA verloren, so wird der Mechanismus zur Behebung eines Tokenverlustes gestartet, da sich der Eingeladene nun im verbundenen Zustand befindet (siehe Abschnitt 4.3.2).

## B.2. Tokenübergabe

Der normale Ablauf der Tokenübergabe kann durch den Verlust des TF oder des TF\_ACK Paketes gestört werden. Beide Fehlerfälle werden in Abbildung B.4 und B.5 gezeigt.

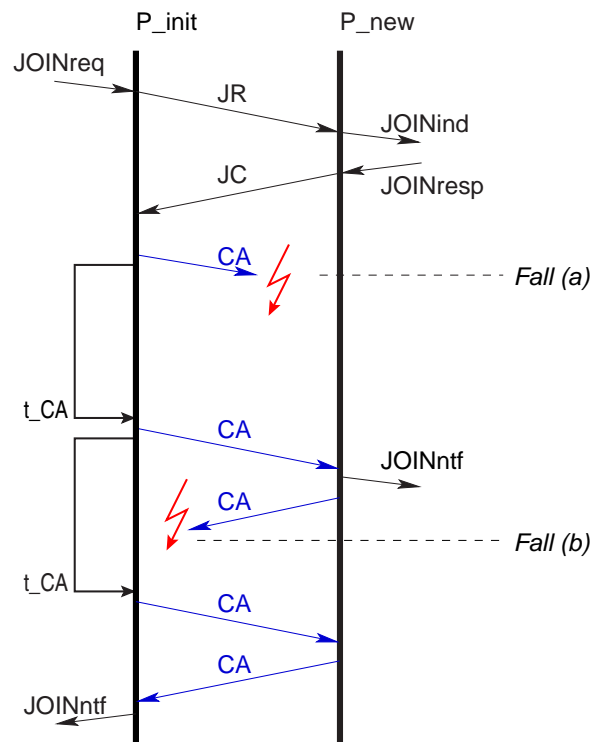


Abbildung B.3.: Verlust des CA (a) vom Initiator und (b) vom Eingeladenen

### Verlust des TF

Geht das TF verloren, so wird es bis zu dreimal wiederholt. Danach wird davon ausgegangen, dass der neue Tokenhalter nicht erreichbar ist. Daraufhin wird er durch das *Forced Leave* aus der Konferenz ausgeschlossen.

### Verlust des TF\_ACK

Bei einem Verlust des TF\_ACK, wird der vorherige Tokenhalter das TF wiederholen, das mit einem erneuten TF\_ACK bestätigt wird. Hat der Tokenhalter bereits den Token weitergegeben, so prüft er, ob das empfangene TF ein neuer Token ist oder ob dies ein Token ist, dessen Bestätigung verloren gegangen war.

## B.3. Datenübertragung

Ein Fehler während der Datenübertragung kann nur durch den Verlust des DT Paketes auftreten, der durch einen Bruch des logischen Rings hervorgerufen wird (siehe Abbildung B.6). Der Ring kann dabei einerseits durch den Ausfall

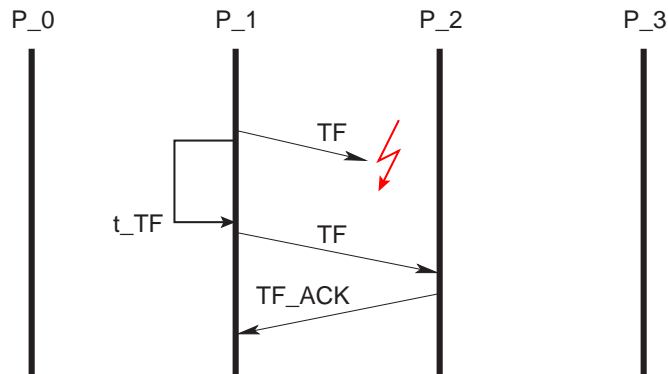


Abbildung B.4.: Verlust des TF

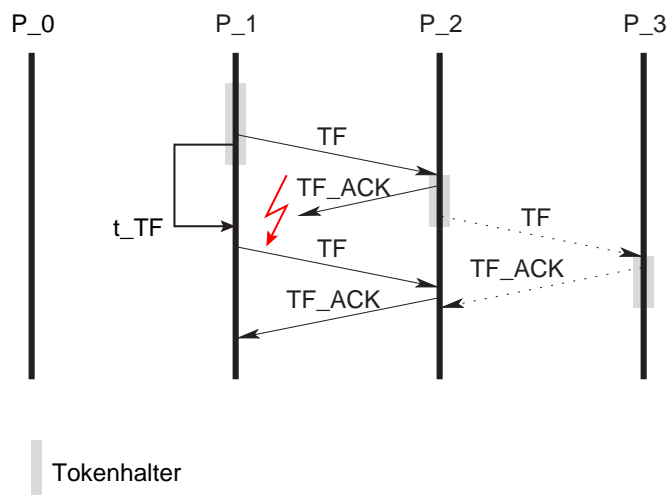


Abbildung B.5.: Verlust des TF\_ACK

eines Teilnehmers als auch durch den Ausfall einer Verbindung zwischen den Teilnehmern unterbrochen werden. In diesen Fällen trifft das vom Tokenhalter gesendete Datenpaket nicht wieder bei ihm ein. Daraufhin wird das Paket bis zu dreimal wiederholt. Ist bis dahin die Übertragung nicht erfolgreich, wird der *Error Detection* Mechanismus mit einem anschließenden *Forced Leave* gestartet, um den Fehler zu finden und zu beheben.



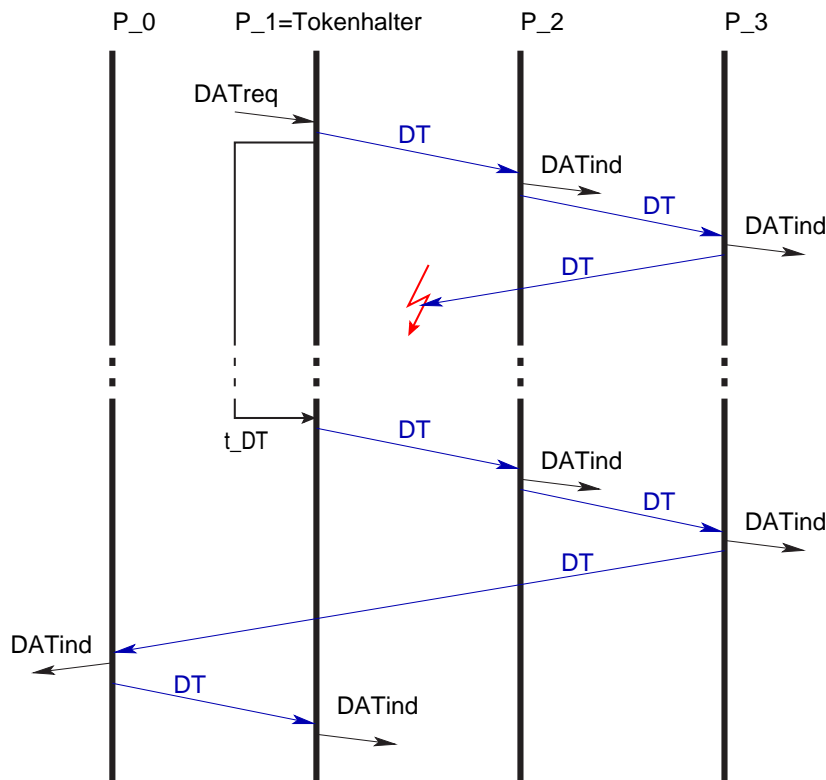


Abbildung B.6.: Verlust des DT

## B.4. Verlassen der Konferenz

Das Verlassen der Konferenz kann durch den Verlust der Pakete LR, LC oder LA gestört werden.

### Verlust des LR

Mit dem LR wird auch der Token weitergegeben. Geht das LR verloren, so wird es bis zu dreimal wiederholt (siehe Abbildung B.7). Trifft bis dahin keine Bestätigung ein, so wird die Konferenz ohne Weiteres verlassen. Dieser Fall gleicht dem eines Tokenverlust, da der Tokenhalter quasi ohne Ankündigung aus der Konferenz ausgetreten ist. In diesem Fall wird durch einen der verbleibenden Teilnehmer der *Error Detection* Mechanismus gestartet.

### Verlust des LC

Nach einem Verlust des LC, wird das LR vom austretenden Teilnehmer bis zu dreimal wiederholt (siehe Abbildung B.8). Geht das LC immer verloren, so

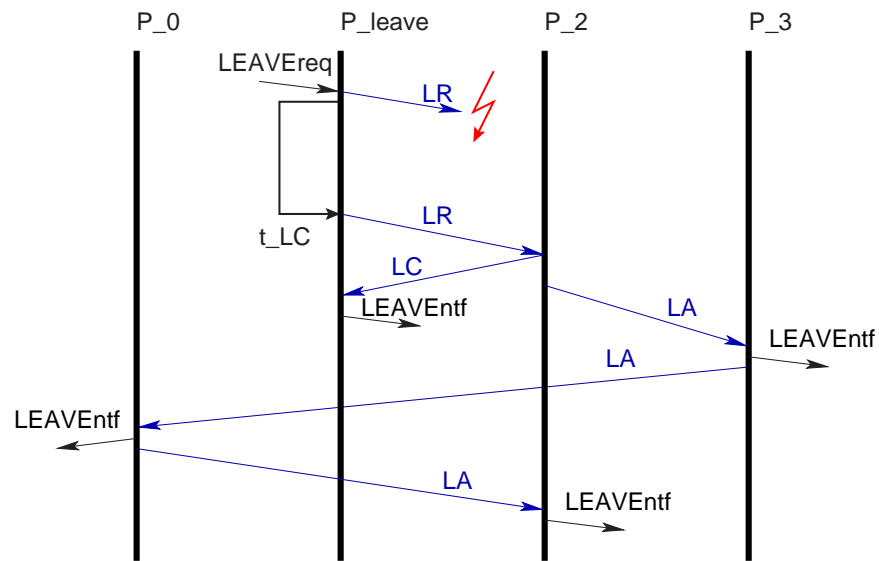


Abbildung B.7.: Verlust des LR

wird, wie oben erläutert, die Konferenz ohne Weiteres verlassen. In diesem Fall existiert jedoch ein neuer Tokenhalter, der den Austritt des Teilnehmers bekannt gibt.

### Verlust des LA

Der Verlust eines **LA** wird wie der Verlust eines DT Paketes behandelt (siehe Abbildung B.9). Der neue Tokenhalter, wiederholt das **LA** bis zu dreimal. Ist bis dahin die Übertragung ohne Erfolg, werden der *Error Detection* und der *Forced Leave* Mechanismus gestartet.

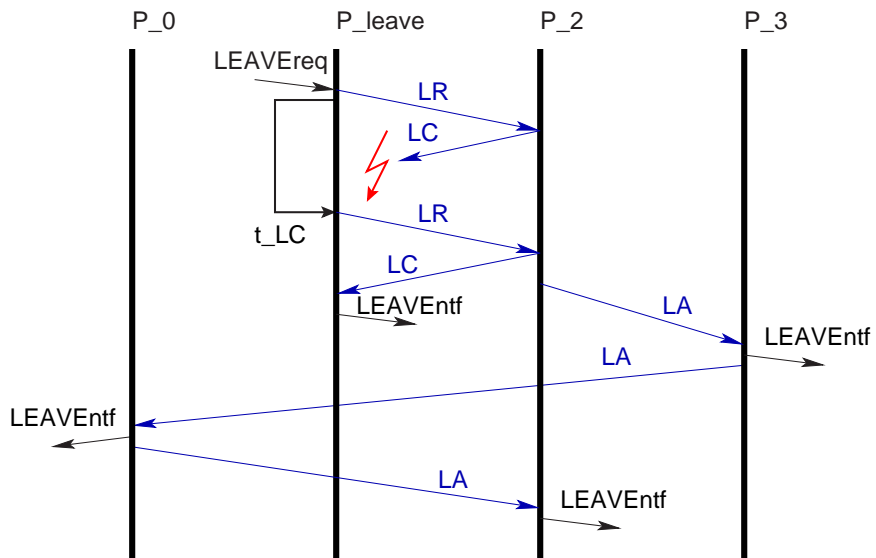


Abbildung B.8.: Verlust des LC

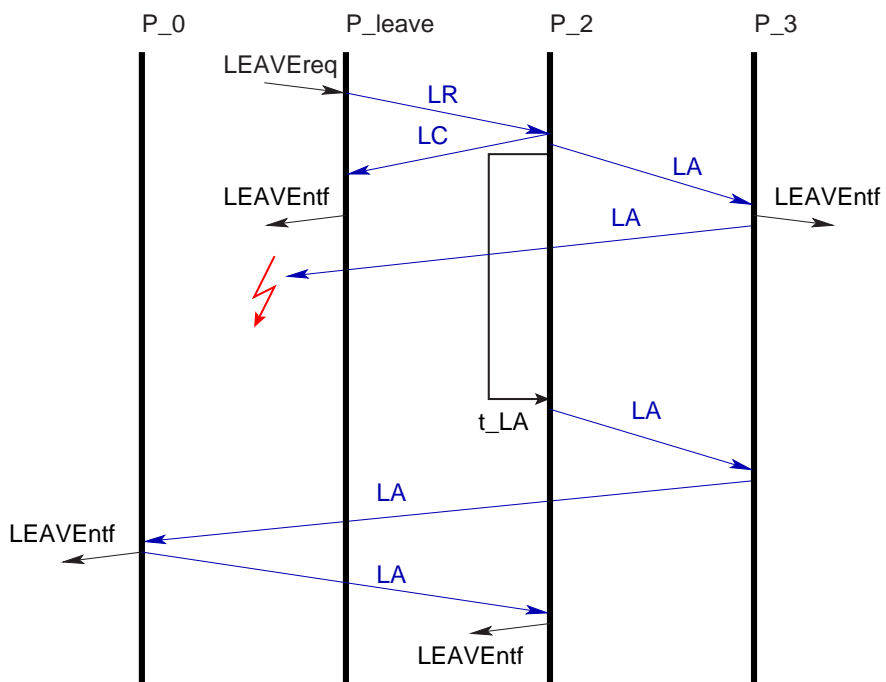


Abbildung B.9.: Verlust des LA



---

## C. GCP<sub>B</sub>

---

Das *Group Communication Protocol for BRAVIS* (GCP<sub>B</sub>) ist eine dritte Variante des GCP, die die ATM-Multicast-Verbindungen des GCP<sub>ATM</sub> auf Unicast-Verbindungen umsetzt. Der BRAVIS-Ansatz enthielt ursprünglich das im Rahmen meiner Arbeit entwickelte GCP<sub>IP</sub>, das während der Konzeption des BRAVIS-Systems für OCTOPUS implementiert wurde. Zu Vergleichszwecken und aufgrund der noch nicht vollständigen Implementierung von GCP<sub>IP</sub> zum Zeitpunkt des Beginns der Entwicklung von BRAVIS wurde GCP<sub>B</sub> durch die Projektmitarbeiter an unserem Lehrstuhl implementiert. Das Problem der Quittungsimplosion wurde durch diverse Verzögerungen bei dem Versand der Quittungen umgangen. Dies zog jedoch hohe und nicht vorhersagbare Gesamtverzögerungen mit sich (siehe Abbildung C.1). Die drei durchgeführten Messreihen mit jeweils 500 Messwerten lieferten nicht deterministische Ergebnisse, die auch nicht auf eine Approximation für mehr als 6 Teilnehmer schließen ließen.

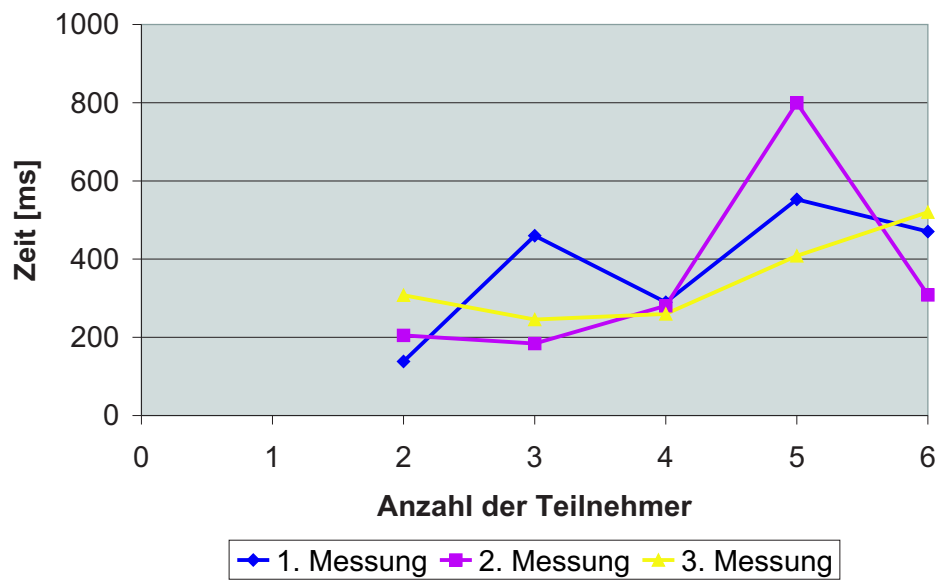


Abbildung C.1.: Messung der tatsächlichen, durchschnittlichen Verzögerung durch  $GCP_B$

---

## Literaturverzeichnis

---

- [1] 4FRONT TECHNOLOGIES: *Digital Audio for UNIX*. <http://www.opensound.com>.
- [2] AGARWAL, D.: *Totem: A Reliable Ordered Delivery Protocol for Inter-connected Local-Area Networks*. PhD thesis, University of California, Santa Barbara, August 1994.
- [3] AGARWAL, D., P. MELLIAR-SMITH, and L. MOSER: *Totem: A Protocol for Message Ordering in a Wide-Area Network*. In *Proceedings of the First ISMM International Conference on Computer Communications and Networks, San Diego, CA*, pages 1–5, June 1992.
- [4] ALSA: *Advanced Linux Sound Architecture*. <http://www.alsa-project.org>.
- [5] AMIR, E. and S. M. R. KATZ: *Receiver-driven Bandwidth Adaptation for Light-weight Sessions*. In *ACM Multimedia '97*, 1997.
- [6] AMIR, E., S. MCCANNE, and M. VETTERLI: *A layered DCT coder for Internet video*. In *Proc. IEEE Int. Conf. Image*, Lausanne, Switzerland, 1996.
- [7] ARMSTRONG, S., A. FREIER, and K. MARZULLO: *Multicast Transport Protocol*. RFC 1301, February 1992.
- [8] BANERJEE, S. and B. BHATTACHARJEE: *A Comparative Study of Application Layer Multicast Protocols*.
- [9] BANERJEE, S., B. BHATTACHARJEE, and C. KOMMAREDDY: *Scalable application layer multicast*. Technical report, UMIACS TR-2002., 2002.

- [10] BEIER, I.: *Gestaltungsprinzipien von Videokonferenzsystemen für CSCW-Anwendungen*. Doktorarbeit, BTU Cottbus, 2000.
- [11] BEIER, I. and H. KÖNIG: *A Protocol Supporting Distributed Group and QoS Management*. In *Proceedings, IEEE Conference for Multimedia Systems – Networking (PROMS-MmNet'97)*, pages 213 — 222, 1997.
- [12] BEIER, I. and H. KÖNIG: *GCSVA - A Multiparty Video Conferencing System with Distributed Group and QoS Management*. In *7th International IEEE Conference on Computer Communications and Networks (IC3N'98)*, pages 594 — 598, 1998.
- [13] BEIER, I. and H. KÖNIG: *Distributed Group and QoS Management for a Multiparty Video Conferencing System*. In YONGCHAREON, T., F. A. AAGESEN, and V. WUWONGSE (editors): *Intelligence in Networks. IFIP TC6 WG6.7 5th International Conference, SMARTNET'99*, pages 335 — 348. Kluwer Academic Publisher, 1999.
- [14] BIERSACK, E.: *Performance Evaluation of Forward Error Correction in ATM Networks*. In *ACM SIGCOMM'92*, pages 248 — 258, August 1992.
- [15] BLAKE, S., D. BLACK, M. CARLSON, E. DAVIES, Z. WANG, and W. WEISS: *An Architecture for Differentiated Services*. RFC 2475, December 1998.
- [16] BOIVIE, R., N. FELDMANN, Y. IMAI, W. LIVENS, D. OOMS, and O. PARIDAENS: *Explicit Multicast (Xcast) Basic Specification*. Internet Draft Proposal, June 2002.
- [17] BOLOT, G. C. and A. VEGA-GRACIA: *Control Mechanisms for Packet Audio in the Internet*. In *IEEE Infocom '96, San Francisco, CA*, pages 232 — 239, April 1996.
- [18] BOLOT, J.C. and T. TURLETTI: *Experience with Control Mechanisms for Packet Video in the Internet*. *Computer Communication Review*, 1998.



- 
- [19] BOLOT, J.C., T. TURLETTI, and I. WAKEMAN: *Scalable Feedback Control for Multicast Video Distribution in the Internet*. In *ACM SIGCOMM'94*, pages 58—67, 1994.
- [20] BORMAN, C., D. KUTSCHER, J. OTT, and D. TROSSEN: *SCCP: Simple Conference Control Protocol*. Internet Draft Proposal, August 2001.
- [21] BRADEN, R.: *Resource reSerVation Protocol (RSVP) - Version 1 Functional Specification*. RFC 2205.
- [22] BRADEN, R.: *Requirements for Internet Hosts – Communication Layers*. RFC 1122, October 1989.
- [23] BRAND, O. und M. ZITTERBART: *Steuerung von Konferenz- und Kollaborationsanwendungen*. Praxis der Informationsverarbeitung und Kommunikation (PIK) 20(1997)4, Seiten 209 — 216, 1997.
- [24] CASTRO, M., P. DRUSCHEL, A. KERMARREC, and A. ROWSTRON: *SCRIBE: A large-scale and decentralized application-level multicast infrastructure*. IEEE Journal on Selected Areas in communications (JSAC), 2002.
- [25] CHAN, D. and S. CHANSON: *Scalability Support for Multimedia Communications*. ACM/Springer Multimedia Systems Journal, 6:75 — 87, 1998.
- [26] CHANSON, S., A. HUI, H. KÖNIG und M. ZÜHLKE: *Das OCTOPUS – Videokonferenzsystem*. Praxis der Informationsverarbeitung und Kommunikation (PIK) 23(2000)4, Seiten 189 — 198, 2000.
- [27] CHANSON, S., A. HUI, E. SIU, I. BEIER, H. KOENIG, and M. ZUEHLKE: *OCTOPUS – A Scalable Global Multiparty Video Conferencing System*. In *7th International IEEE Conference on Computer Communications and Networks (IC3N'99)*, pages 97 — 102, 1999.
- [28] CHU, Y.-H., S. G. RAO, S. SESHAN, and H. ZHANG: *Enabling conferencing applications on the internet using an overlay multicast architecture*. In *ACM SIGCOMM 2001*, San Diego, CA, August 2001. ACM.

- [29] CHU, Y.-H., S. G. RAO, and H. ZHANG: *A case for end system multicast*. In *ACM SIGMETRICS 2000*, pages 1—12, Santa Clara, CA, June 2000. ACM.
- [30] COMMWEB.COM: *Datacom / IT Statistics*. <http://www.commweb.com/article/COM20021016S0002>.
- [31] CONFERZONE - YOUR ECONFERENCE RESOURCE: *Statistics on a Growing Industry . . . .* [http://www.conferzone.com/about/download/cz\\_stat.pdf](http://www.conferzone.com/about/download/cz_stat.pdf).
- [32] DANTE – SEQUIN: *SERVICE Quality across Independently managed Networks*. <http://www.dante.net/sequin/>.
- [33] DAUGHERTY, G.: *PacketAssist<sup>TM</sup> — An Architectural Approach to Delivering Quality of Service (QoS) for IP Video*. Whitepaper: <http://www.vcon.de>.
- [34] DAVIKO - GESELLSCHAFT FÜR DIGITALE AUDIOVISUELLE KOMMUNIKATION MBH: *daViKo*. <http://www.daviko.com>.
- [35] DEERING, S. E.: *Multicast routing in internetworks and extended LANs*. In *Symposium proceedings on Communications architectures and protocols*, pages 55–64. ACM Press, 1988.
- [36] DELGROSSI, L. and E. BERGER: *Internet Stream Protocol, Version 2 (ST2), Protocol Specification - Version ST2+*. RFC 1819, August 1995.
- [37] DE MIGUEL, T. P., S. PAVON, J. SALVACHUA, J. QUEMADA, A. AZCORRA, M. PETIT, D. LARRABEITI, T. ROBLES, and G. HUECAS: *ISABEL: A CSCW Application for the Distribution of Events*. In *COST 237 Workshop on Multimedia Networks and Systems, Barcelona*, November 1996.
- [38] DE MIGUEL, T. P., S. PAVON, J. SALVACHUA, J. Q. VIVES, P. L. C. ALONSO, J. FERNANDEZ-AMIGO, C. ACUNA, L. RODRIGUEZ YAMAMOTO, V. LAGARTO, and J. VASTOS: *ISABEL Experimental Distributed Cooperative Work Application over Broadband Networks*. In

- STEINMETZ, R. (editor): *Multimedia: Advanced Teleservices and High-Speed Communication Architectures*, pages 353—362. Springer-Verlag, vol. LNCS 868, September 1994.
- [39] DIGITAL SIGNAL PROCESSING GROUP (DSPG) OF THE UNIVERSITY OF VICTORIA, VICTORIA, BC, CANADA: *The JasPer Project Home Page*. <http://www.ece.uvic.ca/~mdadams/jasper>.
- [40] DOMMEL, H.-P. and J. J. GARCIA-LUNA-ACEVES: *Floor Control for Multimedia Conferencing and Collaboration*. *Multimedia Systems*, 5(1):23—38, 1997.
- [41] DWYER, D., S. HA, J.-R. LI, and V. BHARGHAVAN: *An Adaptive Transport Protocol for Multimedia Communication*. In *International Conference on Multimedia Computing and Systems*, pages 23—32, 1998.
- [42] FIRST VIRTUAL COMMUNICATIONS: *Web Conferencing and video conferencing software for Business Communications on the Web*. <http://www.fvc.com>.
- [43] FLOYD, S., V. JACOBSON, C. LIU, S. MCCANNE, and L. ZHANG: *A Reliable Multicast Framework for Light-weight Sessions and Application Level Framing*. In *IEEE/ACM Transactions on Networking*, volume 6, pages 784 — 803, December 1997.
- [44] FORGIE, J.: *ST - A Proposed Internet Stream Protocol*. IEN 119, September 1979.
- [45] FRANCIS, P.: *Yoid: Extending the Internet Multicast Architecture*. <http://www.icir.org/yoid/>.
- [46] GEYER, W. and W. EFFELSBURG: *The Digital Lecture Board - A Teaching and Learning Tool for Remote Instruction in Higher Education*. In *EDMEDIA '98, World Conference on Educational Multimedia and Hypermedia, Freiburg, Germany*, June 1998.
- [47] GMD FOKUS: *USMInT - Universal Scalable Multimedia in the Internet*. <http://www.fokus.gmd.de/research/cc/gclone/projects/usmint/>.

- [48] GRAFF, C.: *IPv4 Option for Sender Directed Multi-Destination Delivery*. RFC 1770, March 1995.
- [49] GRINNEMO, K.-J. and A. BRUNSTROM: *Enhancing TCP for Applications with Soft Real-Time Constraints*. In *The Convergence of Information Technologies and Communications (ITCom), Denver, USA*, volume 4518, pages 18 — 31, August 2001.
- [50] GRINNEMO, K.-J. and A. BRUNSTROM: *Evaluation of the QoS offered by PRTP-ECN - A TCP-Compliant Partially Reliable Transport Protocol*. In *The 9th International Workshop on Quality of Service (IWQoS 2001), Karlsruhe, Germany*, pages 217 — 231, June 2002.
- [51] GRUMANN, M.: *Entwurf und Implementierung eines zuverlässigen Multicast-Protokolls zur Unterstützung sicherer Gruppenkommunikation in einer TeleTeaching-Umgebung*. Diplomarbeit, Universität Mannheim, Fakultät für Mathematik und Informatik, 1997.
- [52] GTK+: *The GIMP Toolkit*. <http://www.gtk.org>.
- [53] HANDLEY, M.: *SAP: Session Announcement Protocol*. RFC 2974, October 2000.
- [54] HANDLEY, M. and V. JACOBSON: *SDP: Session Description Protocol*. RFC 2327, April 1998.
- [55] HANDLEY, M., H. SCHULZRINNE, E. SCHOOLER, and J. ROSENBERG: *SIP: Session Initiation Protocol*. RFC 2543, March 1999.
- [56] HEINANEN, J., F. BAKER, W. WEISS, and J. WROCLAWSKI: *Assured Forwarding PHB group*. RFC 2597, June 1999.
- [57] INDEPENDENT JPEG GROUP: *Independent JPEG Group*. <http://www.ijg.org/>.
- [58] INTEL CORPORATION: *Intel® Proshare® Conferencing Products*. <http://www.intel.com/support/proshare/>.
- [59] INTERNET2 QoS WORKING GROUP: *Internet2 QoS Working Group*. <http://qos.internet2.edu/wg/>.

- [60] ISIS MULTIMEDIA NET GMBH: *ISIS Multimedia Net GmbH & Co. KG (Düsseldorf) - Homepage*. <http://www.isis.de>.
- [61] ISO/IEC: *Information Technology - Quality of Service - Methods and Mechanisms - Working Draft*. ISO/IEC JTC 1/SC 21 N9681, July 1995.
- [62] ISO/IEC: *QoS - Basic Framework - CD Text*. ISO/IEC JTC 1/SC 21 N9309, January 1995.
- [63] ITU-T: *Recommendation G.711 (11/88) Pulse code modulation (PCM) of voice frequencies*. <http://www.itu.int>, November 1988.
- [64] ITU-T: *Recommendation G.721 (11/88) 32 kbit/s adaptive differential pulse code modulation (ADPCM)*. <http://www.itu.int>, November 1988.
- [65] ITU-T: *Recommendation G.722 (11/88) 7 kHz audio-coding within 64 kbit/s*. <http://www.itu.int>, November 1988.
- [66] ITU-T: *Recommendation G.728 (09/92) Coding of speech at 16 kbit/s using low-delay code excited linear prediction*. <http://www.itu.int>, September 1992.
- [67] ITU-T: *I.350 (03/93) General aspects of quality of service and network performance in digital networks, including ISDNs*. <http://www.itu.int>, March 1993.
- [68] ITU-T: *Recommendation H.261 - Video codec for audiovisual services at p x 64 kbit/s*. <http://www.itu.int>, March 1993.
- [69] ITU-T: *Recommendation H.322 - Visual telephone systems and terminal equipment for local area networks which provide a guaranteed quality of service*. <http://www.itu.int>, March 1996.
- [70] ITU-T: *Recommendation G.723.1 (03/96) Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s*. <http://www.itu.int>, March 1996.

- [71] ITU-T: *Recommendation G.729 (03/96) Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear-prediction (CS-ACELP)*. <http://www.itu.int>, March 1996.
- [72] ITU-T: *E.800 (08/94) Terms and definitions related to quality of service and network performance including dependability*. <http://www.itu.int>, August 1998.
- [73] ITU-T: *Recommendation H.263 – Video coding for low bit rate communication*. <http://www.itu.int>, February 1998.
- [74] ITU-T: *Recommendation H.321 – Broadband visual telephone systems and terminal equipment*. <http://www.itu.int>, February 1998.
- [75] ITU-T: *Recommendation T.120 – Multipoint Graphic Communication*. <http://www.itu.int>, February 1998.
- [76] ITU-T: *Recommendation H.320 – Narrowband visual telephone systems and terminal Equipment*. <http://www.itu.int>, May 1999.
- [77] ITU-T: *Recommendation G.114 – One-way transmission time*. <http://www.itu.int>, May 2000.
- [78] ITU-T: *Recommendation H.323 – Packetbased multimedia communication systems*. <http://www.itu.int>, November 2000.
- [79] ITU-T: *Recommendation H.324 – Terminal for low bit-rate multimedia communication*. <http://www.itu.int>, March 2002.
- [80] JACOBSON, V., K. NICHOLS, and K. PODURI: *An Expedited Forwarding PHB*. RFC 2598, June 1999.
- [81] JOINT PHOTOGRAPHIC EXPERT GROUP: *JPEG2000 - Links to information*. <http://www.jpeg.org/jpeg2000.html>.
- [82] JOINT PHOTOGRAPHIC EXPERT GROUP: *Welcome to JPEG*. <http://www.jpeg.org/>.
- [83] KAUSER, N. AND CROWCROFT, J.: *End-to End Reliable Multicast Transport Protocol Requirements for Collaborative Multimedia Systems*.

- 
- In *17th IEEE Symposium on Reliable Distributed Systems, (SRDS'98)*, West Lafayette, Indiana, pages 425 — 430, October 1998.
- [84] KOCH, H.: *Lösung des Sichtbarkeitsproblems für das Videokonferenzsystem OCTOPUS*. Diplomarbeit, BTU Cottbus, Institut für Informatik, 2000.
- [85] KOIFMAN, A. and S. ZABELE: *Ramp: A Reliable Adaptive Multicast Protocol*. In *INFOCOM (3)*, pages 1442—1451, 1996.
- [86] KÖNIG, H. und M. ZÜHLKE: *OCTOPUS - Ein globales Videokonferenzsystem der Universitäten Hong Kong und Cottbus*. In: *Internetanwendungen für den Mittelstand. Konferenzband zur INFO'99, Potsdam*, Seiten 86 — 93. Dr. Wilke GmbH Verlag & Vertrieb, 1999.
- [87] KUHMÜNCH, C.: *Videoskalierung und Integration interaktiver Elemente in Teleteaching Szenarien*. Doktorarbeit, Universität Mannheim, 2001.
- [88] LBNL'S NETWORK RESEARCH GROUP: *wb - LBNL Whiteboard Tool*. <http://www-nrg.ee.lbl.gov/wb/>.
- [89] LECHA, M.: *Entwurf und Implementierung eines ATM-basierten Multicast-Transportprotokolls zur Übertragung multimedialer Ströme*. Diplomarbeit, BTU Cottbus, Institut für Informatik, 1998.
- [90] LEE, J.: *Explicit Multicast over Ethernet*. Internet Draft Proposal, January 2002.
- [91] LIN, J. C. and S. PAUL: *RMTP - A Reliable Multicast Protocol*. In *Proceedings of IEEE INFOCOM'96*, pages 1414 — 1424, March 1996.
- [92] LIU, Y., K. N. SRIJITH, L. JACOB, and A. L. ANANDA: *TCP-CM: A Transport Protocol for TCP-friendly Transmission of Continuous Media*. In *Proceedings of the 21st International Performance, Computing, and Communications Conference (IPCCC 2002)*, Phoenix, AZ, April 2002.
- [93] LONG, J.: *Effects of delayed irregular feedback on unskilled and skilled keying performance*. In *Ergonomics* 19, pages 183 — 202, 1976.

- [94] LUCENT TECHNOLOGIES: *Lucent Technologies' Research Software Distribution*. <http://www.bell-labs.com/topic/swdist/>.
- [95] LYONNET, F.: *Rendez-vous, the next generation internet videoconferencing tool*. <http://www.gaia-interactive.com/rv/>, 1999.
- [96] MAISS, G.: *DFN Videokonferenz – Aufbau eines neuen Dienstes*. DFN Mitteilungen, Heft 55:13 — 14, Februar 2001.
- [97] MAISS, G. und K. SCHAUERHAMMER: *DFN Videokonferenz (DFNVC)*. DFN Mitteilungen, Heft 58:4 — 5, März 2002.
- [98] MARKETING-MARKTPLATZ, INFORMATIONS-POOL FÜR MARKETING, WERBUNG, VERTRIEB, VERKAUFSFÖRDERUNG UND KOMMUNIKATION: *Europamarkt für Videokonferenzsystem-Endgeräte, ein Studienbeitrag im Informations-Pool für Marketing und Management*. <http://www.marketing-marktplatz.de/Marktforschung/Videosyst.htm>.
- [99] MCCANNE, S.: *Scalable Compression and Transmission of Internet Multicast Video*. Technical Report CSD-96-928, Computer Science Division (EECS), University of California, Berkeley, California 94720, July 1997.
- [100] MCCANNE, S., V. JACOBSON, and M. VETTERLI: *Receiver-driven Layered Multicast*. In *ACM SIGCOMM '96*, pages 117—130, 1996.
- [101] MICROSOFT GMBH: *Microsoft gmbh*. <http://www.microsoft.de>.
- [102] MICROSOFT GMBH: *Netmeeting*. <http://www.microsoft.de>.
- [103] MILLER, L. H.: *Response time in man-computer conversational transactions*. In *Proceedings of the Spring Joint Computer Conference*, volume 33, pages 267 — 277. AFIPS Press, Reston, 1968.
- [104] MILLS, D. L.: *Network Time Protocol (Version 3)*. RFC 1305, March 1992.
- [105] MOVING PICTURE EXPERTS GROUP: *MPEG Home Page*. <http://mpeg.telecomitalialab.com/>.



- [106] MOVING PICTURE EXPERTS GROUP: *Short MPEG-1 description*. <http://mpeg.telecomitalialab.com/standards/mpeg-1/mpeg-1.htm>, June 1996.
- [107] MOVING PICTURE EXPERTS GROUP: *Short MPEG-2 description*. <http://mpeg.telecomitalialab.com/standards/mpeg-2/mpeg-2.htm>, October 2000.
- [108] MOVING PICTURE EXPERTS GROUP: *MPEG-21 Overview v.5*. <http://mpeg.telecomitalialab.com/standards/mpeg-21/mpeg-21.htm>, October 2002.
- [109] MOVING PICTURE EXPERTS GROUP: *MPEG-7 Overview*. <http://mpeg.telecomitalialab.com/standards/mpeg-7/mpeg-7.htm>, July 2002.
- [110] MOVING PICTURE EXPERTS GROUP: *Overview of the MPEG-4 Standard*. <http://mpeg.telecomitalialab.com/standards/mpeg-4/mpeg-4.htm>, March 2002.
- [111] MUKHERJEE, B. and T. BRECHT: *Time-lined TCP for the TCP-friendly Delivery of Streaming Media*. In *International Conference on Network Protocols (ICNP), Osaka Japan*, pages 165 — 176, 2000.
- [112] NETWORKED MULTIMEDIA RESEARCH GROUP AT UNIVERSITY COLLEGE LONDON: *Videoconferencing Tool*. <http://www-mice.cs.ucl.ac.uk/multimedia/software/vic/>.
- [113] NETWORK RESEARCH GROUP OF LAWRENCE BERKELEY NATIONAL LABORATORY: *vat - LBNL Audio Conferencing Tool*. <http://www-nrg.ee.lbl.gov/vat/>.
- [114] NICKELSEN, J.: *Mischen und Multiplexen von Audioströmen in computergestützten Konferenzsystemen*. Diplomarbeit, TU Berlin, Fachgebiet Informatik.
- [115] OGUZ, N. AND AYANOGLU, E.: *Performance Analysis of Two-Level Forward Error Correction for Lost Cell Recovery in ATM Networks*. In *INFOCOM'95*, pages 728 — 737, April 1995.

- [116] O'NEIL, T. M.: *Applicationbased Quality of Service*. Whitepaper: <http://www.polycom.com/>.
- [117] OSIP: *The GNU oSIP library - GNU Project - Free Software Foundation (FSF)*. <http://www.gnu.org/software/osip/>.
- [118] OTY GMBH: *oty® Videokonferenz Systeme - Teleservice*. <http://www.multipoint.de>.
- [119] PAPADOPOULOS, C. and G. G. PARULKAR: *Retransmission-based error control for continuous media applications*. In *The 6th International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV)*, 1996.
- [120] PARTRIDGE, C.: *A Proposed Flow Specification*. RFC 1363, September 1992.
- [121] POLYCOM: *California CEO Company Info. Polycom Worldwide*. [http://www.polycom.com/company\\_info/0,1392,pw-451,00.html](http://www.polycom.com/company_info/0,1392,pw-451,00.html).
- [122] POLYCOM: *Your videoconferencing and speaker phone source! Polycom Worldwide*. <http://www.polycom.com/>.
- [123] POPOVICI, E., M. ZUEHLKE, R. MAHLO, and H. KOENIG: *BRAVIS - Ein Ansatz für geschlossene Mehrteilnehmer-Videokonferenzen über IP*. In IRMSCHER, K. and K.-P. FÄHNRICH (editors): *KIVS 2003, Leipzig*, pages 167 — 180. VDE Verlag GmbH, February 2003.
- [124] POPOVICI, E., M. ZUEHLKE, R. MAHLO, and H. KOENIG: *Consistency Support for a Decentralized Management in Close Multiparty Conferences Using SIP*. In *Accepted for 11th IEEE International Conference on Networks (ICON2003), Sydney, Australia*, September 2003.
- [125] PORTABLE APPLICATIONS STANDARDS COMMITTEE OF THE IEEE: *Information technology – Portable Operating System Interface (POSIX) – Part 1: System Application Program Interface (API)*. 1996-07-12 edition, ISO/IEC 9945-1, ANSI/IEEE Std. 1003.1., 1995.
- [126] QUASAR: *QUASAR – Quality of Service Architecture*. <http://www.ind.uni-stuttgart.de/Content/Quasar/index.html>.

- 
- [127] RADVISION: *viaIP<sup>TM</sup> MCU*. <http://www.radvision.com/NR/rdonlyres/AFA62E3E-9ADD-4702-BE01-DE4C44E19%D4D/468/viaIPMCU1.pdf>.
- [128] RAKEL, D.: *Entwicklung eines Prototypen zum Austausch von Videodaten im Videokonferenzsystem BRAVIS*. Studienarbeit, BTU Cottbus, Institut für Informatik, 2002.
- [129] RAMANATHAN, S., P. V. RANGAN, H. M. VIN, and T. KAEPFNER: *Optimal Communication Architectures for Multimedia Conferencing in Distributed Systems*. In *International Conference on Distributed Computing Systems*, pages 46 — 53, 1992.
- [130] RATNASAMY, S., P. FRANCIS, P. HANDLEY, R. KARP, and S. SHENKER: *A Scalable Content Addressable Network*. In *Proceedings of ACM SIGCOMM 2001*, 2001.
- [131] RODEO GROUP INRIA FRANKREICH: *Rendez-Vous, The Next Generation Video Conferencing Tool*. <http://www.lyonnet.org/IVStng/>.
- [132] ROSEN, E., A. VISWANATHAN, and R. CALLON: *Multiprotocol Label Switching Architecture*. RFC 3031, January 2001.
- [133] ROTH, R.: *Erster Klasse durchs Netz - Quality-of-Service-Architekturen für das G-WiN*. DFN Mitteilungen, Heft 57:16 — 18, November 2001.
- [134] ROWSTRON, A. and P. DRUSCHEL: *Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems*. Lecture Notes in Computer Science, 2218, 2001.
- [135] SCHAPHORST, R.: *Videoconferencing and Videotelephony: Technology and Standards*. Artech House Publishers Boston, London, 1999.
- [136] SCHMITT, J. and L. WOLF: *Quality of Service - An Overview*, 1997.
- [137] SCHULZRINNE, H., S. CASNER, R. FREDERICK, and V. JACOBSON: *RTP: A Transportprotocol for Real-Time Applications*. RFC 1889, January 1996.

- [138] SHI, S.: *A Proposal for A Scalable Internet Multicast Architecture*. Technical Report WUCS-01-02, Washington University in St. Louis, 2001.
- [139] SHNEIDERMAN, B.: *Response time and display rate in human performance with computers*. In *Computing Surveys*, 16, pages 265 — 285, 1984.
- [140] SIKORA, TH.: *MPEG-1 and MPEG-2 Digital Video Coding Standards*. [http://wwwam.hhi.de/mpeg-video/papers/sikora/mpeg1\\_2/mpeg1\\_2.htm](http://wwwam.hhi.de/mpeg-video/papers/sikora/mpeg1_2/mpeg1_2.htm).
- [141] SISALEM, D.: *End-To-End quality of Service Control Using Adaptive Applications*. In *IFIP Fifth International Workshop on Quality of Service (IWQOS'97)*, May 1997.
- [142] SISALEM, D. and H. SCHULZRINNE: *A Floor Control Application for Light-Weight Multicast Conferences*. In *ICT'98 - International Conference on Telecommunications, Thessaloniki, Greece*, pages 130—134, June 1998.
- [143] SONY ELECTRONICS INC.: *Videoconference with the Sony Contact PCS-1600 Videoconferencing System*. <http://www.sonycontact.com/>.
- [144] SPANIOL, O. and J MEGGERS: *Active Network Nodes for Adaptive Multimedia Communication*. In YONGCHAREON, T., F. A. AAGESEN, and V. WUWONGSE (editors): *Intelligence in Networks. IFIP TC6 WG6.7 5th International Conference, SMARTNET'99*, pages 1 — 18. Kluwer Academic Publisher, 1999.
- [145] STEINMETZ, R.: *Human Perception of Jitter and Media Synchronization*. *IEEE Journal on Selected Areas in Communications*, 14(1):61 — 72, January 1996.
- [146] STEINMETZ, R. and K. NAHRSTEDT: *Multimedia: Computing, Communications and Applications*. Prentice Hall, 1995.
- [147] STEINMETZ, R. and L. WOLF: *Quality of Service: Where are we?* In *Building QoS into Distributed Systems, 5th Int. Workshop on Quality of Service (IWQoS'97)*, New York, USA, pages 210—221, May 1997.

- [148] STILLER, B.: *Quality-of-Service — Dienstgüte in Hochleistungsnetzen*. International Thomson Publishing GmbH, Bonn, 1996.
- [149] SUN MICROSYSTEMS: *SUNForum Overview*. <http://www.sun.com/desktop/products/software/sunforum/>.
- [150] SUN MICROSYSTEMS, INC.: *docs.sun.com: XIL Reference Manual*. <http://docs.sun.com/db/doc/802-5862>.
- [151] SUN MICROSYSTEMS, INC.: *Manual pages of SunSolaris*. man pthread.
- [152] SUN MICROSYSTEMS, INC.: *POSIX Threads*. <http://www.sun.com/docs/>.
- [153] TANDBERG: *TANDBERG is a leading global provider of videoconferencing solutions - Homepage*. <http://www.tandberg.net>.
- [154] TEIXEIRA, L. M. L. AND MARTINS, M. I.: *Video Compression: The MPEG Standards*. In *ECMAST'96*, pages 615–634, May 1996.
- [155] TERENA - TRANS-EUROPEAN RESEARCH AND EDUCATION NETWORKING ASSOCIATION: *TERENA - TF-NGN - Next Generation Networking*. <http://www.terena.nl/tech/task-forces/tf-ngn/>.
- [156] THE OPEN MASH CONSORTIUM: *Open Mash*. <http://www.openmash.org>.
- [157] THE UCL NETWORKED MULTIMEDIA RESEARCH GROUP: *Robust Audio Tool (RAT)*. <http://www-mice.cs.ucl.ac.uk/multimedia/software/rat/>.
- [158] THE UCL NETWORKED MULTIMEDIA RESEARCH GROUP: *Session directory*. <http://www-mice.cs.ucl.ac.uk/multimedia/software/sdr/>.
- [159] THE YANKEE GROUP: *Exerpt of: Desktop Video Conferencing Market Analysis: Opportunities and Challenges on the Horizon*. [http://www.avistar.com/assets/docs/Yankee\\_Whitepaper\\_Excerpts.pdf](http://www.avistar.com/assets/docs/Yankee_Whitepaper_Excerpts.pdf), March 2002.

- [160] THÜRMAN, U. und M. ZITTERBART: *IP-Telefonie über Differentiated Services*. In: *Proceedings of the 1st IP-Telephony Workshop (IPTel 2000)*, April 2000.
- [161] TOPOLCIC, C.: *Experimental Internet Stream Protocol, Version 2 (ST-II)*. RFC 1190, October 1990.
- [162] TROSSEN, D. and K.-H. SCHARER: *SCCS: Scalable Conferencing Control Service*. In *Proceedings of 7th IEEE International Conference on Computer Communication and Networks (IC3N'98)*, pages 698 — 706, October 1998.
- [163] TURLETTI, T.: *The INRIA Video conferencing System (IVS)*. *ConneXions - The Interoperability Report Journal*, 8:20 — 24, 1994.
- [164] UCL NETWORK AND MULTIMEDIA RESEARCH GROUP: *Common Multimedia Library*. <http://www-mice.cs.ucl.ac.uk/multimedia/software/common/index.html>.
- [165] UNIVERSITY OF MANNHEIM: *multimedia lecture board*. <http://www.informatik.uni-mannheim.de/informatik/pi4/projects/mlb/>.
- [166] VCON DESKTOP VIDEOCONFERENCING GMBH: *VCON steht für Videoconferencing*. <http://www.vcon.de>.
- [167] VEREIN DES DEUTSCHEN FORSCHUNGSNETZES: *Das Deutsche Forschungsnetz G-WiN*. <http://www.dfn.de>.
- [168] VEREIN DES DEUTSCHEN FORSCHUNGSNETZES: *Videokonferenz Portal*. <https://www.vc.dfn.de>.
- [169] VON WEIHER, S. und B. WAGNER: *Tagebuch der Telekommunikation: Von 1600 bis zur Gegenwart*. vde-verlag gmbh, Berlin und Offenbach, 1991.
- [170] WAINHOUSE RESEARCH: *Wainhouse Research*. <http://www.wainhouse.com/>.
- [171] WAKEMAN, I.: *Congestion Control for Packetised Video in the Internet*. PhD thesis, UCL, 1995.

- 
- [172] WALSH, G.: *availability.com: Desktop Video Conferencing Market Analysis: Opportunities and Challenges on the Horizon*. <http://www.availability.com/research/industry/index.cfm?fuseaction=news%&id=57D0123A-720C-49CC-B51B1E5384BF4AC9>.
- [173] WHETTEN, B., S. KAPLAN, and T. MONTGOMERY: *A high performance totally ordered multicast protocol*. In *Proceedings of INFOCOMM 95*, April 1995.
- [174] WHITE PINE: *CUSeeMe*. <http://www.wpine.com/Products/CUSeeMe/>.
- [175] WWW.MBONE.DE: *Multicast Backbone Over the interNEt*. <http://www.mbone.de>.
- [176] XIE, J., R. R. TALPADE, A. MCAULEY, and M. LIU: *AMRoute: ad hoc multicast routing protocol*. *Mobile Networks and Applications*, 7(6):429–439, 2002.
- [177] XILINX®: *Video Conferencing*. [http://www.xilinx.com/esp/dvt/collateral/video\\_conferencing.pdf](http://www.xilinx.com/esp/dvt/collateral/video_conferencing.pdf).
- [178] YAVATKAR, R., J. GRIFFOEN, and M. SUDAN: *A Reliable Dissemination Protocol for Interactive Collaborative Applications*. In *ACM Multimedia*, pages 333–344, 1995.
- [179] ZHANG, B., S. JAMIN, and L. ZHANG: *Host Multicast: A Framework for Delivering Multicast to End Users*. In *IEEE INFOCOM*, 2002.
- [180] ZHANG, L., S. DEERING, D. ESTRIN, S. SHENKER, and D. ZAPPALA: *RSVP: a new Resource reSerVation Protocol*. *IEEE Network Magazine*, 3:8–18, September 1993.
- [181] ZHAO, W., D. OLSHEFSKI, and H. SCHULZRINNE: *Internet Quality of Service: an Overview*, 2000.
- [182] ZHUANG, S., B. ZHAO, A. JOSEPH, R. KATZ, and J. KUBIATOWICZ: *Bayeux: An Architecture for Scalable and Fault-tolerant WideArea Data Dissemination*. In *Eleventh International Workshop on Network*

*and Operating System Support for Digital Audio and Video (NOSSDAV 2001)*, June 2001.

- [183] ZUEHLKE, M. and H. KOENIG: *GCP – A Group Communication Protocol for Supporting Closed Groups in the Internet*. In MARTIKAINEN, O., K. RAATIKAINEN, and J. HYVÄRINEN (editors): *Smart Networks. IFIP TC6 WG6.7 7th International Conference on Intelligence in Networks (SMARTNET 2002)*, Saariselkä, Finland, pages 97 — 102. Kluwer Academic Publishers, 2002.
- [184] ZUEHLKE, M. and H. KOENIG: *Voting Based Bandwidth Management in Multiparty Video Conferences*. In BOAVIA, F., O. MONTEIRO, and J. ORVALHO (editors): *Protocols and Systems for Interactive Distributed Multimedia. IEEE Joint International Workshop on Interactive Distributed Multimedia Systems / Protocols for Multimedia Systems (iDMS-PROMS 2002)*, Coimbra, Portugal, pages 202 — 215. Springer-Verlag, vol. LNCS 2515, November 2002.