

Universität Karlsruhe
Fakultät für Informatik
76128 Karlsruhe

Netzwerk-Management und Hochgeschwindigkeits- Kommunikation

Teil XIII

Seminar WS 1995/96

Herausgeber:
Elmar Dorner
Markus Hofmann
Claudia Schmidt
Jochen Seitz

Universität Karlsruhe
Institut für Telematik

Interner Bericht 14/96

Zusammenfassung

Der vorliegende Interne Bericht enthält die Beiträge zum Seminar „Netzwerk-Management und Hochgeschwindigkeits-Kommunikation“, das im Wintersemester 1995/96 zum 13. Mal stattgefunden hat.

Die Themenauswahl kann grob in folgende fünf Blöcke gegliedert werden:

1. Der erste Block befaßt sich mit der Problematik, unterschiedliche Dienste und damit auch verschiedene Anforderungen an die zu erbringende Dienstqualität in ein Kommunikationssystem zu integrieren.
2. Im zweiten Block wird das Management von FDDI-Netzen näher untersucht, wobei besonderes Augenmerk auf der Unterstützung von isochronem Datenverkehr liegt.
3. Der dritte Block befaßt sich mit aktuellen Problemen aus dem Bereich der Mobilkommunikation, insbesondere der Wegwahl in Mobil-Netzen.
4. Im Zentrum des vierten Blocks steht das Internet und dessen Migration hin zu ATM. Dabei spielen vor allem Gruppenkommunikation und Dienstlokalisierung eine große Rolle.
5. Der letzte Block schließlich stellt die Verwaltung von ATM-Netzen in den Mittelpunkt, indem verschiedene Normungsvorschläge zu diesem Thema ausgeführt werden.

Abstract

This Technical Report includes student papers produced within small lessons called seminar of “Network Management and High Speed Communications”. For the thirteenth time this seminar has attracted a huge number of diligent students, proving the broad interest in topics of network management and high speed communications.

The topics of this report may be divided into five blocks:

1. First, the topics of different transport services with according constraints on the required quality of service, that shall be integrated in a communication subsystem, are addressed.
2. Second, the management of FDDI networks is introduced, especially focused on the possibility of isochronous data transmission.
3. The third block deals with actual problems in mobile communication, e.g., routing in mobile networks.
4. The Internet and its migration towards ATM technology is the main point in the fourth block. Special topics are group communication and service localization.
5. Finally, the management of ATM and the current work on related standards are introduced.

Vorwort

Das Seminar „Netzwerk-Management und Hochgeschwindigkeits-Kommunikation“ erfreute sich in den letzten Jahren immer größerer Beliebtheit. Gerade heutzutage sind Stichworte wie „ATM“, „Quality of Service“, „Mobil-Kommunikation“ oder „Internet“ in aller Munde. Daher sind die Forschungsgebiete in diesen Bereichen auch von allgemeinem Interesse, so daß sie eine derartige Vielzahl von innovativen Arbeiten aufweisen können, deren Behandlung in anderen Lehrveranstaltungen so detailliert nicht möglich ist.

Jetzt liegt auch der nunmehr dreizehnte Seminarband als Interner Bericht vor. Durch das engagierte Mitarbeiten der beteiligten Studenten konnte so zumindest ein Ausschnitt aus dem komplexen und umfassenden Themengebiet klar und übersichtlich präsentiert werden. Für den Fleiß und das Engagement der Seminaristin und Seminaristen sei daher an dieser Stelle recht herzlich gedankt.

Die ausgesprochen gute Resonanz bei den Studenten hat uns veranlaßt, auch im Sommersemester 1996 ein derartiges Seminar — natürlich mit geändertem aktuellem Inhalt — durchzuführen, so daß bald ein weiterer Interner Bericht mit neuen Forschungsergebnissen aus innovativen Tagungsbeiträgen erscheinen wird. Doch vorerst sollen im vorliegenden Band folgende Themengebiete vorgestellt werden:

Der erste Schritt zu einem dienstintegrierenden Internet

Das heutige Internet ist, bedingt durch die Gleichbehandlung aller Datenströme, für eine Dienstintegration, d.h. die gleichzeitige Übertragung von Audio, Video und klassischen Daten in einem einzigen Netzwerk, nicht geeignet. Dieser Beitrag beschreibt die Aktivitäten der „Integrated Services“-Arbeitsgruppe der IETF (Internet Engineering Task Force), die sich mit der Spezifikation eines erweiterten Dienstmodells für das Internet beschäftigt.

OSI95: Der OSI-Transportdienst der 90er Jahre

Das ESPRIT-Projekt OSI95 (Akronym für „High Performance OSI Protocols with Multimedia Support on HSLANs and B-ISDN“) begann 1990 und befaßt sich mit qualitätsorientierten Transportdiensten. Dieser Beitrag präsentiert die Entwicklungen, die das Projekt beeinflussten, sowie die neu entworfenen Transportdienste (verbindungsorientiert und verbindungslos).

Im Vergleich: QoS-Management für multimediale Anwendungen

Dieser Beitrag betrachtet drei neue Ansätze, die im Hinblick auf die Unterstützung von Kommunikationsdiensten für multimediale Anwendungen entwickelt wurden. Die grundlegenden Konzepte und Ziele der einzelnen Ansätze werden zunächst vorgestellt. Basierend darauf wird ein Vergleich der Ansätze durchgeführt, wobei die Architekturen und die Funktionen zum QoS-Management eine zentrale Rolle spielen.

Station Management in FDDI-Netzen

Der vorliegende Beitrag beschreibt die Aufgaben und Funktionen der Station Management Komponente innerhalb eines FDDI-Knotens. Schwerpunkt der Betrachtungen ist dabei das Zusammenspiel der einzelnen Teilkomponenten bei der Überwachung des Ringes und der Bereitstellung von Verwaltungsdiensten.

Transportprotokolle mit Unterstützung von isochronem Datenverkehr

Die Übertragung von Audio-/Video-Datenströmen, wie sie bei Multimedia-Anwendungen häufig benötigt wird, wird durch herkömmliche Transportprotokolle nur schlecht unterstützt. Mit der Zielsetzung, dies optimal zu unterstützen, wurden mehrere Protokolle entwickelt. Im Rahmen dieses Beitrags wurden ein Vergleich und eine Bewertung der vorgestellten Ansätze durchgeführt.

Mobile Datenkommunikation in GSM-Netzen

Mit zunehmender Bedeutung der weltweiten Rechnerkommunikation wächst auch der Wunsch zur Mobilität. Was beim Telefon mittlerweile schon zu einer Selbstverständlichkeit geworden ist — die weltweite Erreichbarkeit — befindet sich im Bereich der Datenkommunikation noch im Anfangsstadium. Der vorliegende Beitrag stellt neben Struktur und Verwaltungseinheiten des GSM-Netzes auch die neuen Dienste, die durch die Einführung von GSM möglich wurden, vor.

Routingverfahren in Mobil-Netzen

Ein Hauptproblem bei der möglichst nahtlosen Integration von mobilen Netzwerkteilnehmern in bestehende Netze stellt — neben der Adressierung — die Wegewahl dar. Im Rahmen des Beitrags wurden verschiedene Ansätze zur Einbindung von Mobilteilnehmern in das Internet besprochen. Dabei zeigt sich, daß es noch kein optimales „Mobile-IP“ in Bezug auf die Wegewahl gibt.

Multicast-Routing im Internet

Mit zunehmender Bedeutung der verteilten Teamarbeit und dem damit verbundenen Anwachsen der Kommunikationsgruppen im Internet werden skalierbare und effiziente Routing-Verfahren immer wichtiger. Im Rahmen dieser Seminararbeit wird diese Problematik am Beispiel zweier bekannter Multicast-Routingverfahren motiviert. Darauf aufbauend werden mit Core Based Trees (CBT) und Protocol Independent Multicast (PIM) alternative Lösungsansätze vorgestellt.

Protokollmechanismen zur Lokalisierung von Diensten im Internet

Im Zuge der stetig steigenden Popularität des Internets und dem damit verbundenen Anwachsen der Netzbelastung werden Netzdienste im Internet häufig repliziert und

in unterschiedlichen Regionen angeboten. Dies ermöglicht dem Anwender, den für ihn günstigsten Dienst auszuwählen und somit den Datenverkehr auf einen lokalen Bereich zu beschränken. Der Beitrag diskutiert mehrere Verfahren und Protokollmechanismen, welche die Lokalisierung eines geeigneten Dienstes unterstützen.

Protokolle zur Gruppenkommunikation in globalen Netzen

Der Datenaustausch innerhalb einer Gruppe von Benutzern stellt bei modernen verteilten Anwendungen eine grundlegende Kommunikationsbeziehung dar. An einer Videokonferenz können beispielsweise mehrere hundert Personen beteiligt sein, die über den gesamten Erdball hinweg verteilt sind. Der Datenaustausch in einer solch umfassenden Gruppe von Benutzern erfordert neuartige Protokollmechanismen. In dem vorliegenden Beitrag werden unterschiedliche Protokollarchitekturen zur Unterstützung einer solchen Multicast-Kommunikation gegenübergestellt.

IP und ATM - Konkurrenz oder Symbiose?

Häufig werden in Diskussionen über zukünftige Entwicklungen im Telekommunikationsbereich ATM-Netze als direkte Konkurrenz zum heutigen Internet dargestellt. Im Rahmen des Seminarbeitrags werden die Grundlagen beider Protokollarchitekturen kurz vorgestellt und darauf aufbauend eine Bilanz über die Vor- und Nachteile der Technologien gezogen. Schließlich werden Möglichkeiten einer Koexistenz beider Ansätze bzw. einer Integration von ATM in das Internet diskutiert.

Das ATM-Management-Modell

Mit ATM sollen integriert verschiedene Dienste und Anforderungen erbracht werden können. Entsprechend muß die Verwaltung eines derartigen Netzes auch wesentlich umfangreicher und somit komplexer gestaltet werden. Das ATM-Forum hat zur Verwaltung eines ATM-Netzes ein Management-Modell definiert, das auf unterschiedliche Schnittstellen im Netz und somit auf unterschiedliche Netzbetreiber eingeht. Dieses Modell wird hier vorgestellt.

ATM-Management mittels SNMPv2

Auch in der Internet-Gemeinde, in der ATM zunehmend populärer und immer heißer diskutiert wird, sind bereits Ansätze zur Verwaltung eines ATM-Netzes definiert worden. Thema des vorliegenden Beitrags ist ein Request for Comments, der das Management von ATM auf der Basis des Simple Network Management Protocols version 2 (SNMPv2) festlegt. Dabei ist nicht nur das Managementprotokoll von Interesse, sondern vielmehr auch die Managementinformation, die im Rahmen des SNMPv2-Managements zur Verfügung gestellt werden muß.

ATM-LAN-Emulation und seine Verwaltung

Obwohl ATM als Grundlage für Breitband-ISDN für Weitverkehrsnetze konzipiert ist, kann die Technologie auch im Bereich der lokalen Netze gewinnbringend eingesetzt werden. Um eine weitere Verwendung bereits existierender Soft- und Hardware in lokalen Netzen zu ermöglichen, wurde im Rahmen des ATM-Forums die LAN-Emulation für ATM definiert. Da hierzu mehrere zentrale Komponenten benötigt werden, muß entsprechend auch eine Verwaltung eines derartigen ATM-basierten lokalen Netzes standardisiert werden. Einer der Standards dafür wird im abschließenden Beitrag vorgestellt.

Inhaltsverzeichnis

Tobias Schulz-Hess:

Der erste Schritt zu einem dienstintegrierenden Internet	1
---	----------

Alexander Rösner:

OSI95: Der Transportdienst der 90er Jahre	17
--	-----------

Robert Clauß:

Im Vergleich: QOS-Management für multimediale Anwendungen . . .	31
--	-----------

Thomas Merkle:

Station-Management in FDDI-Netzen	47
--	-----------

Roland Forberger:

Transportprotokolle mit Unterstützung von isochronem Datenverkehr	61
--	-----------

Martin Kemmler:

Mobile Datenkommunikation in GSM-Netzen	73
--	-----------

Simon Oser:

Routingverfahren in Mobil-Netzen	93
---	-----------

Frank Schadt:

Multicast-Routing im Internet	109
--	------------

Jens Clever:

Mechanismen zur Gruppenkommunikation in globalen Netzen	123
--	------------

Virginie J. L. Njeck Nounga:

IP und ATM: Konkurrenz oder Symbiose?	139
--	------------

Manfred Rohrmüller:

Protokollmechanismen zur Lokalisierung von Diensten im Internet . .	151
--	------------

Michael Ebner:

Das ATM-Management-Modell	167
--	------------

Rudi Somen:

ATM-Management mittels SNMPv2	185
--	------------

Daniel Schempf:

ATM-LAN Emulation und seine Verwaltung	199
---	------------

Abbildungsverzeichnis	213
--	------------

Tabellenverzeichnis	215
--------------------------------------	------------

Literatur	217
----------------------------	------------

Der erste Schritt zu einem diensteintegrierenden Internet

Tobias Schulz-Hess

Kurzfassung

Das heutige Internet ist für eine Dienstintegration, d.h. die gleichzeitige Übertragung von Audio, Video und klassischen Daten in einem einzigen Netzwerk nicht geeignet. Dies ist insbesondere darauf zurückzuführen, daß bei der Übertragung alle Datenströme gleich behandelt werden und somit die Dienstanforderungen von Audio- und Videoströmen nicht erfüllt werden können.

Die Arbeitsgruppe *intserv* des IETF hat sich mit diesem Problem beschäftigt; die aktuellen Ergebnisse ihrer Arbeit werden hier vorgestellt. Zum einen wird auf einige neue Dienste eingegangen, die von *intserv* vorgeschlagen wurden, zum anderen wird diskutiert, wie die neuen Dienste bei Beibehaltung des bisherigen Dienstes eingeführt werden sollen.

1 Einführung

1.1 Was ist die IETF?

Jeder, der schon intensiver im Internet gearbeitet hat, wird festgestellt haben, daß das Internet streckenweise sehr überlastet ist. Bei dem exponentiellen Wachstum des Internets steht ein Zusammenbruch des Netzes demnach in nicht allzuweiter Ferne.

Um dieser Katastrophe zu begegnen, wurde im Januar 1986 ein Forum für technische Koordination, das IETF (Internet Engineering Task Force, vgl. [IETFa]), von Mitgliedern der US Defense Advanced Projects Agency (DARPA) gegründet, die am ARPANET, dem US Defense Data Network (DDN) und dem Internet Core Gateway System arbeiten. Seit dieser Zeit wuchs das IETF zu einer großen, offenen, internationalen Gemeinschaft von Netzwerkdesignern, -konstrukteuren, -verkäufern und -forschern, die sich alle mit der Verbesserung und Erweiterung der Internetarchitektur und einem reibungslosen Betrieb des Internets beschäftigen.

Die Aufgaben, die sich das IETF gesetzt hat, sind nach [IETFb]:

1. Vorschlagen von Lösungen zu aktuellen, dringenden, betrieblichen und technischen Problemen im Internet.
2. Spezifizierung der Entwicklung oder Benutzung von Protokollen und der Architektur des Internets um technische Probleme im Internet zu lösen.
3. Erleichterung des Technologietransfers von der Internet Research Task Force (IRTF) zur allen Internetnutzern.
4. Bereitstellung eines Forums zum Austausch von wichtigen Informationen innerhalb der Internetgemeinschaft zwischen Verkäufern, Benutzern, Forschern, Anbietern und Netzwerkmanagern.

1.2 Was ist die Arbeitsgruppe *intserv* und was sind ihre Aufgaben?

Innerhalb der IETF gibt es neun Bereiche, für Aufgaben wie Sicherheit, Routing oder Entwicklung von IP Next Generation (IPng), die auch untereinander in engem Kontakt stehen. Diese Bereiche sind in verschiedene Arbeitsgruppen untergliedert.

In dem Bereich, der für den Transport innerhalb des Internets verantwortlich ist (*Transport Area*), gibt es eine Arbeitsgruppe, die versucht einen möglichst kleinen gemeinsamen Nenner zu finden, um das Internet in eine zuverlässige diensteintegrierende Kommunikationsinfrastruktur zu wandeln. Experimente haben ergeben, daß diensteintegrierende Protokolle besser den Bedürfnissen von Internetanwendungen, speziell von Audio- und Videoanwendungen, gerecht werden. Diese Arbeitsgruppe nennt sich Integrated Services, kurz *intserv*.

Die Aufgaben von *intserv* gliedern sich (nach [IETFc]) in drei Schritte:

1. Zuerst müssen die Dienste definiert und dokumentiert werden, die angeboten werden sollen.
2. Dann muß dokumentiert werden, wie die Anwendungen und die Netzelemente mit diesen neuen Diensten umzugehen haben.
3. Schließlich soll noch darauf eingegangen werden, wie sich diese neuen Dienste zusammen mit dem alten Dienst in einem Netz zusammen nutzen lassen.

2 Neue Dienste im Internet

Bisher wurde im Internet nur ein einziger Dienst angeboten, der sogenannte bestmögliche Dienst (*best effort service*). Bei diesem Dienst werden keinerlei Aussagen über die Zustellzeit der Pakete gemacht, was in der Praxis die unangenehme Auswirkung hat, daß sich einzelne Pakete teilweise erheblich verzögern können. Damit ist dieser Dienst auch für Anwendungen, die Pakete in einer gewissen Zeit zugestellt haben wollen (Echtzeitanwendungen), ungeeignet. Solche Echtzeitanwendungen gibt es bereits heute, beispielsweise Audio- und Videoanwendungen, die aber im Internet nur unter sehr günstigen Bedingungen funktionieren.

Diesen Mangel will die Arbeitsgruppe *intserv* dadurch ausgleichen, indem neue Dienste eingeführt werden, die alle mehr oder weniger einen Echtzeitdienst darstellen. Im Moment sind drei verschiedene Dienste ausgearbeitet, die im weiteren einzeln vorgestellt werden; konkret sind das der Dienst mit kontrollierter Verzögerung, der vorhersagbare und der garantierte Dienst. Dabei genügt es nicht, den Dienst isoliert für sich zu behandeln, sondern es muß auch auf die interne Behandlung des Dienstes in den Netzelementen (Endrechner, Router, etc.) sowie auf die exportierten Informationen zwischen den Netzelementen eingegangen werden, denn gerade an diesen Stellen kommt es zu den Überlastungen, die für die Echtzeitdienste aufgelöst werden müssen. Deshalb werden alle Dienste in der gleichen Weise vorgestellt, indem zuerst die Idee, die hinter dem Dienst steckt, angesprochen wird, dann besprochen wird, was das einzelne Netzelement für Leistungen zu erbringen hat, weiterhin, mit welchen Informationen sich die Netzelemente

untereinander versorgen müssen und schließlich, wie sich dieser Dienst dann für den Anwender darstellt.

2.1 Dienst mit kontrollierter Verzögerung (*Controlled Delay Service*)

Motivation Dieser Dienst ist von der Arbeitsgruppe *intserv* nach [SPW95] für Anwendungen gedacht, die sich selbständig und dynamisch an den notwendigen Dienst und auch an die Übertragungsverzögerung anpassen können. Dies sind Anwendungen, bei denen die Datenpakete auch mit einer gewissen Verzögerung ankommen können und die ihren Dienst je nach ihren Bedürfnissen und denen des Netzes ändern können. Konkret können das Playback-Anwendungen, wie Audio- und Videokonferenzsysteme sein, die dann auch die Abspielzeitpunkte selbständig verschieben können. Verlangen die Anwendungen jedoch gewisse garantierte Paketzustellzeiten, sollten sie einen anderen Dienst wählen.

Der Dienst mit kontrollierter Verzögerung gibt keine Auskunft darüber, wie sehr sich die einzelnen Pakete in Zukunft verzögern können. Vielmehr gibt er nur Information darüber, wie sehr sich die einzelnen Pakete in der Vergangenheit verzögert haben. Außerdem verspricht dieser Dienst eher, Überlastungen zu verhindern, indem zusätzlicher Verkehr nicht zugelassen wird. Was genau unter „Überlastung“ zu verstehen ist, wird von *intserv* nicht spezifiziert, sondern bleibt dem einzelnen Netzbetreiber überlassen. Obwohl also der Dienst eine gewisse Kontrolle der Paketverzögerung zuläßt, schränkt er diese Verzögerung weder statistisch noch statisch ein.

Interne Datenbehandlung der Netzelemente Die Netzelemente müssen sicherstellen, daß die Paketverzögerungen überwacht werden. Dies muß durch eine aktive Zugangskontrolle (*admission control*) erfolgen. Eine Überlastung des Netzelements ist nicht im Sinne des Dienstes mit kontrollierter Verzögerung, das Element muß vielmehr in der Lage sein, das Datenaufkommen abzulehnen, das dazu führen würde, daß sich größere Verzögerungen ergeben. Es gibt jedoch keine quantitative Beschreibung, was nun unter einer „großen Verzögerung“ genau zu verstehen ist.

Es gibt in diesem Dienst drei verschiedene logische Stufen. Ein Netzelement darf auch weniger Stufen implementieren, aber nach außen müssen diese drei Stufen sichtbar sein. Diese Stufen haben verschieden Grade der Kontrolle der Paketverzögerung, wobei Stufe eins eine genauere Kontrolle aufweist als Stufe zwei, die wiederum eine genauere Kontrolle zeigt als Stufe drei. Dies bedeutet jedoch nicht, daß die Pakete einer niedrigeren Stufe eine kürzere Verzögerung erfahren. Alle drei Stufen sollten einen besseren Dienst bieten (z.B. eine genauer gemessene Paketverzögerung), als der unkontrollierte bestmögliche Dienst.

Der Dienst mit kontrollierter Verzögerung muß eine recht geringe Paketverlustrate aufrechterhalten können. Obwohl Pakete verloren gehen können, stellt jeder merkliche Datenverlust einen „Fehler“ des Zugangskontrollalgorithmus dar, denn dieser Algorithmus ist dafür verantwortlich, daß genau solche Verluste nicht auftreten (vgl. dazu auch Abschnitt 3.1).

Exportierte Informationen Für jede der drei Stufen des Dienstes mit kontrollierter Verzögerung exportiert das Netzelement drei Meßergebnisse von Verzögerungsmessungen (insgesamt also neun Werte). Jeder dieser Parameter stellt die maximale Verzögerung über ein bestimmtes Zeitintervall T dar. Diese Parameter werden auch „Charakterisierungsparameter“ genannt, da sie den Dienst nach außen hin charakterisieren. Die drei Zeitintervalle T sind eine Sekunde, 60 Sekunden (eine Minute) und 3.600 Sekunden (eine Stunde). Die Parameter müssen nicht exakt stimmen, sie können sogar Mittelwerte über einen Zeitraum von $2T$ darstellen.

Diese Charakterisierungsparameter müssen nicht notwendigerweise auf genauen Messungen basieren. Im einzelnen können diese Messungen auf Schätzungen über Paketverzögerungen oder auf Gesamtmessungen der Warteschlangenlänge basieren. Außerdem müssen kurzfristige Schwankungen bei der Verzögerung nicht berücksichtigt werden. Diese Lockerungen wurden eingeführt, um Entwerfer von Netzelementen nicht mit der Ausarbeitung genauer Meßalgorithmen zu belasten.

Resultierender Dienst Der resultierende Dienst von Endelement zu Endelement ist also ein Dienst, der den Anwendungen mehrere Stufen anbietet, aus denen sie die passende Stufe auswählen können. Weiterhin verspricht der Dienst, daß alle Verzögerungsstufen so kontrolliert werden, daß sie einen besseren Dienst als den bestmöglichen Dienst darstellen. Absolute Verzögerungen werden von diesem Dienst jedoch genausowenig garantiert, wie auch kurzfristige Schwankungen nicht verhindert werden.

Dieser Dienst ist unter anderem für adaptive Playback-Anwendungen (*playback applications*) gedacht. Diese Anwendung mögen in der Lage sein, verschiedene Grade von de facto Paketverlusten (die darauf basieren, daß Pakete zu spät ankommen) hinzunehmen. Zu einer gegebenen Stufe des Dienstes mit kontrollierter Verzögerung kann die Anwendung die Verlustrate ändern, indem sie die Verzögerungsanpassungsfunktion und damit auch den Abspielpunkt mehr oder weniger ändert. Die Anwendung muß selbst die Einstellungen für die minimale Abspielverzögerung einerseits und dem Prozentsatz der Pakete, die zu spät ankommen um noch verwendet werden zu können, andererseits ideal bestimmen und wählen. Wenn selbst die beste Einstellung für eine gegebene Stufe nicht ausreichen sollte, kann die Anwendung immer noch eine Stufe mit einer besser kontrollierten Verzögerung wählen.

2.2 Vorhersagbarer Dienst (*Predictive Service*)

Motivation Wie in [SP95b] angegeben werden beim vorhersagbaren Dienst eine maximale und (optional) eine obere Schranke der vergangenen Messungen für die Paketverzögerung angegeben. Die Wahrscheinlichkeit, daß diese Schranke nicht überschritten wird ist recht hoch, wenn angenommen wird, daß keine größeren Ausfälle bei den Zwischenelementen auftreten oder sich der Weg der Pakete nicht ändert. Wenn es die Netzressourcen zulassen, versucht der Dienst, die Pakete in einer kürzeren als der maximalen Zeit zuzustellen.

Der vorhersagbare Dienst ist gedacht für Playback-Anwendungen, die gelegentlich fehlende Abspielpunkte dulden, aber dennoch eine garantierte maximale Verzögerung ver-

langen. Die Verwendung von Verzögerungsschranken hat dabei zwei Funktionen: Zum einen kann eine Anwendung, die nicht bereit ist, ständig ihren Dienst und ihre Abspielzeitpunkte zu ändern, bereits im Voraus wissen, was für Paketverzögerungszeiten sie zu erwarten hat. Zum anderen erlauben diese Schranken solchen Anwendungen, ihre Abspielpunkte richtig zu setzen.

Interne Datenbehandlung der Netzelemente Auch bei diesem Dienst gibt es drei verschiedenen logische Stufen, die nicht alle von dem Netzelement implementiert sein müssen, jedoch nach außen hin unterscheidbar sind. Jede Stufe ist mit einer Verzögerungsschranke verbunden, wobei Stufe eins die geringste Paketverzögerungszeit aufweist und Stufe drei die größte. Das Netzelement muß sicherstellen, daß die Paketverzögerungen unter der gesetzten Schranke bleiben. Gelegentliche Verletzungen gegen diese Schranken werden zwar geduldet, aber diese Verletzungen sollten so selten wie nur möglich auftreten. Dazu wird eine aktive Zugangskontrolle verwendet, wie sie auch schon beim Dienst mit kontrollierter Verzögerung verlangt wurde und in Abschnitt 3.1 genauer erklärt wird. Dies muß eine aktive Steuerung sein, so daß das Element in der Lage sein muß, die Zugangsberechtigung der Dienstanfragen prüfen zu können. Eine Überlastung des Elements stellt keine Lösung für den vorhersagbaren Dienst dar.

Exportierte Informationen Es werden für jede Stufe des Dienstes zwei Schranken nach außen hin bekanntgegeben. Die erste Schranke ist für die Verzögerung der Pakete. Für diese Schranke werden ähnlich dem Dienst mit kontrollierter Verzögerung drei Werte angegeben, die das Verkehrsaufkommen über die letzte Sekunde, Minute und die letzte Stunde darstellen. Unabhängig vom Verkehrsaufkommen wird das Netzelement aber versuchen, die angegebenen Werte zu unterbieten. Deshalb kann optional auch noch eine zweite Schranke mit angegeben werden, die das aktuelle Verkehrsaufkommen darstellt.

Resultierender Dienst Der resultierende Dienst von Endelement zu Endelement ist also ein Dienst, der den Anwendungen mehrere Stufen von maximalen Verzögerungen anbietet, von denen sich die Anwendungen die passende Stufe auswählen können. Eine adaptive Anwendung, der die aktuelle Stufe nicht ausreicht, kann zu einer besseren Stufe wechseln um den Dienst zu verbessern (oder zu einer schlechteren Stufe, wenn der aktuelle Dienst besser als benötigt ist). Außerdem verspricht der Dienst, daß, von wenigen Ausnahmen abgesehen, die zu erwartenden Paketverzögerungen unter der angegebenen Schranke bleiben.

Dieser Dienst geht davon aus, daß adaptive Anwendungen selbst in der Lage sind, die Paketverzögerungen zu messen und sich daran anzupassen, wobei die Schranken nur Hinweise sind. Sie mögen dann hilfreich sein, wenn die Anwendung versucht, neue Abspielpunkte zu berechnen oder wenn sie entscheiden muß, ob sie die Stufe des Dienstes wechselt, jedoch gibt es keine Garantie, daß diese Schranken in einigen Fällen nicht auch überschritten werden.

2.3 Garantierter Dienst (*Guaranteed Service*)

Motivation Der garantierte Dienst stellt — wie in [SP95a] vorgeschlagen — mathematisch beweisbar sicher, daß die Verzögerung von Endelement zu Endelement eine vorher gesetzte Schranke nicht überschreiten, vorausgesetzt alle Elemente auf diesem Pfad unterstützen den garantierten Dienst. Darüberhinaus wird noch angenommen, daß keine hardwarebedingten Ausfälle bei Elementen auf dem Paketweg auftreten und sich der Paketweg auch nicht ändert. Derartige Fehlerfälle müssen von höheren Protokollen behandelt werden.

Dieser Dienst ist entstanden für Anwendungen, die eine garantierte Verzögerung unbedingt verlangen. Dies sind Anwendungen, die einen Paketverlust (auf Grund von zu spät angekommenen Paketen) nicht dulden und für Anwendungen mit Echtzeitbedürfnissen. Der garantierte Dienst stellt sicher, daß die Pakete innerhalb einer garantierten Zeit eintreffen und auch nicht wegen Überfüllung von einzelnen Warteschlangen verloren gehen. Programmierer von Playback-Anwendungen sollten davon ausgehen, daß die Pakete oft sehr viel früher eintreffen, als in der vom garantierten Dienst zugesagten Zeit, und deshalb im empfangenden Endsystem zwischengespeichert werden müssen bis die Anwendung sie verarbeiten kann.

Interne Datenbehandlung der Netzelemente Es werden bei dem garantierten Dienst mehrere Stufen des Dienstes angeboten. Die Stufe des Dienstes ist in jedem Netzelement charakterisiert durch eine Bandbreite R und eine Puffergröße B . R repräsentiert dabei den Anteil der Bandbreite des Übertragungsabschnitts (*link*) und B die Puffergröße im Router, die der Verbindung zugeteilt ist. Die Verzögerung einer Verbindung, die durch einen Token Bucket (r, b) beschrieben werden kann, und dessen Link im Router eine Bandbreite R besitzt, wird durch $\frac{b}{R}$ gegeben, so lange R nicht kleiner ist als r .

Das Netzelement muß sicherstellen, daß die Verzögerung für jedes Datenpaket geringer ist als $\frac{b}{R} + \frac{C}{R} + D$, wobei C und D die maximale Abweichung vom gewählten Modell (*fluid model*) beschreiben. C wird dabei in Bytes gemessen, D in Mikrosekunden. Zum Beispiel ist für Weighted Fair Queueing C von der MTU (*maximum transmission unit*; vgl. Abschnitt 3.2) der abgehenden Verbindung gegeben und D Null.

Die Annahme, daß keine Pakete verloren gehen dürfen, wird dadurch erreicht, indem die Puffergröße des Routers gleich b des Token Bucket plus einiger Fehlerwerte ist.

Exportierte Informationen Der Dienst wird charakterisiert durch die zwei Fehlerwerte, C und D , die anzeigen, inwieweit die Implementation des garantierten Dienstes vom Modell abweicht. Die globalen Parameterwerte C_{tot} und D_{tot} werden durch einfaches Aufsummieren aller lokalen Werte (C, D) berechnet. Diese globalen Werte müssen den Endelementen bekanntgegeben werden. Weiterhin müssen Zwischenwerte von C_{tot} und D_{tot} bei allen Netzelementen, die sich auf dem Datenweg befinden, abgeliefert werden, damit diese (z.B. durch entsprechend große Warteschlangen) sicherstellen können, daß keine Daten in einer Warteschlange verloren gehen.

Resultierender Dienst Der resultierende Dienst von einem Endelement zum anderen ist also ein Dienst mit einer garantierten Verzögerung, bei dem auch garantiert keine Pakete in Warteschlangen verloren gehen. Der Dienst basiert auf einem Modell mit spezifizierten Fehlergrenzen. Die Obergrenze für die Verzögerung von Ende zu Ende ist $\frac{b}{R} + \frac{C_{tot}}{R} + D_{tot}$.

Dieser Dienst ist für Anwendungen gedacht, die eine sichere Garantie dafür brauchen, daß ein Paket nicht später als eine gewisse (vorher bekannte) Zeit, nachdem es vom Sender abgeschickt wurde, beim Empfänger ankommt. Bei der Implementation muß jedoch beachtet werden, daß die Schranke für die maximale Verzögerung sehr hoch sein kann, weil sie unter allen Umständen eingehalten werden muß. Mehrere Studien haben ergeben, daß die tatsächliche Verzögerung für die meisten Pakete sehr viel geringer ist als die garantierte Verzögerung.

<i>Name</i>	<i>bestmöglicher Dienst</i> best effort service	<i>Dienst mit kontrollierter Verzögerung</i> controlled delay service	<i>Vorhersagbarer Dienst</i> predictive service	<i>Garantierter Dienst</i> guaranteed service
Kurzbeschreibung	Bisheriger Internet-Dienst	Ähnlich dem bestmöglichen Dienst mit Kontrolle der Verzögerungen	Paketverzögerung meist unter einer Schranke	Paketverzögerung garantiert unter einer Schranke
Zustellzeit der Pakete	unbekannt	bekannt aus Vergangenheit	meist unter gegebener Schranke	garantiert unter gegebener Schranke
Unterteilung	keine	3 Stufen	3 Stufen	mehrere Stufen
Zugangskontrolle	nein	ja	ja	ja
Bei Überlastung:	verzögern	verzögern	verwerfen	verwerfen
Echtzeitdienst	nein	nein	bedingt	ja

Tabelle 1. Zusammenfassung aller Dienste

3 Änderungen bei den Netzelementen

3.1 Steuerung der Auslastung in den Netzelementen

Netzelemente sind all die Stellen im Netz, bei denen Daten zwischengespeichert oder verarbeitet werden, also beispielsweise Router, Switches oder auch Endrechner. Für Netzelemente, die die im vorherigen Abschnitt besprochenen Dienste anbieten, werden im

folgenden die Rahmenbedingungen vorgestellt. Als wichtigste und elementarste Eigenschaft müssen die Netzelemente in der Lage sein, ihre Auslastung (*load*) zu messen und auch zu steuern (vgl. hierzu auch [Wro95b]). Erst damit sind die im weiteren verlangten Eigenschaften und Verhaltensweisen zu erfüllen.

Verhalten zwischen den Endelementen Aus der Sicht der Anwendung sollte das Verhalten zwischen den Endelementen dem Verhalten zwischen Endelementen bei Verwendung des bestmöglichen Dienstes ohne Last weitestgehend entsprechen. Unter der Annahme, daß das Netzwerk korrekt arbeitet, sind das konkret die folgende Verhaltenspunkte, die nach außen hin sichtbar sein sollten:

- Ein sehr hoher Prozentsatz aller Pakete wird erfolgreich beim Empfänger abgeliefert. Die Verlustrate sollte in etwa der Verlustrate des Transportmediums entsprechen.
- Ein sehr hoher Prozentsatz aller Pakete sollte den Empfänger mit einer Verzögerung erreichen, die nicht deutlich größer ist, als die minimale Verzögerung eines erfolgreich zugestellten Pakets.

Um sicherzustellen, daß diese Bedingungen eingehalten werden, muß jeder Dienst eine Zugangskontrolle bestehen. Bei dieser Zugangskontrolle geben zunächst die Anwendungen, die einen der im letzten Abschnitt besprochenen Dienste in Anspruch nehmen, einen Schätzwert an, was für ein Datenaufkommen sie erzeugen werden. Im Gegenzug versuchen alle Netzelemente zwischen den Endrechnern sicherzustellen, daß dieses Datenaufkommen angemessen bearbeitet wird. Erst wenn alle Netzelemente genügend Ressourcen zur Verfügung stellen können, um diese Dienstanfrage angemessen bearbeiten zu können, wird der Dienstanfrage stattgegeben. Kann dieses Datenaufkommen nach bestandener Zugangskontrolle von einem Netzelement nicht mehr bewältigt werden oder werden kurzzeitig mehr Pakete verschickt, als bei der Zugangskontrolle angegeben wurden, werden je nach gewähltem Dienst die Pakete verzögert oder weggeworfen. Somit kann es auch bei einigen IS-Diensten (wie z.B. dem Dienst mit kontrollierter Verzögerung) weiterhin zu ähnlichen Charakteristiken bei Überlast wie bei dem bestmöglichen Dienst kommen. Durch die Zugangskontrolle wird jedoch die Wahrscheinlichkeit, daß eine solche Stausituation auftritt, stark vermindert.

Datenbehandlung der Netzelemente Für die Prüfung bei der Zugangskontrolle muß das Netzelement seine Ressourcen prüfen. Es muß sicherstellen, daß eine angemessene Bandbreite und genügend Verarbeitungsressourcen vorhanden sind, um den Dienst bedienen zu können. Es genügt jedoch nicht, nur einige dieser Ressourcen bei der Zugangskontrolle zu berücksichtigen, vielmehr müssen alle Ressourcen berücksichtigt werden, die für die angeforderte Operation wichtig sind. Beispiele für solche Ressourcen sind die Bandbreite, der Puffer des Links eines Routers oder Switchs und die Kapazität der Prozessoren zur Paketweiterleitung.

Diese Ressourcendaten werden nicht direkt dazu benutzt, um die Verzögerung oder Verlustrate der Pakete zu berechnen. Aber wenn ein Netzelement einen Dienst annimmt,

impliziert das Netzelement damit, daß es in der Lage ist, einen Dienst derart anzubieten, wie er so (oder so ähnlich) auch von einem bestmöglichen Dienst unter unbelasteten Bedingungen angeboten werden könnte. Bei einer solchen Definition sind folgende Punkte eingeschlossen:

- Die durchschnittliche Verzögerung der Pakete ist nicht oder nur geringfügig größer als die „Burstzeit“ (*burst time*). Die Burstzeit ist definiert als die Zeit, die für die Übertragung der maximalen Datenmenge mit maximaler Datenrate bei der angeforderten Datenbeschreibung benötigt wird.
- Eine sehr geringe Paketverlustrate auf Grund von Überlastungen muß aufrechterhalten werden. Diese Verlustrate kann beispielsweise auf eine Knappheit bei Verarbeitungsressourcen wie Pufferplatz oder Bandbreite zurückzuführen sein. Obwohl hin und wieder ein solcher Verlust auf Grund von Stauungen auftreten kann stellt doch jeder größere Verlust einen Fehler des Zugangskontrollalgorithmus dar.

3.2 Allgemeine Charakterisierungsparameter

Charakterisierungsparameter sind Parameter, die die Netzelemente beschreiben, sie charakterisieren. Weiter oben (unter Abschnitt 2) wurden bereits Charakterisierungsparameter beschrieben, die aber abhängig von dem angebotenen Dienst sind. Darüberhinaus gibt es (nach [Wro95a]) einige Parameter, die nicht in den einzelnen Diensten definiert werden, sondern von allgemeinem Interesse sind. Diese Parameter von allgemeinem Interesse werden deshalb auch allgemeine Charakterisierungsparameter genannt, auf die im folgenden eingegangen wird.

Viele der Parameter sind sowohl als lokale als auch als globale Parameter vorhanden. Als lokale Parameter beschreiben sie das einzelne Netzelement, eventuell noch die Verbindung zum nächsten Netzelement. Als globale Parameter charakterisieren sie die komplette Verbindung zwischen zwei Endrechnern. Diese globalen Parameter werden auch den Anwendungen zur Verfügung gestellt, die dann die Parameter nach ihren Bedürfnissen auswerten können.

Anzahl der IS-Hops IS steht für „Integrated Services“. Ein Netzelement, das die in Abschnitt 2 besprochenen Dienste und Parameter anbietet, heißt dann IS-kennend. Es muß dabei nicht alle Dienste implementieren, jedoch müssen diese Dienste nach außen hin sichtbar sein. Die Anzahl der IS-Hops gibt nun die Anzahl der IS-kennenden Netzelementen an, die zwischen den Endelementen die Verbindung aufrechterhalten. Dabei wird dieser Wert von jedem Netzelement einfach um eins erhöht, bevor der Wert weitergegeben wird.

Anzahl der IP-Hops Dieser lokale Parameter gibt an, wieviele IP-Netzelemente zwischen dem letzten IS-Netzelement und diesem liegen. Diesen Parameter zu bestimmen kann dann problematisch sein, wenn das Netzelement seine Nachbarn nicht kennt. In einem solchen Fall muß dieser Parameter mit dem Setup-Protokoll mitgeliefert werden, beispielsweise indem die TTL-Werte überwacht werden.

Bandbreite Dieser lokale Parameter gibt die Bandbreite des Netzelements an, als globaler Parameter gibt er bei Verbindungen die Bandbreite der Verbindung an. Berechnet wird der globale Wert, indem immer das Minimum aus dem lokalen Wert und dem vorher berechneten globalen Wert gebildet wird. Global informiert dieser Wert also über die minimale Bandbreite, die auf dem Weg vom Sender zum Empfänger auftritt.

Der Wert wird gemessen in Bytes pro Sekunde (*bps*) und kann von einem Byte pro Sekunde bis 40 Terabyte pro Sekunde (was für die maximale theoretische Bandbreite für eine Glasfaserleitung gehalten wird) reichen. Natürlicherweise sind speziell für große Bandbreiten nur die ersten Dezimalstellen von Interesse, weshalb dieser Wert in einer Gleitpunktzahl mit einer Genauigkeit von mindestens 0,1% dargestellt werden kann. Diese Genauigkeit wird von einfach-genauen IEEE Gleitpunktzahlen bereits erreicht.

Sollte die Anzahl der IS-Hops kleiner sein als die Anzahl der IP-Hops, sollten die Endelemente diesen Wert eher als eine Empfehlung als einen verlässlichen Wert betrachten.

Latenzzeit Dieser lokale Parameter gibt die Latenzzeit des Übertragungsabschnitts (*link*) zusammen mit der des Netzelements an, wobei die Latenzzeit definiert ist als die kleinste mögliche Paketverzögerung auf dem Paketweg. Diese Verzögerung kann von optischen Effekten herrühren, von Beschränkungen in der Verarbeitung der Pakete oder von beiden. Um den globalen Wert zu bilden, werden alle Werte zwischen den Endelementen aufsummiert. Dieser globale Wert gibt dann dem Endelement die minimale Verzögerung auf dem Paketweg vom Sender zum Empfänger an.

Die Latenzzeit wird gemessen in Mikrosekunden. Ein einzelnes Element darf dabei Werte zwischen 1 und 2^{28} (etwas über zwei Minuten) annehmen. Die Gesamtlatenzzeit darf Werte bis $2^{32} - 1$ (etwas über eine halbe Stunde) annehmen. $2^{32} - 1$ ist die oberste Grenze und darf deshalb nicht überschritten werden. Die Genauigkeit, mit der das Element diesen Wert messen muß, darf bis zu 100 Mikrosekunden betragen. Jedoch sollte das Element so exakt wie möglich messen und diese geforderte Genauigkeit möglichst unterschreiten.

MTU, maximum transmission unit Als lokaler Parameter gibt die MTU die maximale Übertragungseinheit von einem Netzknoten an, also wie groß die einzelnen Pakete bei einer Übertragung zwischen zwei Netzknoten maximal sein darf. Als globaler Parameter gibt die MTU die minimale größte Übertragungseinheit zwischen den Endrechnern an. Entsprechend wird der globale Parameter aus den Minima der einzelnen Werte berechnet.

4 Heterogenität des alten und der neuen Dienste im Internet

Die Arbeitsgruppe *intserv* geht momentan davon aus, daß sie, während sie neue Dienste definiert und zu neuen Standards erklärt, nicht erwarten kann, daß diese Entwicklungen sofort in die Tat umgesetzt werden. Ein Router, der auf Anfragen nach den erweiterten

Diensten reagiert, unterstützt damit die Architektur der integrierenden Dienste, selbst wenn er einige (oder alle) Dienste nicht implementiert hat. Ein solcher Router wird im folgenden „IS-kennend“ bezeichnet (IS steht für *integrated services*, vgl. auch Abschnitt 3.2). Ob nun ein Dienst oder bestimmte Stufen eines Dienstes implementiert sind, oder nicht, bleibt eine Sache der Netzbetreiber und -entwickler. Einige Router werden jedoch auch in Zukunft keinen anderen Dienst außer dem bisherigen bestmöglichen Dienst anbieten. Während private IP-Netzwerke wohl sehr schnell einen hohen Grad an Homogenität erreichen werden, wird im Internet als Ganzes eine Heterogenität aller angebotenen Dienste unabwendbar sein und muß folglich in der allgemeinen Architektur berücksichtigt werden.

Im folgenden wird über einen der möglichen Ansätze diskutiert, mit dieser Heterogenität fertig zu werden. Dazu wird die Situation angenommen, daß verschiedene definierte Dienste, die eine Zugangskontrolle verlangen, zusammen mit dem bestmöglichen Dienst angeboten werden. Die Diskussion wird beschränkt auf das Problem der Heterogenität in einer Menge aus IS-kennenden Routern. Speziell wird das Problem von IS-nicht-kennenden Routern nicht behandelt, die nur den bestmöglichen Dienst kennen und auf Anfragen anderer Dienste nicht reagieren können.

Gäbe es keine Mechanismen um mit der Heterogenität in der Menge der angebotenen Dienste fertig zu werden, müßten Router, die einen von einer Anwendung verlangten Dienst nicht unterstützen, diese Dienstanfrage ablehnen. Dann könnten die Anwendungen nur diejenige Dienste nutzen, die von allen Routern auf dem Weg zwischen Sender und Empfänger angeboten werden. Diese Einschränkung auf den „besten allgemein bekannten Dienst“ beschränkt die den Anwendungen zur Verfügung stehenden Dienste merklich. Zum Beispiel würden Anwendungen, die auf einem an einem Ethernet angeschlossenen Endsystem laufen, das nur den bestmöglichen Dienst anbieten kann, keinen anderen als den bestmöglichen Dienst bei allen Routern auf dem ganzen Weg sicherstellen können. Dieses Verhalten würde deshalb die Entwicklung des Internet zu einer integrierenden Dienstarchitektur drastisch bremsen, wenn nicht sogar verhindern. Deshalb sind Mechanismen, die auf das Problem der Heterogenität abzielen, notwendige Komponenten für eine integrierende Dienstarchitektur.

Einige Fälle der Heterogenität können als trivial abgetan werden. Wenn Router eine Dienstanfrage erhalten, können sie immer frei zu einer „besseren“ Stufe wechseln. Weniger trivial ist schon, daß, unter der Voraussetzung daß der vorhersagbare Dienst besser ist als der Dienst mit kontrollierter Verzögerung, dann der Router immer den Dienst mit kontrollierter Verzögerung durch den vorhersagbaren Dienst ersetzen kann. In solchen „geordneten“ Fällen kann darüber nachgedacht werden, ob das Netzelement sowohl den schlechteren als auch den besseren Dienst überhaupt anbieten soll.

4.1 Ersatzdienste

Der von der Arbeitsgruppe *intserv* ausgearbeitete Lösungsweg basiert auf dem Prinzip der „Ersatzdienste“ (*replacement services*). Wenn ein Dienst angefordert wird, kann er, falls der Router diesen Dienst nicht anbietet, durch einen anderen Dienst ersetzt werden (unter den unten stehenden Bedingungen). Die fundamentalen Prinzipien, die diesem

Lösungsweg zugrundeliegen, sind, daß (1) der Dienst, den die Anwendung erwartet, modelliert ist auf der idealen Ende-zu-Ende-Verbindung des angeforderten Dienstes (das heißt dem Dienst, der zur Verfügung gestellt werden würde, wenn alle Router auf dem Weg den angeforderten Dienst anbieten würden) und (2) das Netzwerk dafür verantwortlich ist, die Anwendung darüber zu informieren, in wie weit dies eine realistische Annäherung ist. Dadurch erhalten Anwendungen keinen Einblick in die genaue Natur der Heterogenität, also wie sich der angeforderte Dienst auf dem Weg aus den verschiedenen Diensten zusammensetzt.

Die Verwendung von Ersatzdiensten sind aus zwei Gründen realisierbar. Erstens können lokale Bedingungen und Dienstdefinitionen einem Router ermöglichen, einen Dienst durch einen anderen ohne einen wahrnehmbaren Unterschied für die Anwendung in den Diensteseigenschaften zu ersetzen. Zum Beispiel mag ein sehr schwach belastetes Ethernet, das nur den bestmöglichen Dienst anbieten kann, durchaus so geringe Verzögerungen besitzen, wie sie auch von dem garantierten Dienst verlangt werden. Zweitens variieren die Anwendungen im Grad der Zusicherung, die sie für den Dienst benötigen. Eine adaptive Anwendung, die den Verlust von Paketen toleriert, mag sich auch mit dem bestmöglichen Dienst in den Routern zufrieden geben, die einen Echtzeitdienst nicht anbietet.

Der Ersatzdienst wird unterteilt in einen „zuverlässigen“ und einen „unzuverlässigen“ Ersatzdienst. Ein zuverlässiger Ersatz trifft die meiste Zeit die Dienstanforderungen des zu ersetzenden Dienstes. Die Arbeitsgruppe *intserv* will für den Grad der Zuverlässigkeit keinen genauen Wert angeben, aber um eine Diskussionsgrundlage zu schaffen, sollte ein zuverlässiger Ersatz diese Bedingung zu mindestens 95% erfüllen. Obwohl nicht genau spezifiziert ist, wie eine zuverlässige Ersetzung zu arbeiten hat, wird angenommen, daß sich im Laufe der Zeit gewisse Richtlinien und informelle Standards herausbilden werden. Der springende Punkt hinter dem Begriff einer zuverlässigen Ersetzung ist die implizite Annahme, daß sich die Verwendung eines Ersatzdienstes gewöhnlich nicht merklich von dem angeforderten Dienst unterscheidet.

Unzuverlässige Ersetzungen sind die schwächere Form von Ersatzdiensten, welche keine Zusicherung über die Qualität des resultierenden Dienstes geben. Deshalb können unzuverlässige Ersatzdienste, beliebig schlecht sein. Beispielsweise können der Dienst mit kontrollierter Verzögerung oder der bestmögliche Dienst in einem häufig überlasteten Netzwerk als unzuverlässige Ersetzungen für den garantierten Dienst angesehen werden. Einige Anwendungen mögen bereit sein, unzuverlässige Ersatzdienste zu verwenden wenn keine andere Möglichkeit existiert, denn in einigen Fällen ist ein schlechter Dienst immer noch besser als gar kein Dienst. Unzuverlässige Ersatzdienste werden immer angeboten, speziell da der bestmögliche Dienst eine unzuverlässige Ersetzung für alle anderen Dienste darstellt.

Die Wahl, welcher Dienst als Ersatzdienst für eine Anfrage verwendet werden soll und ob diese Ersetzung als zuverlässig oder unzuverlässig eingestuft wird, bleibt dem Router überlassen. Damit können Dienste mit hohen Anforderungen durch schwächere Dienste ersetzt werden, solange keine Überlast auftritt. In einigen Fällen wird die Anwendung genau die Unterschiede der erreichten Qualität zwischen dem garantierten oder dem vorhersagbaren Dienst und seiner zuverlässigen Ersetzungen erkennen und die Ersetzungen

dann nicht verwenden, aber in den meisten Fällen werden diese Unterschiede keine Rolle spielen.

4.2 Routersubstitutionen

In diesem Abschnitt wird beschrieben, wie Router Ersatz für Dienste, die nicht angeboten werden, substituieren. Die Behandlung des Punktes, wann die Benutzung eines Ersatzes erlaubt ist, und wie viel Kontrolle die Anwendung über diesen Prozeß haben darf, wird auf später verschoben (vgl. Abschnitt 4.3). Für den Augenblick wird angenommen, daß wann immer ein Router eine Anfrage für einen nicht angebotenen Dienst erhält, er versucht, einen Ersatzdienst zu verwenden.

Anfragen für nicht angebotene Dienste können in zwei Kategorien unterteilt werden: zum einen in Dienste, die zwar bekannt sind aber vom Router nicht angeboten werden, zum anderen in Dienste, die weder bekannt noch angeboten werden. Dies bedeutet, daß eine zuverlässige Ersetzung substituiert wird, wenn eine angeboten wird, und eine unzuverlässige Ersetzung substituiert wird, wenn keine zuverlässige Ersetzung angeboten wird. Die Dienstanfrage wird dann behandelt, als ob sie eine Anfrage für den Ersatzdienst wäre. Wenn der Ersatzdienst eine Zugangskontrolle verlangt, verwendet der Router den Zugangskontrollalgorithmus, um den Zugang für den neuen Dienst zu prüfen. Der Dienst wird abgewiesen, wenn der Zugangskontrolltest scheitert; weitere Substitutionen werden nicht zugelassen. (Der Grund für diese Beschränkung wird später klar. Kurz erklärt erlaubt diese Beschränkung den Routern, die Ersatzdienste benutzen, auch Charakterisierungsparameter über diese Ersatzdienste zu exportieren.) Wenn der Dienst zugelassen wird oder der Ersatzdienst keine Zugangskontrolle verlangt, ist die Dienstanfrage angenommen. Wenn ein Reservierungsprotokoll wie RSVP verwendet wird, um die Dienstanfrage zum Router zu befördern, wird die Dienstanfrage von dem Reservierungsprotokoll entlang des Datenpfads zurückbefördert.

Wenn ein Router eine Anfrage für einen unbekannten Dienst bekommt kann er nicht einmal die Brauchbarkeit der Ersatzdienste beurteilen. Das beste, was er tun kann, ist den bestmöglichen Dienst anzubieten.

4.3 Verwendung von Ersatzdiensten

Wenn mehrere Dienstanfragen an einem Router zusammenkommen und von dort alle in die selbe Richtung weitergeleitet werden, können diese Dienste (wenn sie die Zugangskontrolle bestanden haben) zu einem einzigen gemeinsamen Dienst zusammengesetzt werden (*merging*). An geeigneten anderen Routern muß dieser eine gemeinsame Dienst dann wieder in die einzelnen Dienste aufgesplittet werden.

Während die Verwendung von Ersatzdiensten die Zusammensetzung von heterogenen Diensten erlaubt, könnten Anwendungen verschiedenen Toleranzstufen für die Verwendung von Ersatzdiensten haben. Einige tolerante Anwendungen mögen immer geneigt sein, Ersatzdienste zu verwenden (egal, ob zuverlässig oder unzuverlässig), während andere Anwendungen es vorziehen würden, daß ihre Dienstanfrage abgelehnt wird, wenn der

angeforderte Dienst nicht angeboten wird. Mechanismen, um den verschiedenen Bedürfnissen von Anwendungen zu begegnen, werden später diskutiert. Es überrascht nicht, daß das Ausstatten der Anwendungen mit zusätzlicher Funktionalität und Flexibilität auch zusätzliche Komplexität in den dazugehörigen Mechanismen mit sich führt.

Der einfachste Entwurf entzieht den Anwendungen alle Wahlfreiheit und erlaubt nur eine einheitliche Verwendung von Ersatzdiensten. Dabei wird der Router immer versuchen, den besten Ersatzdienst zu verwenden, wenn der angeforderte Dienst nicht angeboten wird. Während dieser Entwurf einfach zu verwirklichen ist in dem Sinne, daß es keine zusätzliche Komplexität bei den Dienstanfragen oder der Zusammensetzung verlangt, geht er nicht auf Anwendungen ein, die keine Ersatzdienste verwenden wollen. Aber, wie weiter oben schon erwähnt, in Umgebungen, in denen eine Ende-zu-Ende-Charakterisierung der Dienstverfügbarkeit den Anwendungen angeboten wird, wissen die Anwendungen bereits bevor sie eine Dienstanfrage stellen, ob Ersetzungen benutzt werden. In einer solchen Umgebung hat eine Anwendung, die keine Ersatzdienste verwenden will, die Möglichkeit, einen Dienst, der nicht durchgehend von Ende zu Ende angeboten wird, nicht anzufordern. Wenn also die Ende-zu-Ende-Charakterisierung immer verfügbar wäre dann wäre dieser Entwurf einfach und wünschenswert. Das Design muß aber auch die Situationen berücksichtigen, in denen diese Charakterisierung nicht verfügbar ist.

Eine Erweiterung von diesem einfachen Entwurf ermöglicht Anwendungen, das Ausmaß, in dem Ersetzungen benutzt werden, zu steuern. In diesem Entwurf wird ein Bit zur Behandlung der Ersetzung bei jeder Dienstanfrage eingeführt. Wenn dieses Bit bei einer Dienstanfrage gesetzt ist, darf ein Ersatzdienst verwendet werden. Wenn das Bit nicht gesetzt ist muß ein Router, der den angeforderten Dienst nicht anbietet, die Dienstanfrage ablehnen und eine Fehlermeldung zurückgeben. Dies ermöglicht Anwendungen, ihre Vorlieben für die Verwendung von Ersatzdiensten auszudrücken. In Umgebungen, in denen die Ende-zu-Ende-Charakterisierung für die Verfügbarkeit der Dienste bekannt ist, bedeutet das gesetzte Bit, daß Anwendungen bereit sind, jegliche Ersatzdienste entlang des Pfades zu nutzen. In Umgebungen, in denen diese Charakterisierung vorhanden ist, sollten Anwendungen immer dieses Bit setzen, da Anwendungen, die keine Ersetzungen wünschen, bereits vor der Anfrage von den Ersetzungen wissen und deshalb keine Anfrage stellen sollten.

Mit der Einführung dieses zusätzlichen Bits bei Dienstanfragen und wenn man Anwendungen erlaubt, dieses Bit zu kontrollieren, taucht das Problem auf, wie Dienste zusammengefügt (*merging*) werden sollen, die dieses Bit verschieden gesetzt haben. Wie soll vorgegangen werden, wenn bei einer Dienstanfrage das Bit zur Behandlung der Ersetzung gesetzt ist und bei der anderen nicht?

In [BS95] sind zwei Möglichkeiten ausgearbeitet, wie dieses Zusammenfügungsproblem gelöst werden kann. Bei der ersten Möglichkeit wird ein spezielles Vereinigungsverhalten vorgeschrieben. Genauer, wenn zwei Dienste zusammengefügt werden, die das Bit verschieden gesetzt haben, wird der vereinigte Dienst das Bit nicht gesetzt haben. Dieser Versuch, das Problem zu lösen, geht den konservativen Weg, nämlich daß es besser ist, einer „flexiblen“ Anwendung die Verwendung von Ersatzdiensten zu verbieten, als eine

„unflexible“ Anwendung mit einem niedrigeren Dienst als dem angeforderten zu versehen. Wenn die Anwendungen sich korrekt verhalten, sollte die unflexible Anwendung ihre Dienstanfrage einstellen oder ändern, was den anderen Anwendungen die Möglichkeit gibt, Ersatzdienste zu benutzen (obwohl ihr Dienst unterbrochen werden könnte). Durch diesen Mittelweg können Anwendungen ihre Wünsche ausdrücken, aber schwierige Zusammenfügungsalgorithmen werden umgangen. In diesem Fall kann das gesetzte Bit zur Behandlung der Ersetzung nur als Empfehlung angesehen werden. Anwendungen, die dieses Bit nicht setzten, oder die noch nicht ausgereift sind, können also den Anwendungen, die bereit sind, Ersatzdienste zu verwenden, Dienste wegnehmen. Bei diesem Lösungsweg sollten dann, wenn eine Ende-zu-Ende-Charakterisierung für alle Endelemente verfügbar ist, alle Anwendungen ihr Bit setzen und es würde kein kompliziertes Zusammenfügevverhalten mehr geben. So lange es diese Charakterisierung von Ende zu Ende nicht gibt, müssen einige flexible Anwendungen ihren Dienst für unflexible Anwendungen unterbrechen.

Behandelt man das Bit zur Behandlung der Ersetzung höher als eine Empfehlung, muß mehr Information verwaltet werden, wenn Diensteanfragen zusammengefügt werden sollen. Genauer verlangt die Zusammenfügung von Dienstanfragen, die das Bit unterschiedlich gesetzt haben auch die Verwaltung von zwei Dienst Anforderungsspezifikationen (*service request specification*, *RSpec*) in der zusammengefügtten Anfrage. Eine Spezifikation beschreibt den besten Dienst, der von einer Anwendung angefordert wurde, die bereit ist, Ersatzdienste zu nutzen, die andere Spezifikation steht für den besten Dienst, der von einer Anwendung angefordert wurde, die für die Ersatzdienste nicht bereit ist.

Die Verarbeitung dieser zusammengefügtten Anfrage ist also komplizierter. Der Router muß eine angemessene Stufe eines Dienstes wählen, entscheiden, wann er einen Fehler zurückzugeben hat und wann und wie die Dienstanfrage geändert werden muß. Trotzdem hat dieser Entwurf den Vorteil, den Anwendungen eine größere Funktionalität zur Verfügung zu stellen.

5 Ergebnis

Die Arbeitsgruppe *intserv* hat ihre Arbeit beinahe abgeschlossen. Einige Details sind noch offen und müssen noch geklärt werden, doch im großen und ganzen ergibt sich ein recht rundes Bild bei der Verwendung und Integration der neuen Dienste. Interessanterweise wurden bei der Ausarbeitung der neuen Dienste sehr viele Kompromisse eingegangen, da man zum einen die Entwickler nicht zu sehr einschränken wollte, zum anderen aber auch möglichst gute Dienste anbieten wollte. Trotz dieser vielen Kompromisse scheinen diese Dienste nicht in ihrer Leistungsfähigkeit beschränkt. Auf jeden Fall ist es besser, diese Kompromisse bereits beim Entwurf der Dienste einzugehen, als später bei der Implementation durch Nachbesserungen die Dienste zu beschränken.

Obwohl die Brauchbarkeit der neuen Dienste durch zahlreiche Simulationen unterstrichen wurde, bleibt nun die Praxistauglichkeit abzuwarten. Zunächst müssen diese Dienste in Soft- und Hardware implementiert werden und danach zuerst in kleinen Netzen unter Alltagsbedinungen getestet und verifiziert werden. Bis diese Dienste die breite Masse

erreichen, werden wohl noch einige Jahre vergehen, jedoch lassen die hier besprochenen Theorien auf bessere Zeiten im Internet hoffen.

OSI95: Der Transportdienst der 90er Jahre

Alexander Rösner

Kurzfassung

In den letzten Jahren haben sich die Anforderungen an die Kommunikationsdienste sowohl im lokalen als auch im Weitverkehrsbereich derart tiefgreifend verändert, daß sie mit den traditionellen OSI-Transportdiensten und Protokollen weder beschrieben, geschweige den erfüllt werden können. Es besteht also die Notwendigkeit einer Anpassung der OSI-Protokolle und Dienste an die heutigen Gegebenheiten [Dan94]. „High Performance OSI Protocols with Multimedia Support on HSLANs and B-ISDN“, kurz OSI95, ist ein 1990 entstandenes ESPRIT-Projekt, welches versucht, Lösungen zu dieser Problematik aufzuzeigen. Ich werde im folgenden Auszüge der Ergebnisse der ersten Phase dieses Projekts vorstellen. Dabei beleuchte ich zunächst die Entwicklungen, welche das Projekt beeinflussten, und präsentiere dann im zweiten Teil die OSI95-Erweiterungen in Bezug auf die Qualität von Kommunikationsdiensten. Schließlich gehe ich auf den neuen verbindungslosen bzw. verbindungsorientierten Transportdienst ein.

1 Einleitung

Zu der Zeit als die Transportprotokolle entworfen wurden, die auch heute noch verwendet werden und weit verbreitet sind, war das größte Problem die bei der Übertragung auftretende hohe Fehlerquote. Nahezu alle Anwendungen brauchten eine sehr hohe Übertragungszuverlässigkeit. So ist es verständlich, daß die Protokolle des OSI-Turms, wie TP4, aber auch TCP, sich in erster Linie dem Problem einer zuverlässigen Übertragung und der Beseitigung der bei der Übertragung auftretenden Fehler widmen. Einzig die gleichzeitig übertragbare Datenmenge und die Verzögerung der Datenpakete machen hier die Qualität eines Netzsegmentes aus. Dabei wird in Kauf genommen, daß eben diese Qualitäts- und Leistungsmerkmale mit wachsender Nutzung des Netzes immer schlechter werden, da alle Netzwerkbenutzer anarchisch von den Ressourcen Gebrauch machen, und so keinerlei Qualitätsgarantien gegeben werden können. Heutige interaktive Anwendungen sind fehlertoleranter im Bezug auf Übertragungszuverlässigkeit, allerdings weitaus sensibler gegenüber Durchsatz und Verzögerung. Damit wird mit den neuen Technologien in LANs und in Weitverkerksnetzen ein (aus jetziger Sicht) übertriebener Fehlerkorrekturoverhead überflüssig und die unkontrollierte Netzbenutzung untragbar. Auch sind die heutigen Transport-Protokoll-Implementierungen nicht in der Lage, die theoretisch möglichen Übertragungsleistungen von schnellen Netzen zu erbringen. Dies ist auch nicht verwunderlich, denn die Eröffnung einer Autobahn erhöht ja auch noch lange nicht die Geschwindigkeit von Fahrrädern.

1.1 Änderungen der Leistungsanforderungen

In der Zwischenzeit gibt es Dienste im LAN-Bereich, die sich keiner der OSI-Schichten zuordnen lassen, wie etwa der *multicast service*, der weder eindeutig zur Netzwerk-, noch

zur Transportschicht gehört. Hinzu kommt, daß Multimedia-Anwendungen garantierte, hohe Bandbreiten und geringe Unterschiede in den Laufzeiten der einzelnen Datenpakete (*jitter*) benötigen. Es ist meist gar nicht nötig, daß alle Bitfehler, die bei einer Übertragung auftreten, korrigiert werden. Ein Beispiel dafür ist etwa ein Datenendbenutzer eines Multimediasdienstes, der es nicht bemerkt, wenn beim Übertragen eines Videobildes ein Bit in dem für einen Bruchteil einer Sekunde sichtbaren Einzelbild einer Videokonferenz verkehrt angezeigt wird, der aber sehr wohl bemerkt, wenn sein Videobild ruckelt. Diese Fehlertoleranz impliziert eine ganz andere Qualitätsanforderung an die Netzdienste, als sie „herkömmliche“ Anwendungen haben.

Auch Klient/Server-Anwendungen benötigen kurze Laufzeiten und ein erweitertes Adressierungsverfahren. So findet die Serversuche oft mit *broadcast* oder *multicast*-Mechanismen statt, und die Tatsache, daß viele Klienten während der Erbringung eines Dienstes durch den Server blockieren, schreit geradezu nach kurzen Laufzeiten. Diese kurzen Lauf- oder Antwortzeiten lassen sich mit dem OSI-Turm aber kaum erbringen.

Aber damit nicht genug: Interaktive Grafik-Anwendungen erzeugen keinen kontinuierlichen Datenstrom, sondern sogenannte *bursts*, in welchen nur für einen kurzen Zeitraum eine große Bandbreite benötigt wird.¹ Danach werden wieder geringe Datenmengen übertragen.

Die neuen Anforderungen sind eigentlich zu unterschiedlich, als daß sie von den derzeit eingesetzten Protokollen (selbst nach einigen Änderungen derselben) erbracht werden können. Es ist also nötig die Transport-Problematisierung unter den Gesichtspunkten der neuen Anforderungen von Grund auf neu zu überdenken. Der Entwurf eines derart neuen Transportdienstes ist das Hauptziel des ESPRIT-OSI95-Projektes.

1.2 Lösungsmöglichkeiten

Es gibt zwei Ansätze zur Lösung der oben beschriebenen Probleme. Der erste fordert einen komplett neuen Entwurf der Transportprotokolle. Der zweite sieht einen Teil der Probleme in den schlechten Protokollimplementierungen und möchte die bestehenden Protokolle nur modifizieren, um die tatsächlich möglichen Übertragungsleistungen besser auszunutzen. Folgende Schwächen bestehender Protokolle und Implementierungen fallen auf:

- *timeout*: Die Benutzung von *timeouts* ist in hohem Maße ineffektiv. Speziell in Hochgeschwindigkeitsnetzen ist *timeout*-basierte Staukontrolle unbrauchbar.
- *Speicherung*: Es gibt Implementierungsprobleme, die unabhängig von den verwendeten Protokollen sind. So hat etwa die Größe der verwendeten Puffer in einem Netzknoten einen starken Einfluß auf das Laufzeitverhalten².
- *Parallelisierung*: Einige Probleme im Zusammenhang mit der *Leistung* lassen sich durchaus parallelisieren, es wäre also hilfreich, wenn die Protokollspezifikationen solche Ansätze unterstützen oder zumindest nicht prinzipiell unterbinden würden.

¹ z.B., wenn ein Benutzer mit der Maus wackelt

² Mit Hilfe von Warteschlangen läßt sich dies beschreiben.

Während früher die Fehlerrate das größte Problem in der Datenübertragung war, sind es heute die Staukontrolle und die Garantie der Dienstqualität. Wenn in einem unter Last stehendem Netzsegment Datenpakete verloren gehen, werden die Daten bei TCP und TP4 nach einem *timeout* abermals versendet, was die Last nur noch weiter erhöht. Es muß abgewägt werden, ob sich der große Aufwand einer Stauvermeidung, verglichen mit dem weitaus geringeren Aufwand einer Stauerkennung beziehungsweise Auflösung, lohnt. Dabei sollte grundsätzlich zwischen der Systemsicht des Netzbetreibers und der des Netzbenutzers unterschieden werden. Entweder schützen die stauvermeidenden Maßnahmen den Netzbetreiber vor einem vollständigen Kollaps des Netzes, oder sie bieten dem Benutzer den bestmöglichen Dienst. Dieser Zwiespalt zwischen den Wünschen des Netzbetreibers und denen des Benutzers führt zur Notwendigkeit einer genauen Spezifikation des Dienstes, einer Art Verkehrsdienstvertrag zwischen Betreiber und Nutzer, in dem die einzelnen Qualitätsparameter des Dienstes (QoS³) genau festgelegt werden. OSI95 bietet eine Semantik zur Formulierung der Dienstqualitätswünsche eines Benutzers unter Berücksichtigung der Möglichkeiten des Betreibers. Welche Dienstqualitäten notwendig und sinnvoll sind, und wie sie vereinbart werden, beschreibt der nächste Abschnitt.

2 Die Qualität von Transportdiensten

Bevor ich näher auf die Dienstqualität eingehe, zunächst eine Definition: *Die Qualität eines Dienstes ist ein Maß für den Grad an Zufriedenheit der Dienstbenutzer.* In OSI95 ist also nur die Sicht des Benutzers relevant für die Bestimmung der Dienstqualität. Für den Dienstbenutzer errechnet sich die Dienstqualität also anhand des tatsächlichen Wertes einiger Ende-zu-Ende Parameter. Die wichtigsten Leistungsmerkmale sind dabei der Durchsatz, die Verzögerung und die Zuverlässigkeit der Übertragung. Um diese Leistungsmerkmale messen zu können, müssen sie erst einmal genau definiert werden.

2.1 Die Sicht des Benutzers

Der Benutzer eines Transportdienstes hat eine sehr eingeschränkte Sicht auf das Netz, er hat eigentlich nur die Möglichkeit, die Dienstprimitive, die an seinem Dienstzugangspunkt (TSAP, *transport service accesspoint*) auftreten, zu beobachten.

Neben dem Auftreten der Dienstprimitive am TSAP kann der Benutzer nur noch die Zeit zwischen zwei Primitiven messen. Die Zeitmessung spielt eine wichtige Rolle bei der Definition der Leistung (*performance*). Jeder Qualitätsparameter wird in Abhängigkeit vom Auftreten zweier oder mehrerer in Beziehung stehender Dienstprimitiven definiert. Welcher Art diese Beziehung ist, hängt von dem gewünschten Dienst ab. Bei einem verbindungsorientierten Dienst wird die Dienstqualität beim Aufbau der Verbindung ausgehandelt. Im Falle eines verbindungslosen Dienstes werden sie vom Dienstanutzer bestimmt und an eine einzelne SDU⁴ gebunden. Nach [CCITT I.350] müssen benutzerorientierte Dienstqualitäten Eigenschaften besitzen, die [ADL94]

³ Quality of Service

⁴ Service data unit

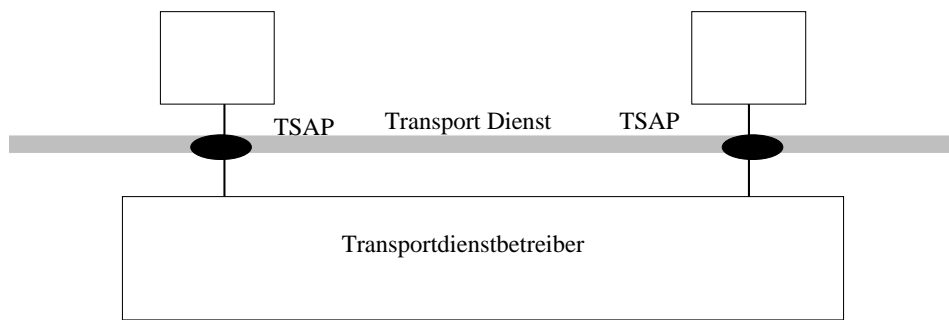


Abbildung 1. Ein Modell eines Transportdienstes.

- sich auf vom Benutzer wahrnehmbare Effekte konzentrieren und nicht auf die Ursachen dieser Effekte im Netzwerk,
- unabhängig vom Aufbau des Netzwerkes sind,
- aus Benutzersicht alle Aspekte des Dienstes, die objektiv am SAP gemessen werden können, überdecken,
- dem Benutzer am SAP vom Netzbetreiber garantiert werden können,
- in einer netzwerkunabhängigen Syntax formuliert werden.

Damit ist klar, daß die Dienstqualität nur auf vom Benutzer am SAP wahrnehmbaren Effekten aufbauen.

2.2 Was sind QoS Parameter ?

Ein Dienstqualitätsparameter muß die Anforderungswünsche des Benutzers ausdrücken können, und er sollte überwachbar⁵ sein. Ich gebe im folgenden Definitionen einiger der wichtigsten QoS-Parameter an, wie sie in OSI95 vorgeschlagen werden.

Verzögerung (*transit delay*.)

Die Verzögerung ist das Zeitintervall zwischen dem Auftreten des DATA.request Primitive am SAP des sendenden Benutzers und Eintreffen des entsprechendem DATA.indication Primitive am SAP des Empfängers. Diese Definition ist gültig sowohl für verbindungslose als auch für verbindungsorientierte Dienste.

Verzögerungsschwankung (*transit delay jitter*)

Die Verzögerungsschwankung ist die Differenz zwischen der längsten und kürzesten Verzögerung, die in einer Verbindungsrichtung beobachtet wird⁶.

Durchsatz (*throughput*)

Der Durchsatz einer Verbindung ist abhängig von der Länge⁷ der einzelnen Datenpakete (SDUs) und von der Übertragungsrichtung. Der **Sendedurchsatz**(*sending user's*

⁵ am besten durch den Betreiber

⁶ Das wäre eigentlich der *max jitter*, aber so sehr geht man in [2] wohl nicht ins Detail

⁷ Der Anzahl der Oktets

throughput) ist die Anzahl der Oktets, der zuletzt übertragenen SDU dividiert durch das Zeitintervall zwischen dem letzten und dem nächsten T-DATA.request.

Der **Empfangsdurchsatz**(receiving user's throughput) ist die Anzahl der Oktets der zuletzt übertragenen SDU dividiert durch das Zeitintervall zwischen dem korrespondierenden letzten und dem nächsten T-DATA.indication.

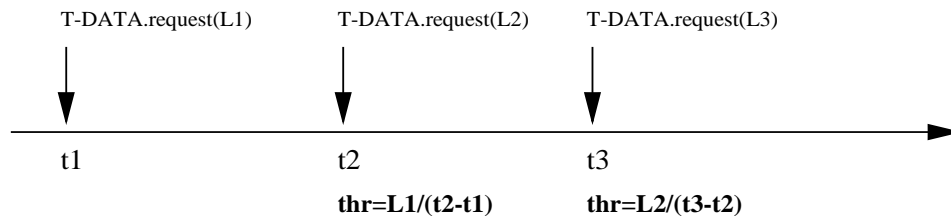


Abbildung 2. OSI95 Definition des Sendedurchsatzes

Der *Gesamtdurchsatz* ist der kleinere von beiden. Diese OSI95-Definition ist etwas unterschiedlich von der ISO-Definition und erleichtert die Durchsatzüberwachung durch den Betreiber, da sie den Momentandurchsatz definiert.

3 Formen der Aushandlung

Nun, da die einzelnen Parameter der Dienstqualität definiert sind, müssen wir uns ein paar Gedanken darüber machen, wie sie ausgehandelt werden, d.h. wie ein Verkehrs-dienstvertrag zwischen Nutzern und Betreibern festgelegt wird. Im Falle einer Punkt-zu-Punkt-Verbindung sind die drei Verhandlungspartner der sendende Benutzer (Sender), der Netzwerk-Betreiber (Betreiber) und der empfangende Benutzer (Empfänger). Alle Verhandlungsformen basieren auf dem klassischen Austausch der vier Primitive *request*, *indication*, *response* und *confirmation*. Für einige Leistungsparameter⁸ sind hohe Werte gewünscht; für andere⁹ niedrige. Ich spreche im folgenden von *verstärken* und *abschwächen* von Leistungsparametern, je nach dem ob sie qualitativ verbessert oder verschlechtert werden sollen.

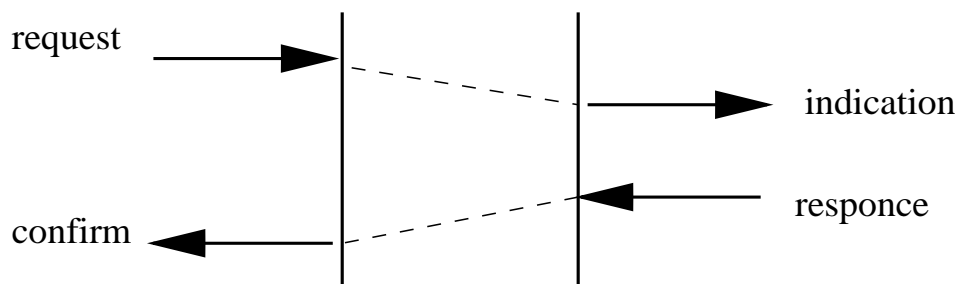


Abbildung 3. Klassischer 4-Primitiven-Austausch

⁸ etwa Durchsatz

⁹ jitter

In einer ersten Form der Verhandlung, die hier diskutiert wird, schlägt der Sender dem Betreiber den Wert eines Qualitätsparameter mit einem Request-Primitiv vor (Abb.4a). Der Betreiber darf diesen Wert nach Belieben abschwächen (Abb.4b) und leitet ihn an den Empfänger mit dem Indication-Primitiv weiter (Abb.4c). Dieser darf den Wert des Parameters wiederum weiter abschwächen, bevor er ihn dem Response-Primitiv übergibt (Abb.4d). Am Ende der Verhandlung steht dann der beste Wert, den alle drei Verhand-

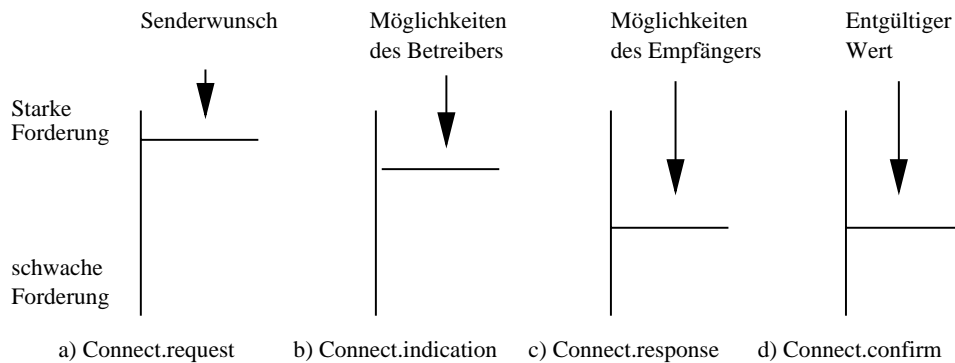


Abbildung 4. Dreiecksverhandlung zum Informationsaustausch

lungspartner erfüllen können. Der Dienst wird nur dann vom Betreiber abgelehnt, wenn dieser nicht in der Lage ist, diesen Dienst überhaupt zu erbringen, und der Sender hat jederzeit die Möglichkeit einen Verbindungsabbruch zu veranlassen, wenn er mit dem Ergebnis der Verhandlung nicht zufrieden ist. Der ISO-Transportdienst nutzt diese Form der *Dreiecksverhandlung zum Informationsaustausch* (Sender-Betreiber, Betreiber-Empfänger, Empfänger-Sender) für seine leistungsorientierten Qualitätsparameter.

Mit dem eben beschriebenen Verhandlungsschema kann der Sender seine Qualitätsanforderungen nicht genau genug beschreiben. Daher gibt in einer erweiterten Variante, der *Dreiecksverhandlung eines Vertragswertes*, der Sender dem Request-Primitiv zwei Werte des Qualitätsparameters mit, einen Maximalwert und einen Minimalwert (Abb.5). Wenn der Betreiber den Request annimmt, darf er den Wert des Qualitätsparameters nicht weiter als bis zum Minimalwert abschwächen, bevor er beide Werte an den Empfänger mittels connect.indication weiterreicht. Der Empfänger wählt nun einen Wert zwischen Minimalwert und dem neuen Maximalwert aus und teilt diesen dem Betreiber mittels connect.response mit. Dieser informiert den Sender mit einem connect.confirmation über den endgültigen Wert des angeforderten Qualitätsparameters.

3.1 Bestmögliche Dienstqualität

Bei dieser Form der Dienstqualität wird der vereinbarte Qualitätsparameter vom Betreiber nicht überwacht. Der Dienstbetreiber ist also auch nicht in der Lage die Benutzer darüber zu informieren, wenn der ausgehandelte Qualitätsparameter unter den vereinbarten Wert fällt. Dies tritt etwa ein, wenn der Betreiber selbst unfähig ist, die Dienstqualität weiterhin zu erbringen. Dabei ist es vollkommen irrelevant, mit welchem der beiden oben beschriebenen Verfahren der Qualitätsparameter ausgehandelt wurde. Es ist klar, daß diese Form der Qualitätshandhabung keinerlei Gewähr für den tatsächlichen Wert des

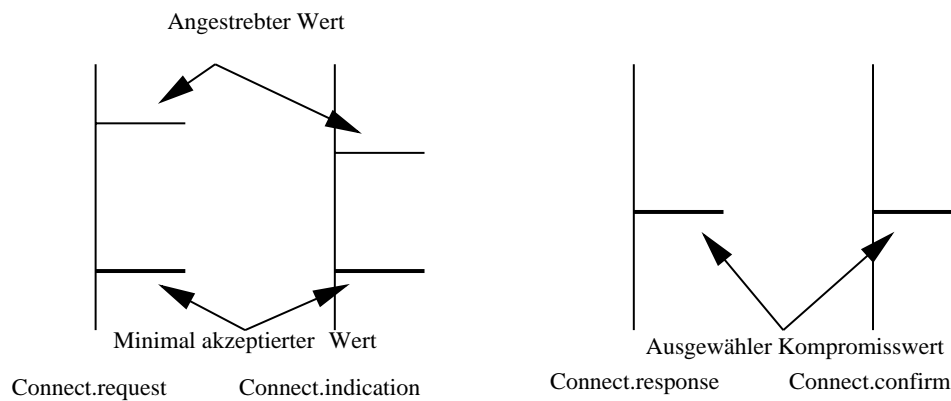


Abbildung 5. Dreiecks Aushandlung eines Vertragswertes

ausgehandelten Qualitätsparameters gibt. Wenn der Benutzer also „Glück“ hat, dann wird seine Forderung erfüllt, wenn er „Pech“ hat, dann eben nicht. Ist letzteres der Fall, dann wird er noch nicht einmal darüber informiert

Dieser für heutige Anforderungen unhaltbare Zustand ist leider die Realität, wie sie uns täglich beim Transportdienst und den Protokollen von ISO begegnet. Der einzige Qualitätsparameter, der in Protokollen wie etwa TP4 unterstützt wird und garantiert werden kann ist die Fehlerrate, was bei weitem nicht ausreichend ist.

3.2 Garantierte Dienstqualität

Was wir benötigen, ist ein Transportdienst mit garantierter Dienstqualität, bei dem die Qualitätssicherung allein Aufgabe des Betreibers ist. Grundvoraussetzung dafür sind ausreichende, alokierbare und managbare Ressourcen auf Seiten des Betreibers und auf Seiten des Benutzers ausreichende Rechenleistung, um den vom Betreiber angebotenen Dienst richtig entgegennehmen zu können. Dabei treten folgende Schwierigkeiten auf:

- Die Reservierung von Ressourcen muß auf dem ganzen Pfad vom Sender bis zum Empfänger vorgenommen werden.
- Eine Qualitätsvereinbarung sollte rekonfigurierbar sein.
- Es gibt immer wieder unvorhersehbare Einflüsse auf ein Netzwerk, die sich negativ auf die Dienstqualität auswirken.
- Der Benutzer möchte eine hohe Dienstqualität, der Betreiber aber möglichst viele Verbindungen mit Benutzern bedienen (Interessenkonflikt).
- Damit der Betreiber merkt, wann die Dienstqualität gefährdet ist, muß er ihn überwachen und gegebenenfalls die Benutzer informieren.

Bei einer garantierten Dienstqualität dürfen die Anforderungen des Benutzers niemals verletzt werden. Es müssen also Ressourcen tatsächlich reserviert werden. Dies ist aber oft zu teuer und auch nicht sinnvoll. Meist genügt ein weniger aufwendiges Verfahren. Welche Lösungen OSI95 dafür anbietet, zeigt der nächste Abschnitt.

3.3 Dienstqualität in OSI95

Zum Bestimmen der Qualitätsparameter bedient sich OSI95 der folgenden Attribute:

- Pflichtwert (*compulsory value*): Wird dieser Wert mit in die „Verhandlungen“ eingebracht, wird seine Einhaltung vom Betreiber überwacht und es kommt zu einem Verbindungsabbruch (Dienstabbruch), wenn der Wert nicht mehr eingehalten werden kann. Der Betreiber garantiert also nicht, das ein Dienst mit der dem Pflichtwert entsprechenden Qualität erbracht wird, sondern nur, das solange der Dienst überhaupt erbracht wird, er mindestens die dem Pflichtwert entsprechende Qualität hat.
- Schwellwert (*threshold value*): Sobald der Betreiber nicht mehr in der Lage ist, die Einhaltung dieses Wertes zu gewährleisten, werden die Benutzer darüber informiert.
- Maximalwert (*maximal quality value*): Dieser Wert stellt die maximal erwünschte Qualität dar. Denn nicht immer ist es wünschenswert, die Dienstqualität beliebig zu steigern, nur weil der Betreiber noch freie Ressourcen hat. Besonders dann nicht, wenn der Benutzer nicht über ausreichend Verarbeitungsleistung verfügt, oder wenn der Betreiber dem Benutzer die Qualität seiner Leistungen in Rechnung stellt.

Die Einführung des Pflichtwertes ist keine echte Garantie für die Dienstqualität, sie zwingt aber den Betreiber dazu, eine schwierige Aufgabe, nämlich die Überwachung dieses Wertes zu lösen. Hierfür kann es mitunter nötig sein, daß der Benutzer den Betreiber zur Aushandlungszeit mit zusätzlichen Informationen versorgen muß¹⁰. Der Schwellwert (*threshold value*) kann auch ohne den Pflichtwert (*compulsory value*) verwendet werden; es kommt dann nur zu einer Warnung, wenn er unterschritten wird, jedoch nicht zu einem Verbindungsabbruch. Alle drei Werte können mit dem in Bild 5 gezeigten Verfahren ausgehandelt und vom Betreiber überwacht werden. Zusammengefaßt zeigt sich dann folgendes Verhalten des Betreibers in Abhängigkeit der Dienstqualitäts-Semantik:

	Best Mögliche Qualität	Maximale Qualität	Schwellwert	Pflichtwert	Garantierte Qualität
Verbindliches Ergebnis	NEIN	JA	NEIN	NEIN	JA
Überwachung des Wertes	NEIN	JA	JA	JA	NEIN
Abbruch wenn erreichter Wert schwächer ist	NEIN	–	NEIN	JA	–
Anzeige wenn erreichter Wert schwächer ist	NEIN	–	JA	NEIN	–

Tabelle 2. Verhalten des Betreibers in Abhängigkeit der Dienstqualitäts-Semantik

¹⁰ Eine Anfrage auf einen 2Mb/s Durchsatz mit 10KByte SDUs ist etwas anderes als eine Anfrage auf einen 2Mb/s Durchsatz mit 40Byte SDUs

4 Der verbindungsorientierte Transportdienst in OSI95

Der verbindungsorientierte Transportdienst in OSI95 stellt folgende Neuerungen gegenüber dem ISO/IEC-Transportdienst bereit [Bag94]:

- unidirektionale Verbindungen
- einige unmanagebare Leistungsparameter werden weggelassen, dafür kommen andere wie etwa die Verzögerungsschwankung (*transit delay jitter*) hinzu.
- eine auf dem Konzept von Pflichtwert (*compulsory value*) und Schwellwert (*threshold value*) aufbauende neue Leistungsparameter-Semantik wird eingeführt.
- eine flexiblere Fehlerkontrolle und Fehlerkorrektur.
- eine nach Übertragungsrichtung getrennte Verbindungsabbruchmöglichkeit.
- die Möglichkeit zur Neuaushandlung der Dienstqualität.
- die Übertragung von *out-of-band data*, welche zu einer Verbindung assoziiert werden.

4.1 Verbindungsaufbau und Übertragungsrichtung

Die OSI95-Transportverbindung (*TC, transport connection*) wird auf klassische Weise mit den vier T-CONNECT-Primitiven *request*, *indication*, *response* und *confirm* aufgebaut. Allerdings gibt der neue Parameter *direction to open* in jedem Primitiv die Übertragungsrichtung an. Dieser Parameter kann die Werte *bidirectional*, *direct unidirectional* und *reverse unidirectional* annehmen. Ein weiterer Parameter in den T-CONNECT-Primitiven ist die Größe der kürzesten und der längsten vom Sender verschickten TSDU. Er ist für den Betreiber zur Überprüfung weiterer Qualitätsparameter von Bedeutung.

4.2 Neue Dienstqualitätsparameter

Die TC-Parameter *protection* und *priority* werden bisher aus der ISO/IEC-Definition übernommen. Die Leistungsparameter Durchsatz und Übertragungsverzögerung werden um den *transit delay jitter* erweitert und nun mit der neuen Pflichtwert- und Schwellwert-Semantik nach dem Verfahren gemäß Abbildung 5 gezeigt ausgehandelt. Dabei wird der Durchsatz etwas anders als in ISO/IEC definiert, nämlich wie in 2.2 beschrieben. Ein weiterer Parameter gibt an, ob die Datenpakete *isochron* gesendet werden. Diese Information ist für den Betreiber wichtig, denn sie bestimmt, nach welchem Zeitraum er wieder in der Lage sein muß, ein weiteres T-DATA.req entgegenzunehmen. Mit der neuen Semantik verhält sich der Betreiber dann wie in Tabelle 2 aufgeführt.

Dabei ist T_0 der Zeitpunkt, an dem das letzte T-DATA.req aufgetreten ist, $\Delta t_{max} := L / \text{minimaler compulsory Durchsatz}$, $\Delta t_{thres} := L / \text{threshold Durchsatz}$ und $\Delta t_{min} := L / \text{maximaler Durchsatz}$, wobei L den momentan gemessenen Durchsatz bezeichnet.

Zeitintervall:	Verhalten des Betreibers
$]T_0, T_0 + \Delta t_{min}[$	Der OSI95–TC–Betreiber darf keine T-DATA.req annehmen.
$]T_0 + \Delta t_{min}, T_0 + \Delta t_{thres}[$	Der OSI95–TC–Betreiber darf ein T-DATA.req annehmen, muß aber nicht.
$[T_0 + \Delta t_{thres}, T_0 + \Delta t_{max}[$	Der OSI95–TC–Betreiber sollte in der Lage sein weitere T-DATA.req anzunehmen, anderenfalls müssen beide Benutzer so schnell wie möglich informiert werden.
$[T_0 + \Delta t_{max}, \infty[$	Der OSI95–TC–Betreiber muß jederzeit weitere T-DATA.req annehmen oder anderenfalls die Verbindung so schnell wie möglich abbrechen.

Tabelle 3. Das Verhalten des Betreibers in Abhängigkeit der QoS Semantik

4.3 Fehlerbehandlung

Die vollständige Fehlerbehandlung, wie sie der ISO/IEC-verbundungsorientierte Dienst anbietet, läßt sich bei heutigen Anwendungen¹¹ nur mit viel Aufwand gewährleisten, wenn zusätzlich noch starke Anforderungen an den *jitter* gestellt werden. Hinzu kommt noch, daß einige Anwendungen, wie das oben erwähnte Videokonferenzing, eine gewisse Fehlertoleranz aufweisen, so daß die vollständige Korrektur unnötig ist, und nur Ressourcen verschwendet werden. Der OSI95 verbundungsorientierte Transportdienst bietet nun einige Fehlerbehandlungsstrategien an, die zwischen den Benutzern¹² ausgehandelt werden. Dabei wird zwischen vollständigem Verlust von ganzen oder von Teilen von TSDUs und dem Verfälschen¹³ derselben unterschieden. Im Einzelnen sind die Möglichkeiten für einen Verlust:

- Der Verlust (von Teilen) von TSDUs ist untolerierbar
- Der Verlust (von Teilen) von TSDUs ist tolerierbar, wird aber angezeigt.

Und zur Verfälschung:

- Die Verfälschung (von Teilen) von TSDUs ist untolerierbar.
- Die Verfälschung (von Teilen) von TSDUs ist tolerierbar wird aber angezeigt
- Die Verfälschung (von Teilen) von TSDUs in tolerierbar und wird nicht angezeigt.

Es gibt somit sechs (2 mal 3) Kombinationen von Verlust- und Verfälschungs-Formen, welche zu einer entsprechenden Anzahl von Fehlerbehandlungsstrategien führen, die zwischen den Benutzern ausgehandelt werden. Zur Vereinfachung werden vollständig verlorene TSDUs als TSDUs der Länge Null weitergereicht. Neben dem Strategieparameter

¹¹ Etwa Multimedia

¹² Der Betreiber hat damit nichts am Hut.

¹³ Das hängt vom Kontext ab, in dem die Daten verwendet werden.

gibt es noch die Parameter *dummy replacement indicator* und den *dummy replacement character*, welche jenes Zeichen bestimmen, das an den fehlerhaften Stellen in einer TSDU eingesetzt wird. Zwei weitere Parameter, *lost TDSU error rate* und die *corrupted TSDU error rate* werden nach der bereits besprochenen Semantik für Leistungsparameter ausgehandelt. Alle Fehlerbehandlungsparameter werden immer nur für eine Übertragungsrichtung bestimmt.

Schließlich werden noch die zwei Parameter, welche den Fehlertyp und den betroffenen Bereich innerhalb eine TDSU angeben, eingesetzt. Dies sind die *transfer status parameter*. Als betroffener Bereich wird in diesem Zusammenhang der kleinste zusammenhängende Block (Abb.6) bezeichnet, der alle Fehler enthält.

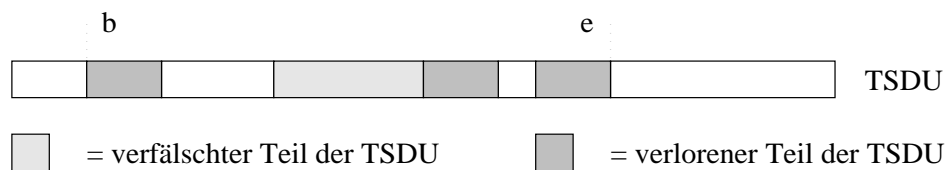


Abbildung 6. Betroffener Bereich im Transferstatus

All diese Parameter sind nur von Bedeutung, wenn sie im Kontext der gewählten Fehlerbehandlungsstrategie sinnvoll sind.

4.4 Neuverhandlung der Dienstqualität

Eine weiterer vom OSI95 Verbindungstransportdienst angebotener Dienst ist die Neuverhandlung von Qualitätsparametern. Dabei kann der Verbindungsinitiator ein T-RENEGOTIATE.request senden und damit die Neuverhandlung einleiten. Nachdem der Empfänger ein T-RENEGOTIATE.response abgegeben hat, paßt dieser sein Sendeverhalten an die neu gewählten Qualitätsparameter an. Der Sender paßt bei Erhalt von T-RENEGOTIATE.confirm sein Sende- und Empfangsverhalten an. Damit der Empfänger weiß, ab wann er sein Empfangsverhalten anpassen soll, gibt es das neue Primitiv T-NEW-QOS welches mit seinem Eintreffen beim Empfänger diesen Zeitpunkt bestimmt.

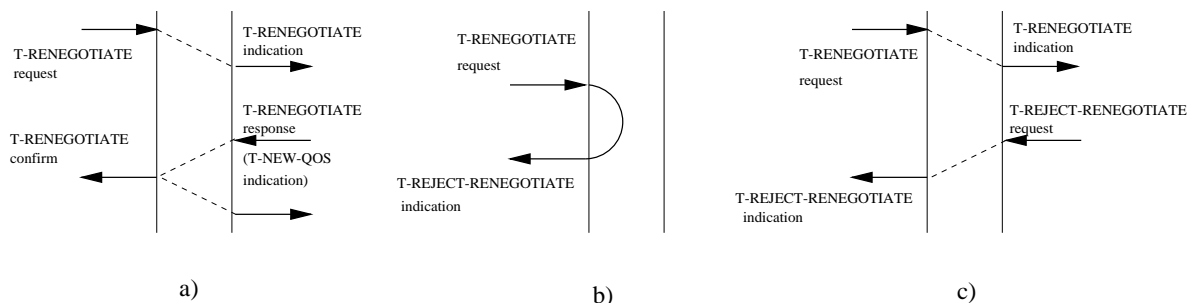


Abbildung 7. Zeitdiagramm einer a) erfolgreichen und b),c) nicht erfolgreichen Neuverhandlung

Es ist nur möglich, die Qualitätsparameter von Verbindungen neuzuverhandeln, die noch

nicht geschlossen sind. Wenn eine Neuverhandlung nicht erfolgreich ist, wie in Bild 6b) bzw 6c) gezeigt, dann werden die alten Qualitätsparameter-Werte beibehalten.

Mit Hilfe von sogenannten *out-of-band-data* besteht die Möglichkeit, daß sich Sender und Empfänger spezielle, die Verbindung betreffende Informationen zukommen lassen¹⁴. Out-of-band-data ist eine Art bestätigter verbindungsloser Dienst, der allerdings mit der Verbindung assoziiert wird.

5 Der verbindungslose Transportdienst in OSI95

Prinzipiell gibt es drei Formen des verbindungslosen Transportdienstes in OSI95 [YBD94]:

- ein unbestätigter verbindungsloser Dienst, der eine verbesserte Version des ISO/IEC-Dienstes darstellt,
- ein bestätigter verbindungsloser Dienst,
- und ein *request/response* verbindungsloser Dienst

Die erste Form erweitert den ISO/IEC-Dienst um einen *compulsory-transit-delay*-Parameter, nach dessen Überschreitung das Datenpaket nicht mehr ausgeliefert wird. Die Parameter *protection* und *priority* bleiben erhalten und der Fehlerwahrscheinlichkeits-Parameter wie er beim ISO/IEC Dienst verwendet wird, wird weggelassen, da er unmanagbar ist.



Abbildung 8. Eine Folge von T-UNITDATA Primitiven

Die zweite Form, der **bestätigte verbindungslose Transportdienst** hat die Primitive T-ACKDATA.req, T-ACKDATA.ind, T-REJECT-ACKDATA.ind und T-ACKDATA.conf. Letzteres zeigt dem Sender die erfolgreiche Auslieferung des T-ACKDATA.ind durch den Betreiber an den Empfänger an. Im Falle eines Fehlers wird stattdessen ein T-REJECT-ACKDATA vom Betreiber an den Sender geschickt.

Die Dienstqualitätsparameter-Parameter bestehen nun neben den bekannten *priority*- und *protection*-Parametern, aus

- *maximum service completion delay*: Er gibt die Zeit an, nach der ein T-ACKDATA.conf eingetroffen sein muß; ansonsten gilt das gesendete Paket als verloren¹⁵.

¹⁴ etwa kann der Empfänger vom Sender das Recht auf Neuverhandlung anfordern.

¹⁵ obwohl es durchaus sein kann, daß nur die Bestätigung verloren ging.

- *compulsory transit delay*: Dieser Parameter steht im T-ACKDATA.req. Sollte er überschritten werden, wird das Datenpaket nicht mehr ausgeliefert und der Sender wird mit einem T-REJECT-ACKDATA.ind hierüber informiert. Hierbei ist der *Zeitraum zwischen dem Auftreten eines T-ACKDATA.req und dem Auftreten des entsprechenden T-ACKDATA.conf am gleichen TSAP* relevant.
- *threshold transit delay*. Wenn dieser Wert überschritten wird, werden Sender und Empfänger darüber in ihren T-ACKDATA.conf bzw T-ACKDATA.ind informiert. Das geht selbstverständlich nur, wenn die entsprechenden Primitive auch auftreten.

In beiden Fällen ist der *transit delay* definiert als *die zwischen dem Auftreten des T-ACKDATA.req und dem des entsprechenden T-ACKDATA.ind beim Partner TSAP verstreicht*. Mit dem T-REJECT-ACKDATA wird außerdem noch die Ursache für das Scheitern der Übertragung zurückgemeldet.

Dieser Dienst eignet sich nur für Anwendungen, die duplizierte TSDUs tolerieren können, also für Anwendungen bei denen es nur wichtig ist, daß ein Datenpaket mindestens einmal und nicht genau einmal beim Empfänger angekommen.

Der **request/response verbindungslose Transportsdienst** stellt eine Erweiterung des eben beschriebenen bestätigten Dienstes dar. Er ist nützlich für Anwendungen, die mit einer Anfrage eine Antwort assoziieren, wie es etwa Klient/Server-Anwendungen tun. Die Primitive heißen nun T-REQRES.req, T-REQRES.ind, T-REQRES.conf und T-REQRES.res. Letzteres trägt die Antwort-TSDU. Das Primitiv T-REJECT-REQRES spielt die gleiche Rolle wie das T-REJECT-ACKDATA Primitiv. Die *transit delay* Leistungsparameter sind bei diesem Dienst nicht notwendig, da nicht die Anfrageübermittlung, sondern das Erhalten der Antwort die Intention einer Dienstinanspruchnahme ist.

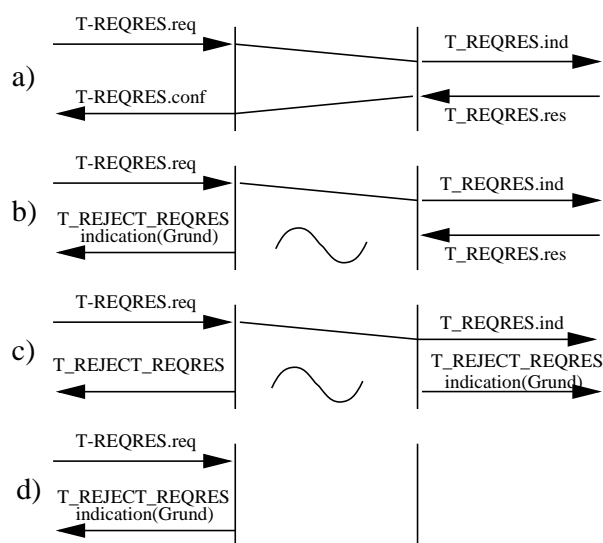


Abbildung 9. Einige request/response Szenarien a) Erfolgreiche Übermittlung b),c) Erfolgreiche Übermittlung, aber unter Umständen werden Duplikate erzeugt. d) Der Betreiber lehnt den Dienst ab.

Ansonsten ist der request/response Transportdienst genauso aufgebaut wie der bestätigte Dienst. Abb.9 zeigt einige möglich Aufrufzenarien.

6 Zusammenfassung

Die Vorschläge von OSI95 zur Verbesserung der Kommunikationsdienste sind ausreichend flexibel, um für die Anforderungen heutiger Anwendungen, und den damit verbundenen Problemen eine Lösung zu bieten. Speziell die Pflichtwert–Schwellwert–Maximalwert–Semantik stellt einen leistungsfähigen Mechanismus dar, um die Qualität von Transportdiensten zu verbessern. Man kann nur hoffen, daß zukünftige Protokollimplementierungen sich dieser wohlüberlegten Mechanismen bedienen.

Im Vergleich: QOS-Management für multimediale Anwendungen

Robert Clauß

Kurzfassung

Verteilte multimediale Anwendungen stellen neuartige Anforderungen an die verwendeten Telekommunikationsdienste, welche deshalb ein spezielles QOS-Management einsetzen müssen. In dieser Ausarbeitung werden zuerst die Anforderungen multimedialer Anwendungen vorgestellt. Dann werden Qualitätsparameter als Mittel Anwendungsbedürfnisse zu spezifizieren betrachtet. Die Ressourcen, die zur Erbringung von multimedialen Kommunikationsdiensten benötigt werden, werden vorgestellt. Daraufhin werden Grundkonzepte und Methoden des multimedialen Ressourcenmanagements untersucht. Dienstgarantien können nur mittels Ressourcenmanagement vergeben werden. Drei unterschiedliche Architekturansätze zum QOS-Management für multimediale Anwendungen werden aufgezeigt, und dann bezüglich ihrer Architektur und Funktionalität verglichen.

1 Einleitung und Motivation

Durch die rasanten Fortschritte im Telekommunikationsbereich und in der Rechnertechnologie in den letzten Jahren gewinnt auch das Gebiet der vernetzten, multimedialen Anwendungen (Networked Multimedia Systems, NMS) immer mehr an Bedeutung. Populäre Beispiele für solche Anwendungen sind Videokonferenzen, Video on demand, Bildtelefonie, Multimedia-Mail-Systeme, Screensharing und Virtual Desktops. All diesen Anwendungen ist gemeinsam, daß sie gleichzeitig mehrere unterschiedliche Medienströme verarbeiten müssen. Unter einem Medienstrom versteht man hierbei einen unidirektionalen Strom digitaler Daten von einer Quelle zu einer Senke. Die digitalen Daten werden entsprechend dem jeweiligen Medium an der Quelle erzeugt — Videodaten können zum Beispiel von einer Kamera aufgenommen und digitalisiert werden — , daraufhin mit Hilfe eines Kommunikationsdienstes zur Senke übertragen und dort weiterverarbeitet bzw. in der Regel dem Medium entsprechend wiedergegeben. Dies führt zu neuen Anforderungen an Telekommunikationsdienste:

- Es muß garantiert werden können, daß ein Dienst in der ausgehandelten Qualität erbracht wird. So sind zum Beispiel viele multimediale Anwendungen verzögerungsempfindlich und benötigen deshalb Zeitgarantien. Video-on-demand-Anwendungen brauchen beispielsweise Übertragungszeitgarantien, um eine gewisse Anzeigequalität garantieren zu können.
- Es muß möglich sein, die Dienstqualität einer laufenden Verbindung dynamisch an veränderte Anwendungsbedürfnisse anzupassen.

- Kommunikationsdienst-, Betriebssystem- und Multimediaressourcen müssen als zusammengehörige Elemente eines einzigen Systems verwaltet werden, da sie wechselseitig voneinander abhängig sind und sich gegenseitig beeinflussen können.
- Häufig müssen Kommunikationsdienste sehr hohen Anforderungen bezüglich Verzögerung, Jitter oder Durchsatz gerecht werden. Bei Videoverbindungen kann der geforderte Durchsatz beispielsweise zwischen 4 MBit/sec und 100 MBit/sec schwanken.
- Beziehungen zwischen den Medienströmen müssen unterstützt werden. Man unterscheidet dabei im wesentlichen Synchronisationsbeziehungen (zum Beispiel Lippen-synchronisation bei Bildtelefonie) und logische Beziehungen (beispielsweise Strom A muß nach Strom B verarbeitet werden).
- Außer den herkömmlichen Punkt-zu-Punkt-Verbindungen werden häufig Multicast-Verbindungen gefordert. Ein Beispiel hierfür sind Videokonferenzen.

Die traditionellen Telekommunikationsdienste können diesen Anforderungen nur teilweise gerecht werden, da einerseits ihre Funktionalität nicht ausreichend und andererseits ihre Architektur nicht problemadäquat ist. In jüngster Vergangenheit wurden deshalb eine Vielzahl neuartiger Ansätze entwickelt.

2 Ressourcen und Qualitätsparameter

Der vom Telekommunikationssystem zu erbringende Dienst wird durch Qualitätsparameter spezifiziert und mit den zur Verfügung stehenden Ressourcen erbracht. Diese Ressourcen werden basierend auf den Qualitätsanforderungen reserviert und zugeteilt.

2.1 Ressourcen

Eine Hauptaufgabe des Multimedia-Kommunikationsdienstes ist die Verwaltung der unterschiedlichen Ressourcen zur Erbringung einer Dienstleistung unter Berücksichtigung der Abhängigkeiten zwischen den Ressourcen. Ressourcen werden dabei in

- Netzwerkressourcen (z Bsp. Bandbreite, Puffer in den Netzknoten)
- Betriebssystemressourcen (z. Bsp. Pufferspeicherblöcke, CPU-Zeit)
- und Multimediaressourcen (z. Bsp. Geräte wie Videokameras, Mikrophone, Lautsprecher)

unterteilt. Außerdem lassen sie sich folgendermaßen klassifizieren:

- Unterscheidung in aktive Ressourcen wie CPUs oder Netzwerkadapter, die Dienste anbieten, und passive Ressourcen wie Hauptspeicher oder Bandbreite, die zur Dienstleistung benötigt werden.
- Unterscheidung in exklusive Ressourcen wie Lautsprecher und geteilte Ressourcen wie Bandbreite.

2.2 Qualitätsparameter

Durch Spezifikation über mehrere Qualitätsparameter sind Multimedia-Kommunikationsdienste sehr flexibel und können von einer Vielzahl unterschiedlicher Anwendungen verwendet werden. Die unterschiedlichen Qualitätsparameter lassen sich drei Ebenen zuordnen. Auf der Anwendungsebene spezifiziert eine Anwendung ihre Bedürfnisse mit Hilfe von anwendungsorientierten Qualitätsparametern. Üblicherweise werden hier Medientyp, Übertragungsart, die gewünschte Medienqualität (beispielsweise Rahmengröße bei Videoanwendungen) und die Beziehungen zwischen den Medienströmen angegeben. Die anwendungsorientierten Parameter werden auf der Systemebene in systemnahe Qualitätsparameter übersetzt, die Betriebs- und Kommunikationssystemanforderungen spezifizieren. Man unterscheidet hier quantitative Parameter (z. Bsp. Fehlerzahl, Übertragungsrate, Prozeßrechenzeit) und qualitative Parameter (z. Bsp. Auslieferungsdisziplin, Synchronisationsart). Auf der Geräteebene werden netzorientierte und geräteabhängige Qualitätsparameter benötigt. Netzorientierte Parameter sind beispielsweise Verzögerung, Bandbreite oder Jitter, geräteabhängige Parameter geben Zeit- und Durchsatzanforderungen an Multimediageräte an.

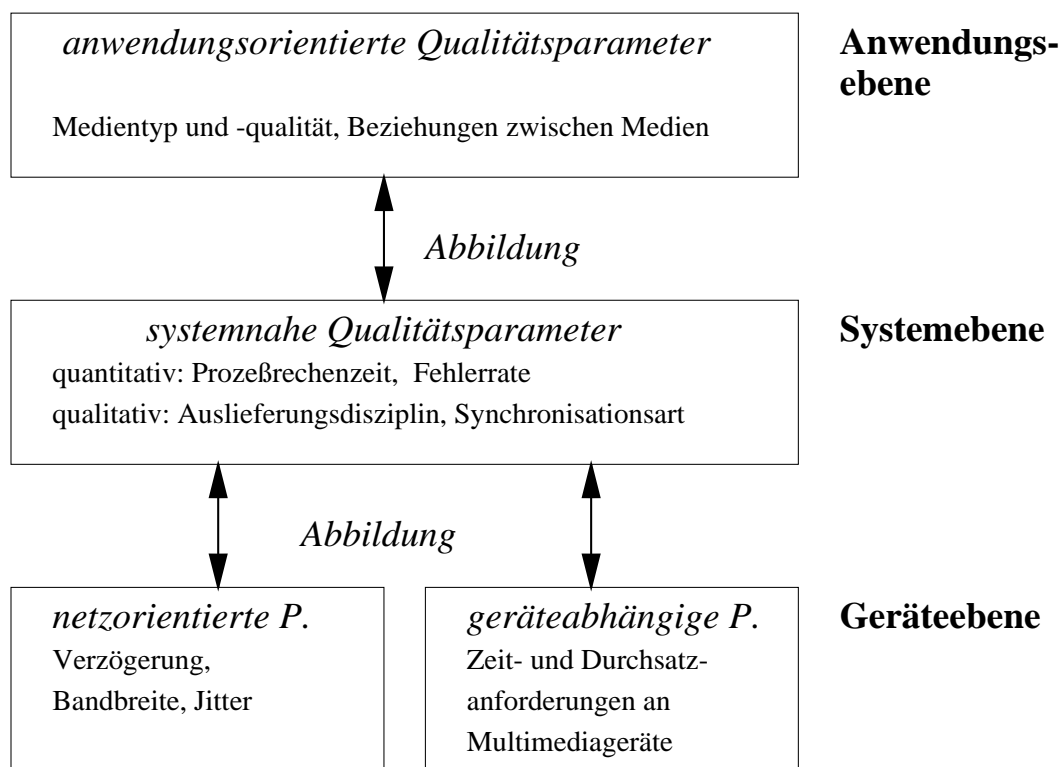


Abbildung 10. QOS-Parameter auf den unterschiedlichen Systemebenen

Bei der Erbringung eines von einer multimedialen Anwendung benötigten Telekommunikationsdienstes geht man dann folgendermaßen vor. Die mittels anwendungsorientierten Parametern spezifizierten Anwendungsbedürfnisse werden in entsprechende systemnahe, netzorientierte und geräteabhängige Parameter übersetzt. Diesen Parametern werden — beispielsweise unter Einsatz bestimmter Heuristiken — die benötigten Ressourcen

zugeordnet, die mit einer Ressourcenmanagementarchitektur bereitgestellt werden. Diese Architektur ermöglicht außerdem die dynamische Anpassung der Dienstqualität an veränderte Anwendungsbedürfnisse während des Betriebs und erlaubt die Ressourcenfreigabe beim Verbindungsabbau.

3 Ressourcenmanagement

Das Ressourcenmanagement bei vernetzten, multimedialen Anwendungen läßt sich grob in Management während der Auf- bzw. Abbauphase und Management während der Übertragung unterteilen. Genauere Informationen zu diesem Thema findet man in [NK95b].

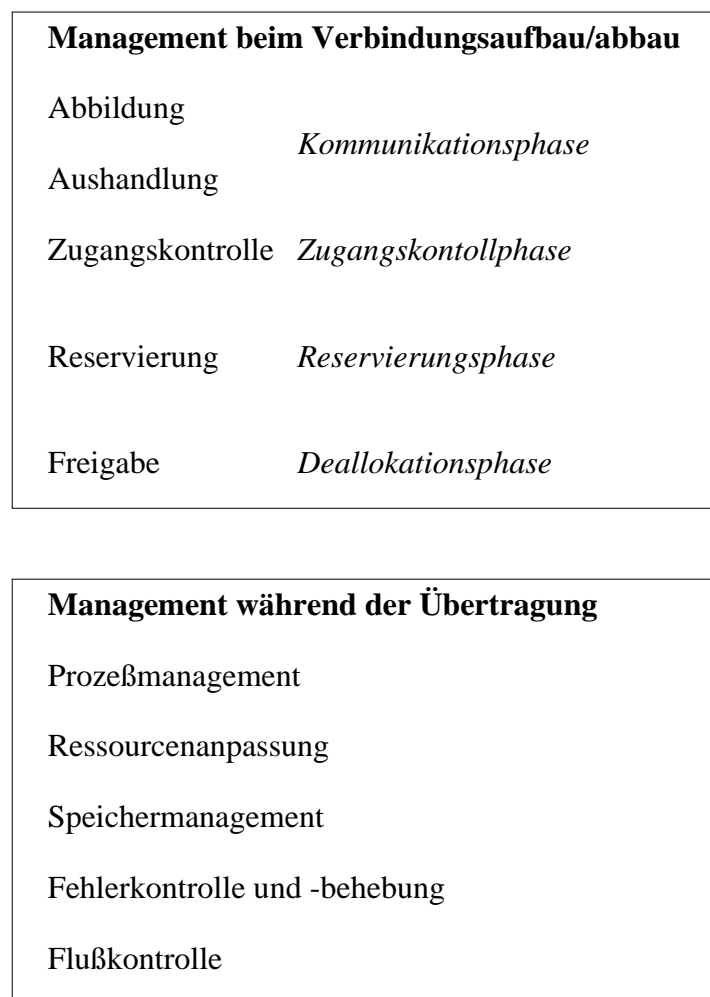


Abbildung 11. Ressourcenmanagement bei multimedialen Anwendungen

3.1 Ressourcenmanagement beim Verbindungsaufbau bzw. -abbau

Das Management während des Verbindungsaufbaus bzw. -abbaus beinhaltet folgende Phasen. Die *Kommunikationsphase* unterteilt sich in Übersetzungsvorgänge und Verhandlungsvorgänge. Wie oben schon angedeutet müssen Qualitätsparameter zwischen

unterschiedlichen Systemebenen übersetzt werden. So hat beispielsweise bei einer Videoverbindung die Rahmengröße und Farbtiefe, die durch die Anwendung spezifiziert werden, keine Bedeutung für das verwendete Netzwerk. Qualitätsparameter müssen von den beteiligten Komponenten ausgehandelt werden. Man unterscheidet vertikale Verhandlungen (Schicht zu Schicht, Dienstbenutzer zu Dienstbringer) zwischen unterschiedlichen Schichten (zum Beispiel Aushandlung einer Puffergröße zwischen Anwendungsschicht und Betriebssystem) und horizontale Verhandlungen (Initiator zu Responder) zwischen den beteiligten Endpunkten. Wenn dabei der Dienstbringer keine Einwirkungsmöglichkeiten hat, spricht man von einer bilateralen Verhandlung, ansonsten von einer Dreiecksverhandlung (es gibt drei Verhandlungsteilnehmer).

In der *Zugangskontrollphase* wird überprüft, ob die angeforderte Dienstqualität auch erbracht werden kann. Dazu werden eine Reihe von Tests durchgeführt, die die Ressourcenverfügbarkeit überprüfen.

Danach folgt die *Reservierungsphase* in der die benötigten Ressourcen reserviert werden. Man kann dabei einen optimistischen Ansatz verfolgen und für den Durchschnittsfall reservieren, oder einen pessimistischen Ansatz und für den schlechtesten Fall reservieren, d.h. man fordert möglichst lange CPU-Zeiten, hohe Bandbreite an. Während der Übertragung findet dann eine Zuteilung statt, die reservierte Ressourcen explizit für die Übertragung bereitstellt.

In der *Deallokationsphase* werden die reservierten Ressourcen wieder freigegeben.

3.2 Ressourcenmanagement während der Datenübertragung

Während der Übertragung werden folgende Managementtätigkeiten durchgeführt: Mit dem *Prozeßmanagement* werden der möglicherweise großen Menge der Multimediaprozesse gemäß einer Verteilungsstrategie Prozessorzeit zugewiesen. In der Regel müssen damit die Echtzeitanforderungen an das System unterstützt werden, d.h. bestimmten Prozessen muß eine Mindestrechenzeit garantiert werden.

Durch spezielles *Puffermanagement* soll der bei der Übertragung anfallende Speicherbedarf erfüllt werden. Häufig erweisen sich traditionelle Speicherverwaltungsmethoden als unzureichend.

Um die ausgehandelte Datentransferrate einzuhalten, werden bestimmte ratenbasierte *Flußkontrollmechanismen* ausgeführt, die dem Nutzer eine minimale Transferrate garantieren, unabhängig von der Auslastung des Kommunikationssystems.

Es werden *Fehlererkennungs- und Fehlerbehebungsmechanismen* eingesetzt, die auch strukturelle Informationen innerhalb eines Datenstroms schützen. Außerdem wird ein Ressourcenmonitor betrieben, mit dem man das Systemverhalten bei den ausgehandelten Qualitätsparametern feststellen kann, und deren Einhaltung überprüfen kann.

Eine weitere Aufgabe ist das *dynamische Anpassen der Dienstqualität* während der Übertragung. Die Änderung kann durch den Nutzer verursacht werden oder vom Dienstbringer zum Beispiel bei Überlastungssituationen im Netzwerk. Es erfolgt dann eine Bekanntmachung dieser geänderten Anforderungen an alle beteiligten Komponenten, eine Neuverhandlung der Dienstqualität und schließlich eine Ressourcenanpassung. Man

unterscheidet dabei zwei Ansätze: Bei der Netzwerkanpassung wird beispielsweise versucht durch dynamisches Rerouting das Netzwerk an die Überlastung anzupassen. Bei der Quellenanpassung wird die Quelldatenrate an die aktuell verfügbaren Ressourcen angepaßt.

4 Der QOS-Broker

An der Universität von Pennsylvania wurde der QOS-Broker entwickelt, der an den Verbindungsendpunkten einer verteilten, multimedialen Anwendung platziert ist und ein QOS-Management für Endsystemverbindungen bietet. Der Broker setzt auf dem darunterliegenden Netzwerkressourcenmanagement auf und arbeitet mit diesem zusammen, um die gewünschte Dienstqualität zu erbringen. Vorgestellt wird er beispielsweise in [NK95a]. Die Hauptaufgabe einer QOS-Broker-Instanz ist dabei die Verwaltung der lokalen Betriebssystem- und Multimediaressourcen und die Verhandlung mit dem lokalen Netzwerkressourcenmanager und dem entfernten QOS-Broker. Der Broker bereitet das Telekommunikationssystem auf eine garantierte Übertragung vor und nimmt an, daß sich das Telekommunikationssystem entsprechend dem von ihm ausgehandelten QOS-Profil verhält. Das QOS-Profil ist eine genaue Spezifikation aller beteiligten Qualitätsparameter auf Anwendungs-, System- und Geräteebe.

4.1 Die QOS-Broker-Architektur

Das den multimedialen Anwendungen zur Verfügung gestellte Kommunikationssystem läßt sich in eine anwendungsnahe, eine transportorientierte und eine netzorientierte Schicht unterteilen. Der QoS-Broker erstreckt sich über die Anwendungs- und Transportebene (siehe Abbildung 12). Das Anwendungssubsystem des Brokers, übernimmt Aufgaben wie I/O-Gerätemanagement, Mediensynchronisation und Medienauslieferung an die Anwendung. Es greift dabei auf Multimedia- und Betriebssystemressourcen über ressourcenspezifische Schnittstellen zu. Das Transportsubsystem übernimmt Aufgaben wie Verbindungsaufbaumanagement, Fehlerbehebung und vergibt Übertragungsaufträge an das darunterliegende Netzwerk. Der Broker führt eine Aushandlung durch und erstellt ein vollständiges QOS-Profil. Mittels diese QOS-Profils werden die benötigten Ressourcen allokiert und das Netzwerk beauftragt.

Abbildung 13 zeigt die Architektur einer QOS-Broker-Instanz. Eine QOS-Broker-Instanz besteht aus den Teilen Broker-Sender und Broker-Empfänger. Der Broker-Sender ist für das Ressourcenmanagement bei abgehenden Verbindungen zuständig, der Broker-Empfänger für das Management eingehender Verbindungen. Beide Teile enthalten jeweils einen Käufer- und einen Verkäufer-Protokollautomaten. Damit wird folgendes Konzept realisiert: es gibt Käufer, die Ressourcen für eine Verbindung bei entfernten Knoten kaufen wollen, und Verkäufer, die ihrerseits Ressourcen für Verbindungen anbieten. Bei den vom Käufer initiierten Kaufverhandlungen werden folgende Aktionen durchgeführt (die Abfolge der einzelnen Aktionen wird im nächsten Kapitel erläutert):

- Der Käufer verwaltet seine lokalen Betriebssystem- und Multimediaressourcen und stellt mittels Zugangskontrolle ihre Verfügbarkeit fest.

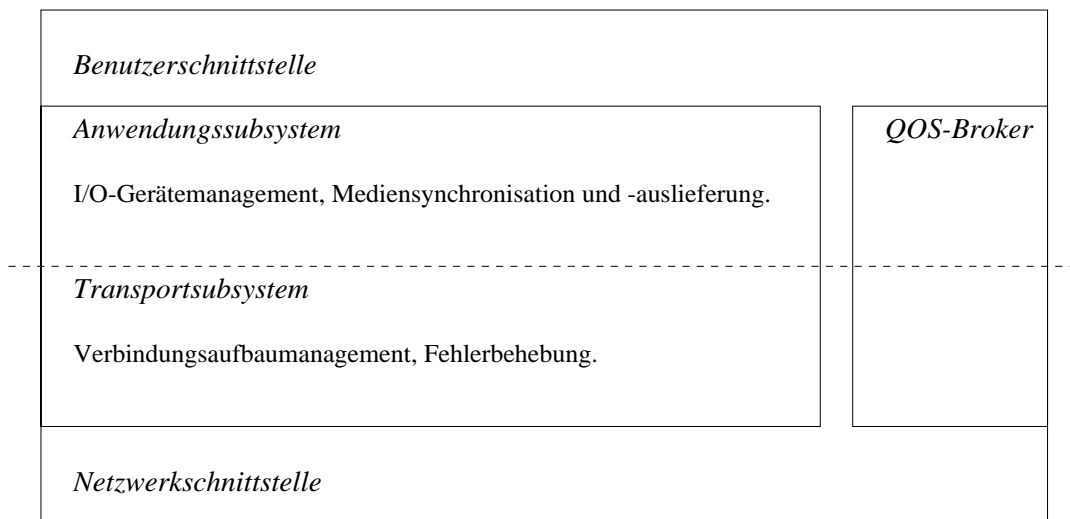


Abbildung 12. Der QOS-Broker innerhalb eines Kommunikationssystems

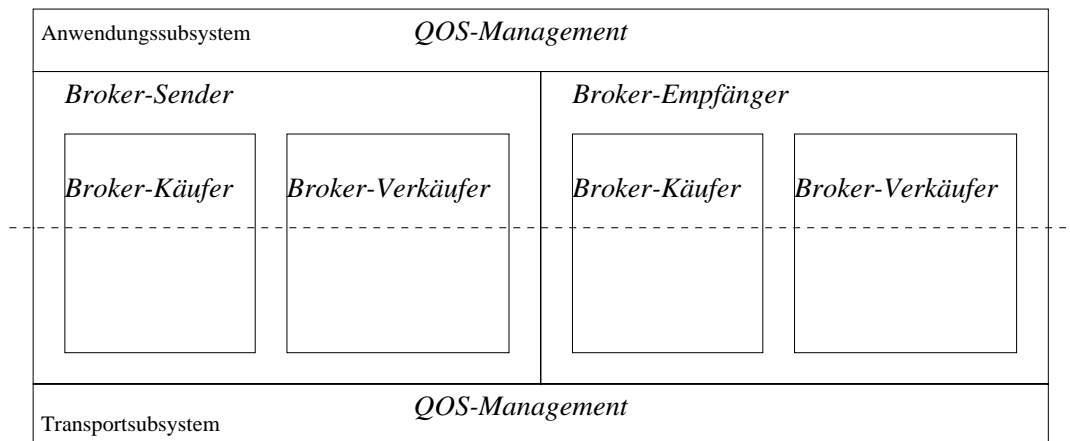


Abbildung 13. Aufbau einer QOS-Broker-Instanz

- Der Käufer sammelt Informationen über die verfügbaren Netzwerkressourcen, die vom Netzwerkressourcenmanagement verwaltet werden.
- Der Käufer sammelt Informationen über die Ressourcen des Zielknotens.
- Der Verkäufer des Zielknotens verwaltet dort die lokalen Ressourcen und liefert dem Käufer Informationen über deren Verfügbarkeit.

4.2 Das QOS-Broker-Protokoll

Das QOS-Broker-Protokoll übernimmt die Aushandlung des QOS-Profiles zwischen den Teilnehmern. Dabei kann die Anzeige eines Verbindungswunsches zu folgenden Reaktionen des Zielknotens führen:

- *accept*, wenn der Verkäufer die geforderten Ressourcen bereitstellen kann.
- *reject*, wenn der Verkäufer nicht genügend Ressourcen hat oder bei Timeouts verursacht durch Netzwerkfehlverhalten.

- *modify*, wenn die vom Verkäufer angebotenen Ressourcen nicht den Forderungen entsprechen aber noch innerhalb der Toleranzgrenze liegen.

Das Protokoll läßt sich in ein Käufer-Protokoll und ein Verkäufer-Protokoll unterteilen. Das Käufer-Protokoll geht im wesentlichen folgendermaßen vor:

- Der Verhandlungsvorgang wird initiiert durch die Eingabe von Anwendungs-Qualitätsparametern für eine gewünschte Verbindung.
- Diese werden in entsprechende System-Qualitätsparameter übersetzt.
- Ein Zugangskontrolldienst untersucht, ob die benötigten lokalen Ressourcen verfügbar sind.
- Alle bisherigen Aktionen finden innerhalb der Anwendungsschicht statt. Nach Zulassung der lokalen Ressourcen wird der Ende-zu-Ende-Verhandlungsprozeß zwischen den beteiligten Brokern auf Anwendungsebene initiiert.
- Ist die Antwort *accept* oder *modify*, findet eine Übersetzung in netzorientierte Qualitätsparameter statt.
- Ein Zugangskontrolldienst innerhalb des Transportsystems überprüft die Verfügbarkeit der für den Transport benötigten lokalen Ressourcen und kommuniziert mit dem Netzwerkressourcenmanagement, um die Verfügbarkeit von Ressourcen wie Bandbreite oder Ende-zu-Ende-Verzögerung festzustellen.
- Danach erfolgt eine Ende-zu-Ende-Aushandlung auf Transportebene.
- Der Broker wartet auf die Antwort und stellt das Verhandlungsergebnis der Anwendung in Form von anwendungsorientierten Qualitätsparametern dar.

Das Verkäufer-Protokoll beinhaltet folgende Schritte:

- Es wartet auf die Ankunft der systemnahen Qualitätsparameter des Käufers auf Anwendungsebene und entscheidet mittels Zugangskontrolle, ob es die geforderten Ressourcen anbieten kann (ob beispielsweise die geforderten Multimediageräte vorhanden sind).
- Es wird eine Verhandlungsantwort zurückgegeben.
- Analog wird eine Verhandlung auf Transportebene durchgeführt.
- Nach erfolgreicher Verhandlung erhält der Verkäufer Informationen über die verfügbaren Netzwerkressourcen vom Netzwerkressourcenmanagement.

Das QOS-Broker-Protokoll unterstützt des weiteren eine dynamische Anpassung und Neuverhandlung von Verbindungseigenschaften. Außerdem sind Gruppenverhandlungen bei Multicastverbindungen möglich.

5 AQOSM und HLQNP

An der Universität Pierre et Marie Curie in Paris wurden der QOS-Manager AQOSM (Applications QOS Manager) und das Verhandlungsprotokoll HLQNP (High Level QOS Negotiation Protocol) entwickelt. Sie sind Teil einer Architektur zur Erfüllung der Anforderungen multimedialer Anwendungen im Endsystembereich. Dabei wird einerseits die genaue Spezifikation der Anwendungsbedürfnisse durch Qualitätsparameter unterstützt. Es wird ein Sendeprotokoll zur Verfügung gestellt, mit dem Medientyp und Übertragungsart angegeben werden können. Außerdem werden Primitive zur Spezifikation von Beziehungen zwischen Medienströmen und zur Definition von Qualitätsabschwächungen bei Nichterbringbarkeit der ursprünglichen Forderungen bereitgestellt. AQOSM übersetzt dann die anwendungsorientierten Qualitätsparameter und liefert den verwendeten System- und Netzwerkkomponenten die Anwendungsbedürfnisse in einer für sie verständlichen Form. Genauere Informationen zu AQOSM findet man in [TW94].

5.1 Die AQOSM-Architektur

AQOSM ist ein Endsystem-QOS-Manager und wird somit auf jedem beteiligten Endsystem platziert. AQOSM erhält die Übertragungswünsche und Anwendungs-Qualitätsparameter von einer multimedialen Anwendung. Die AQOSM-Architektur wird in Abbildung 14 verdeutlicht. Zur Erbringung des geforderten Dienstes verwendet AQOSM die lokalen Systemressourcen, das sind neben den üblichen Multimedia- und Betriebssystemressourcen auch Softwarekomponenten zur Mediensynchronisation, und den vom Netzwerk zur Verfügung gestellten Kommunikationsdienst. Dabei ist AQOSM nicht direkt am Ressourcenmanagement beteiligt, sondern es informiert die beteiligten Systemmanager, die die Ressourcen verwalten, über den Bedarf. Die AQOSM-Instanzen an Quelle und Senke einer Verbindung kommunizieren mittels HLQNP.

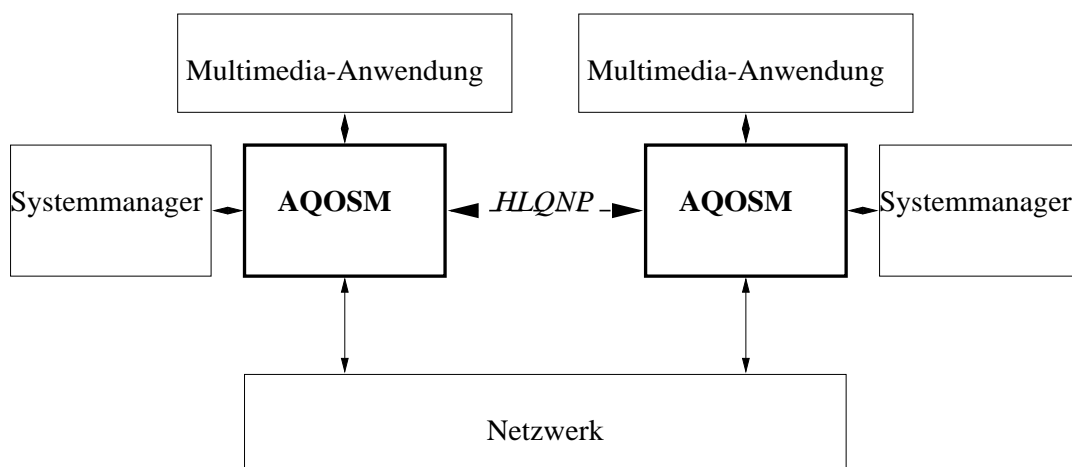


Abbildung 14. Die AQOSM-Architektur

5.2 Interaktionsmodell und HLQNP

Lokal kommunizieren AQOSM, die Systemmanager und der Kommunikationsdienst über eine Reihe von Primitiven:

- AQOSM kann mit einem *Request-Primitiv* einen bestimmten Dienst von einer Ressource anfordern.
- Ressourcen antworten mit einem *Accept-Primitiv*, wenn sie den geforderten Dienst erbringen können, mit einem *Refuse-Primitiv*, wenn der Dienst nicht erbracht werden kann, und mit einem *Modify-Primitiv*, wenn der Dienst nur in abgeschwächter Form erbracht werden kann.
- Die Verfügbarkeit einer Ressource kann von AQOSM mit einem *Check-Primitiv* überprüft werden.
- Schon ausgehandelte Dienste können mit einem *Change-Primitiv* geändert werden und mit einem *Release-Primitiv* beendet werden.
- Mit einem *Notify-Primitiv* kann AQOSM die Anwendung über Probleme bzw. QOS-Änderungen informieren.

Das durch die Primitive definierte Interaktionsmodell wird auch als lokale QOS-Verhandlung bezeichnet. Es erzwingt aber keine garantierten Dienstleistungen der Ressourcen. Um dies zu erreichen, müßte AQOSM die Ressourcen selbst allokalieren und reservieren.

Die verteilten Managementfunktionen leistet AQOSM mit Hilfe von HLQNP. Dazu gehören die Aushandlung von anwendungsorientierten Qualitätsparametern vor Ausführung der Anwendung und die Neuverhandlung von Qualitätsparametern während der Ausführung. Dabei ist Voraussetzung, daß das zugrundeliegende Netzwerk zu dynamischen Qualitätsänderungen fähig ist. Prinzipiell wird HLQNP über einen zuverlässigen out-of-band-Übertragungsdienst abgewickelt. Verhandlungen werden jeweils nur für einen Medienstrom durchgeführt. HLQNP unterstützt die Verhandlungsmodi *negotiated*, in dem der Initiator auf eine Reaktion des Responders wartet, und *forced*, in dem der Initiator sofort nach Abgang des Verbindungsaufbauwunsches senden kann. Das Protokoll wird mit Hilfe von 5 Meldungen abgewickelt:

- *Negotiate* wird vom Initiator abgesendet und übermittelt die Anwendungsanforderungen.
- *Accept* wird vom Responder zurückgeschickt, wenn die Anwendungsanforderungen ohne Änderungen akzeptiert werden.
- *Modify* erlaubt dem Responder in der Verhandlungsphase eine Verbindung mit abgeschwächten Qualitätsparametern zu akzeptieren. Es wird außerdem zur dynamischen QOS-Anpassung während einer Übertragung verwendet.
- Mit *Reject* kann einerseits eine Verbindung vom Responder abgelehnt werden, andererseits wird es zur Anzeige eines Fehlverhaltens des Netzwerkes verwendet.

- Mit der *End* Meldung wird das Ende einer Verbindung angezeigt und die Freigabe der verwendeten Ressourcen veranlaßt.

Das Protokoll führt Informationen zu Medienströmen in der Verhandlungs- bzw. Übertragungsphase in speziellen Datenstrukturen. Dort werden unter anderem die Priorität des Stroms und die vorreservierten bzw. reservierten Ressourcen gespeichert. Mit diesen Informationen kann das Protokoll entscheiden, ob eine neue Verbindung mit höherer Priorität möglich ist, und notfalls Medienströme mit geringerer Priorität bezüglich der Qualitätsparameter abschwächen bzw. ganz beenden, um Ressourcen freizugeben.

6 Die QoS-Architektur

An der Lancaster University wurde die Telekommunikationsdienstarchitektur QOS-A (Quality of Service Architecture) entwickelt. Dabei handelt es sich im wesentlichen um einen Programmierrahmen, der die Spezifikation der Bedürfnisse multimedialer Anwendungen unterstützt und deren Implementierung im Rahmen eines Telekommunikationsdienstes ermöglicht. Es werden ausschließlich ATM-Netzwerke als Vertreter der diensteintegrierenden Netzwerke betrachtet. Die QOS-A ermöglicht ein an multimediale Anwendungen angepaßtes Ressourcenmanagement sowohl im Endsystembereich als auch innerhalb des ATM-Netzwerkes. Das QOS-Management erstreckt sich also über alle Schichten des Kommunikationssystems. Der resultierende Kommunikationsdienst garantiert die Einhaltung einer einmal ausgehandelten QOS-Spezifikation. Die Architektur wird in [CA93] vorgestellt, [CA94] geht auf die Transportschicht der QOS-A ein. Neben der Übersetzung der von den Anwendungen gelieferten QOS-Spezifikationen sind die wichtigsten Aufgaben der QOS-A einerseits die QOS-Aushandlung und Ressourcenreservierung und andererseits die Kontrolle des Dienstbringers durch Monitorfunktionen und die Kontrolle des Dienstbenutzers durch das sogenannte Policing. Dadurch wird überprüft, ob das ausgehandelte Kommunikationsaufkommen vom Benutzer überschritten wird. Die vom Programmierrahmen zur Verfügung gestellten Funktionen lassen sich in Management-Funktionen und QOS-Unterstützungsmechanismen einteilen (siehe Abbildung 15).

6.1 Architektur

QOS-A unterteilt das Kommunikationssystem in 4 Schichten. Auf jeder dieser Schichten müssen neben den traditionellen Funktionen auch Aufgaben zur Unterstützung der Anwendungsanforderungen erledigt werden.

So muß zum Beispiel innerhalb der Anwendungsschicht eine Abbildung der anwendungsorientierten Qualitätsparameter auf folgende Qualitäts-Dimensionen stattfinden:

- Verbindungsaufbau-QOS (entspricht den OSI Qualitätsparametern zum Verbindungsaufbau), Jitter, Verzögerung, Durchsatz und Fehlercharakteristika beziehen sich auf einzelne Medienströme.

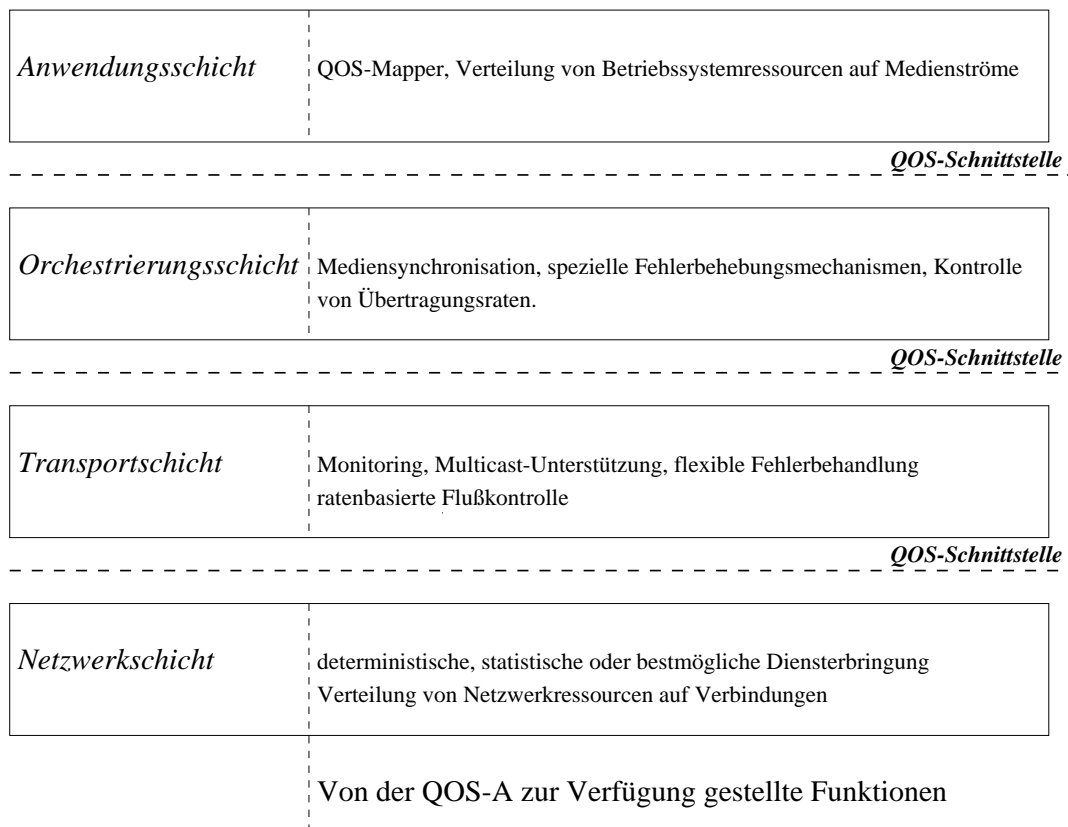


Abbildung 15. Das QOS-A Schichtenmodell

- Synchronisation- und Multicast-Parameter beziehen sich auf mehrere Medienströme

Die Abbildung wird von der QOS-Mapper-Funktion erledigt, die bestimmten Medientypen bzw. -qualitäten die oben aufgeführten Parametern zuordnet. Zwischen den folgenden darunterliegenden Schichten werden die Qualitätsparameter über QOS-Schnittstellen ausgetauscht, die Qualitätsspezifikationen in den Qualitäts-Dimensionen, eine Notation zur Spezifikation der Verpflichtung des Dienstbringers ausgehandelte Qualität zu liefern, eine Kostenspezifikation und Details zur durchzuführenden QOS-Verhandlung enthalten. Die QOS-Unterstützungsmechanismen der Anwendungsschicht beinhalten die Verwaltung der Betriebssystemressourcen und ihre Verteilung auf die einzelnen Medienströme.

Die Aufgabe der Orchestrierungsschicht ist die Bereitstellung eines im Vergleich zur Transportschicht höherwertigen Kommunikationsdienstes. Zum einen wird hier die Mediensynchronisation durchgeführt, zum anderen die Übertragungsrate kontrolliert und Fehlerbehebungsmechanismen angeboten.

Transportdienste, die vergleichbar sind mit dem Dienst der OSI-Transportschicht, bedienen traditionell eher die Bedürfnisse von Dateitransfer oder interaktiven Anwendungen mit geringem Bandbreitenverbrauch. Sie sind daher für den Transport kontinuierlicher, multimedialer Daten ungeeignet. QOS-A sieht deshalb einen speziellen Transportdienst vor, der den Forderungen nach QOS-Aushandlung und Neuverhandlung, Monitoring, flexiblen Fehlerbehandlungsmechanismen und Multicast-Verbindungen gerecht wird. Die wichtigsten QOS-Unterstützungsmechanismen der Transportschicht sind ein ratenbasier-

ter Flußkontrollmechanismus, mit dem man einen bestimmten Durchsatz gewährleisten kann, und die Fehlerkontrolle.

Die Netzwerkschicht ist bei der QOS-A im wesentlichen eine ATM-Schicht, die eine möglichst gute Ressourcenauslastung unter Garantie der Anwendungsbedürfnisse leisten soll. Dazu wird die Information zur Verpflichtung des Dienstbringers verwendet. Man unterscheidet eine deterministische Dienstbringung, die Qualitäten genau einhält, eine statistische Dienstbringung, bei der die Qualität abhängig ist vom Netzwerkzustand, und eine bestmögliche Dienstbringung. Der wichtigste QOS-Unterstützungsmechanismus der Netzwerkschicht ist ein Algorithmus, der Verbindungen abhängig vom Datenaufkommen Netzwerkressourcen zuteilt.

6.2 QOS-Verhandlung, Monitoring und Policing

Über die QOS-Schnittstellen erhält ein Dienstinutzer die Möglichkeit obere und untere Grenzen für akzeptierbare Dienstqualität anzugeben. Der Dienstbringer versucht dann die Dienstqualität so nah an der oberen Grenze wie möglich zu halten. Außerdem können vom Dienstinutzer sogenannte Abschwächungsbedingungen angegeben werden, entlang denen der Erbringer die Dienstqualität einer Verbindung im Bedarfsfall abschwächen kann. Der eigentliche Verhandlungsvorgang läßt sich in 2 Phasen aufgliedern: In einer Vorwärtsphase wird entlang des Weges von der Quelle zur Senke von allen dazwischenliegenden Netzwerknoten- bzw. Endsystemschichten die Verfügbarkeit der geforderten Ressourcen mit einem Zugangskontrollmechanismus überprüft. Sind ausreichen Ressourcen vorhanden, werden alle verfügbaren Ressourcen für die Verbindung allokiert. An der Senke angekommen, wird erstens überprüft, ob die allokierten Ressourcen für die zu erbringende Dienstqualität ausreichen, und zweitens die möglicherweise zu viel reservierten Ressourcen berechnet. In der Rückwärtsphase werden dann von der Senke zur Quelle von allen dazwischenliegenden Schichten die überflüssigen Ressourcen wieder freigegeben. Ergebnis des Verhandlungsvorgangs ist ein Dienstvertrag, der auf allen Schichten eine entsprechende Repräsentation hat und sowohl vom Erbringer als auch vom Nutzer einzuhalten ist.

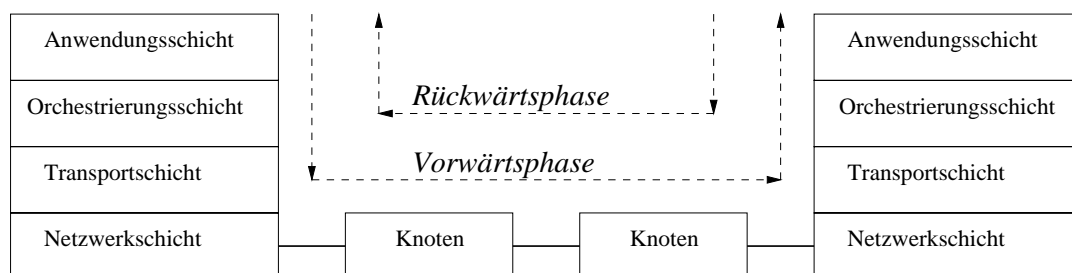


Abbildung 16. Verhandlungsvorgehensweise von QOS-A

In den oberen drei Schichten wird durch Monitorfunktionen die tatsächliche Dienstqualität überprüft und falls nötig und möglich an den Dienstvertrag angepaßt. In der Anwendungsschicht wird die Qualität der Ende-zu-Ende-Kommunikation und ihre Ursachen im Kommunikations- und Endsystem festgestellt. In der Orchestrierungsschicht

werden die Beziehungen zwischen den Medienströmen und gewisse Qualitätsmerkmale einzelner Ströme überprüft. Die Transportschicht kontrolliert die Qualität der Ende-zu-Ende-Kommunikation hinsichtlich der Transportverbindung, untersucht also die Netzwerkparameter unabhängig von den dem Endsysteem zuzuordnenden Ressourcen.

In der Netzwerkschicht wird durch Policing-Mechanismen überprüft, ob der Dienstbenutzer eine ausgehandelte Datenrate auch einhält. So wird zum Beispiel festgestellt, ob eine Anwendung mehr Bandbreite benötigt, als ihr zusteht. Falls dies der Fall ist, wird es der Anwendung angezeigt, die ihre Übertragungsrate dann korrigieren sollte.

7 Vergleich der drei Ansätze

Alle drei Ansätze bieten QOS-Management für multimediale Anwendungen im Endsysteembereich. QOS-A bietet zusätzlich ein QOS-Management, das sich über alle beteiligten Netzwerkknoten erstreckt. Dagegen sind der QOS-Broker und AQOSM auf ein adäquates Netzwerkressourcenmanagement angewiesen. Der QOS-Broker und AQOSM sind durch eine einzige Instanz realisiert, die alle Aufgaben übernimmt und dem Kommunikationssystem hinzugefügt wird. QOS-A sieht dagegen einen speziellen, auf die Bedürfnisse multimedialer Anwendungen zugeschnittenen Protokollturn vor, der implementiert werden muß. Hier gibt es auf allen Schichten Managementinstanzen. Während der QOS-Broker schichtenübergreifend ausgelegt ist, gibt es bei der QOS-A in jeder Schicht eine Managementinstanz. Alle drei Ansätze erlauben multimedialen Anwendungen die Spezifikation benötigter Übertragungsdienstqualität. Alle drei Ansätze sehen Mechanismen zur Aushandlung dieser Qualitätsparameter vor. QOS-A kann garantieren, daß ausgehandelte Dienstverträge auch eingehalten werden. Der QOS-Broker und AQOSM sind darauf angewiesen, daß der Transportdienst die von ihnen Ausgehandelten Qualitätsparameter auch einhält. Alle drei Ansätze ermöglichen die dynamische Anpassung von Qualitätsparametern an veränderte Anwendungsbedürfnisse oder Diensterbringerzustände. Außerdem unterstützt jeder Ansatz eine Synchronisation unterschiedlicher Medienströme. Nur der QOS-Broker und QOS-A unterstützen Multicast-Verbindungen und nur QOS-A bietet Policing-Funktionen.

Der QOS-Broker und AQOSM haben gegenüber QOS-A den großen Vorteil, daß sie aus nur einer Instanz bestehen die gängigen Kommunikationssystemen hinzugefügt werden kann. Dies ist dadurch möglich, daß sie ihr QOS-Management auf den Endsysteembereich beschränken. Dies führt allerdings auch dazu, daß sie auf ein adäquates Netzwerkressourcenmanagement angewiesen sind und bei der Garantie von Dienstqualitäten vom Transportdienst abhängig sind. Der Schwerpunkt beider Architekturen liegt im Verhandlungsbereich — beide bieten ausgereifte Protokolle, die QOS-Verhandlungen unterstützen — und in der Verwaltung der Endsysteemressourcen. Der QOS-Broker hat gegenüber AQOSM den Vorteil, daß er auch Multicast-Verbindungen unterstützt. In einigen Anwendungsbereichen von multimedialen Systemen, beispielsweise bei Videokonferenzen oder Unterstützungstools für Gruppenarbeit, sind solche Multicast-Verbindungen von einiger Bedeutung.

QOS-A bietet die umfangreichste Lösung an und kann nahezu allen Anforderungen heu-

tiger multimedialer Anwendungen gerecht werden. Allerdings sieht QOS-A hierfür ein vollständig neu implementiertes, speziell auf den multimedialen Bereich ausgerichtetes Kommunikationssystem vor, was sich durchaus als nachteilig erweisen kann. Außerdem wird bei der QOS-A kein Aushandlungsprotokoll spezifiziert. QOS-A ist die einzige Architektur, die Dienstqualitäten autonom garantieren kann. Der Schwerpunkt liegt hier deshalb eher auf speziellen Transport- und Netzwerkdiensten, die solche Garantien ermöglichen, und weniger im Verhandlungsbereich.

	QOS-Broker	AQOSM	QOS-A
QOS-Management im			
Endsystembereich	Ja	Ja	Ja
Netzwerkbereich	Nein	Nein	Ja
Architektur	eine Instanz	eine Instanz	mehrere Instanzen
QOS-Spezifikation auf Anwendungsebene	Ja	Ja	Ja
QOS-Aushandlung	Ja	Ja	Ja
Garantierte Dienstleistung	Anhängig vom NW-Rourcenmanager	Abhängig von den Ressourcenmanagern	Ja
dynamische Neuverhandlung	Ja	Ja	Ja
Synchronisation von Medienströmen	Ja	Ja	Ja
Unterstützung von Multicastverbindungen	Ja	Nein	Ja

Abbildung 17. Vergleich der drei Ansätze

8 Zusammenfassung und Ausblick

Im Rahmen dieser Ausarbeitung wurden zuerst die Anforderungen von multimedialen Anwendungen an Telekommunikationsdienste vorgestellt. Qualitätsparameter als Hilfsmittel zur Spezifikation dieser Anforderungen auf den unterschiedlichen Systemebenen wurden untersucht. Die beteiligten Multimedia-, Betriebssystem- und Netzwerkressourcen wurden genannt und Konzepte des Ressourcenmanagements bei multimedialen Anwendungen betrachtet. Es wurden die drei neue Ansätze vorgestellt, die im Hinblick auf die Unterstützung von Diensten für multimediale Anwendungen entwickelt wurden: der QOS-Broker, AQOSM und QOS-A. Beim QOS-Broker und bei AQOSM handelt es sich um Endsystem-QOS-Manager. Sie sehen ihren Schwerpunkt im lokalen Res-

sourcenmanagement und in der Aushandlung von Dienstqualitäten. QOS-A bietet einen Programmierrahmen, der die Implementierung eines Protokollturms ermöglicht, welcher nahezu allen Anforderungen multimedialer Anwendungen gerecht wird. Im abschließenden Kapitel wurden die drei Ansätze miteinander verglichen und die Vor- und Nachteile der einzelnen Architekturen herausgestellt.

Station-Management in FDDI-Netzen

Thomas Merkle

Kurzfassung

Neuartige Multimedia-Anwendungen benötigen häufig synchrone Datenströme. Die FDDI-Spezifikation läßt sowohl asynchronen als auch synchronen Frame-Verkehr zu und eignet sich damit hervorragend für diese Art von Anwendungen. Das FDDI verwendet eine Token Ring-Architektur mit Glasfasern als Übertragungsmedien. FDDI unterstützt mehrere hundert Stationen, die sich über mehrere Kilometer ausdehnen können.

Das Station Management - das heißt SMT - definiert den lokalen Teil des Managements eines FDDI-Rings. Dazu gehört auch die ordnungsgemäße Funktion einer Station im FDDI-Ring. SMT regelt auf einer FDDI-Station das Zusammenwirken und die Überwachung der Prozesse in den FDDI-Komponenten (PMD, PHY, MAC). Die Grundlage des SMT bilden sogenannte Managed Objects (MO), das sind abstrakte Sichtweisen von Netzwerkressourcen, die zur Steuerung des Netzwerks definiert werden.

1 Einleitung

In diesem Beitrag soll anhand der FDDI-Spezifikation ein Überblick über die SMT-Dienste vorgestellt werden. Diese Arbeit basiert auf dem Draft Proposed American National Standard vom 16. Dezember 1992 (Revision 7.2c) [MR92]. Im ersten Teil wird das Zusammenwirken der unterschiedlichen FDDI-Komponenten beschrieben. Darauf aufbauend werden im zweiten Teil die genauen Funktionen der einzelnen Komponenten des SMT gezeigt.

Das Station Management in FDDI-Netzen wird benötigt, um die Ressourcen in diesem Hochgeschwindigkeitsnetz zu verwalten. Es ist ausgelegt für 100 Mbps mit einer Token-Ring-Architektur auf Glasfasermedien.

2 Grundlagen des SMT

2.1 Einordnung in das ISO/OSI-Schichtenmodell

Das Station-Management operiert auf den Schichten 1 bis 2a (siehe Abb. 18) des ISO/OSI-Schichtenmodells und stellt Dienste für das Management dieser Protokoll-Schichten zur Verfügung.

2.2 Aufgaben des SMT

Aufgabe des SMT ist es, auf einer FDDI-Station das Zusammenwirken der FDDI-Komponenten (PMD, PHY, MAC) zu regeln und zu überwachen. SMT-Dienste werden bereitgestellt für folgende Aufgaben:

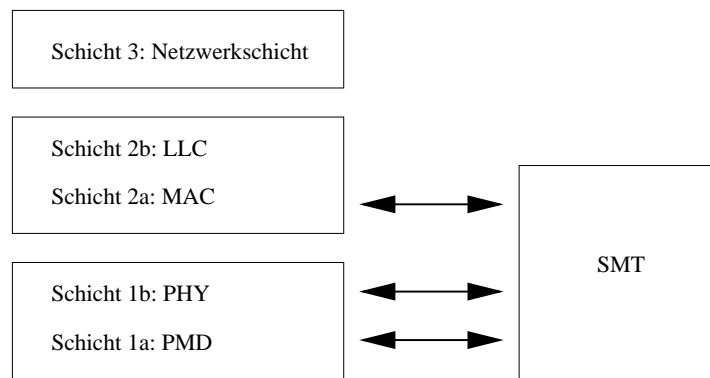


Abbildung 18. Einordnung des SMT in das ISO/OSI-Schichtenmodell.

- Initialisieren von Stationen
- Einfügen und Entfernen von Stationen
- Isolieren und Beheben von Fehlern
- Sammeln von Statistiken
- Kommunikation mit der Netzmanagement-Station

Bevor die einzelnen Punkte näher erläutert werden, folgt im nächsten Abschnitt eine kurze Übersicht über das zugrundeliegende Netzwerkmanagementmodell.

2.3 Das Netzwerkmanagementmodell

Das Modell einer Komponente im Sinne des Netzwerkmanagement wird i. allg. Managed Object (MO) genannt. Es stellt alle Management-relevanten Informationen über die zu verwaltende Komponente zur Verfügung. Ein MO wird definiert durch:

- Attribute: Informationen werden im Objekt in Attributen abgelegt, auf die über Managementoperationen zugegriffen werden kann. Attribute haben einen bestimmten Datentyp, es sind auch mehrdimensionale Strukturen und proprietäre Typen (Adressen, Zähler, ...) möglich.
- Operationen: Der Zugriff auf ein MO wird durch Operationen ermöglicht, die das Objekt an seiner Schnittstelle anbietet. Mögliche Operationen sind: Lesen und Schreiben von Attributen, Erzeugen und Löschen von MO's und MO-spezifische, in Abhängigkeit der modellierten Komponente.
- Meldungen: Antworten auf eine Managementoperation oder selbsttätige Meldung nach Eintritt eines Ereignisses wie Fehler oder Ausnahmesituation (z.B. Zählerwert).
- Verhalten: Der Zugriff auf MO's kann genormt werden, jedoch nicht die Beziehung zwischen dem MO und der modellierten Komponente. Dennoch hat eine Änderung im MO Auswirkungen auf die Komponente und umgekehrt.

Die Managed Objects werden durch Standardisierungsgremien registriert und fortlaufend nummeriert. Diese Nummer wird zur eindeutigen Identifizierung verwendet.

3 Die Station Management-Objekte

3.1 Beziehungen der SMT-Management-Objekte zueinander

Alle Informationen der Management Objects einer zu verwaltenden Einheit werden in der Management Information Base (MIB) gesammelt. Die für FDDI definierten MO-Gruppen mit ihren zugeordneten Funktionsgruppen sind:

MO-Gruppe	zugehörige Komponente	Beispiel
SMT Object	SMT-Funktion	fddiSMTStationId
MAC Object	MAC-Schicht	fddiMACUpstreamNbr
PATH Object	logischer Datenpfad	fddiPATHRingLatency
PORT Object	PHY- und PMD-Komponenten	fddiPORTCurrentPath

Das Station Management wird weiter in drei Gruppen unterteilt:

- CMT: Connection Management,
- RMT: Ring Management und
- SMT Frames and Services.

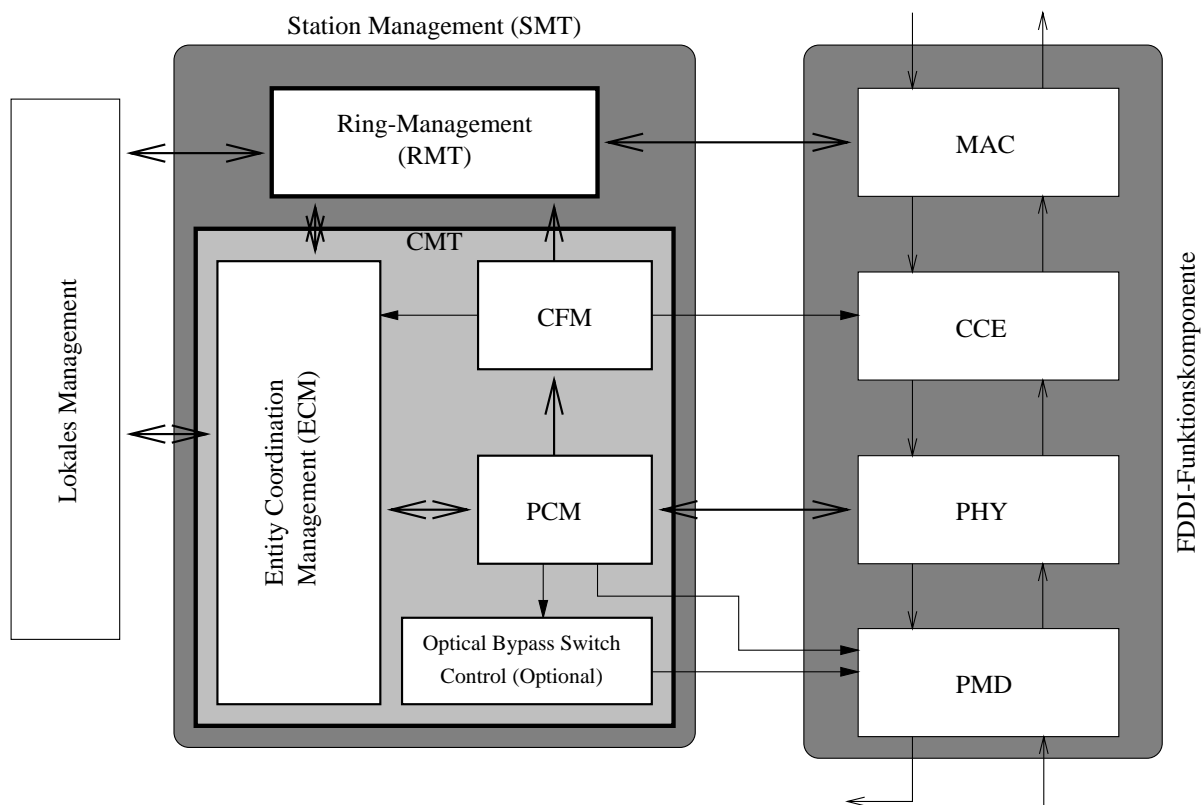
In der Abb. 19 werden die Beziehungen zwischen den Komponenten der FDDI-Station dargestellt.

3.2 Verbindungsmanagement CMT

Das CMT verwaltet hauptsächlich die PHY-Komponenten und ihre Verbindungen. Es sorgt im wesentlichen für:

- Initialisierung der physikalischen Verbindungen,
- Fehlererkennung,
- Rekonfiguration,
- Verbindungstest,
- Anzeige verfügbarer Verbindungen und
- Steuerung des optischen Bypass-Schalters.

Das CMT unterteilt sich weiter in das Entity Coordination Management (ECM), das die PHY's und den Optical Bypass Switch steuert, das PCM und CFM. Die Hauptaufgabe des Physical Connection Managements (PCM) ist die Initialisierung einer Verbindung zwischen einem lokalen und einem benachbarten Port. Das Configuration Management (CFM) steuert das Configuration Control Element (CCE), das seinerseits für die Konfiguration von MAC und PHY in dieser Station verantwortlich ist.



CCE: Configuration Control Element
 CFM: Configuration Management
 CMT: Connection Management
 PCM: Physical Connection Management

Abbildung 19. Interaktion der Management-Komponenten.

3.3 Ring-Management RMT

Das RMT verwaltet und überwacht die MAC-Komponenten und ihre logische Anbindung an den FDDI-Ring. Als weitere Tätigkeit kommt die Überwachung des FDDI-Rings und Fehlerbehebung z.B. bei Ringbruch hinzu.

Zum Management der MAC-Komponenten gehört das Ändern einer MAC-Adresse in eine eindeutige Adresse und das Entfernen der MAC mit duplizierter Adresse vom Ring. Um dies alles erledigen zu können, werden Status-Informationen von MAC und CMT an das RMT weitergeleitet und auch anderen SMT's in anderen FDDI-Stationen am Ring zur Verfügung gestellt. Zur Unterstützung dieser Aufgabe signalisieren dem RMT zahlreiche Timer und Flags Ereignisse innerhalb des MAC-Protokolls.

3.4 SMT-Frames

Damit das SMT Managementinformationen zur Erbringung von Management-Dienstleistungen mit anderen SMT-Einheiten austauschen kann, werden SMT-Frames verwendet. Die SMT-Frames sind in den FDDI-MAC-Frames enthalten (siehe Abb. 20).

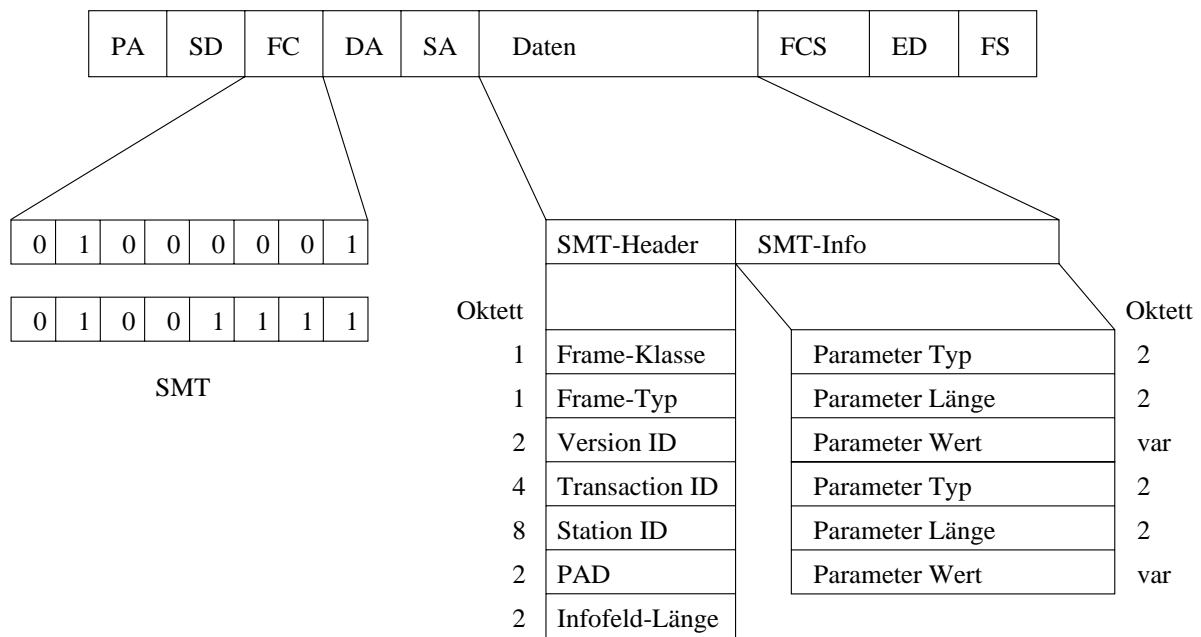


Abbildung 20. SMT Frameformat.

3.5 SMT-Dienste

- Neighbor Notification Process: Mit diesem Dienst kann man die MAC-Adresse des unmittelbaren Nachbarn lernen. Er wird auch dazu verwendet die Erkennung von duplizierten Adressen des RMT zu ergänzen, Verbindungen zum unmittelbaren Nachbarn zu prüfen, Ringfehler zu entdecken und Topologie- und Konfigurationsabbildungen für ein Netzmanagement höherer Ebene zu unterstützen.
- Status Report Protocol: Wird zur Verbreitung von Zustandsinformationen, die für das Netzmanagement nützlich sind, verwendet.
- Parameter Management Protocol: Damit kann die lokale RMT die MO's der SMT's anderer Stationen lesen und ändern.
- Request Denied Frames: Zur Erkennung von SMT-Stationen, die nicht der aktuellsten SMT-Version entsprechen.

4 Station Management

4.1 SMT-Dienste

SMT-to-MAC-Dienste Dieses Interface wird von der SMT-Schicht verwendet, um die MAC-Einheit zu überwachen und zu steuern.

- Initialize-Protocol: SMT (re-)konfiguriert die MAC-Einheit neu.
- Control: Wird dazu verwendet, einen Reset durchzuführen oder einen Beacon-Rahmen zu senden. (weitere: Status mitteilen, Zähler zurücksetzen, gegebenen Zustand erkannt, falsche FCS zu Testzwecken senden).

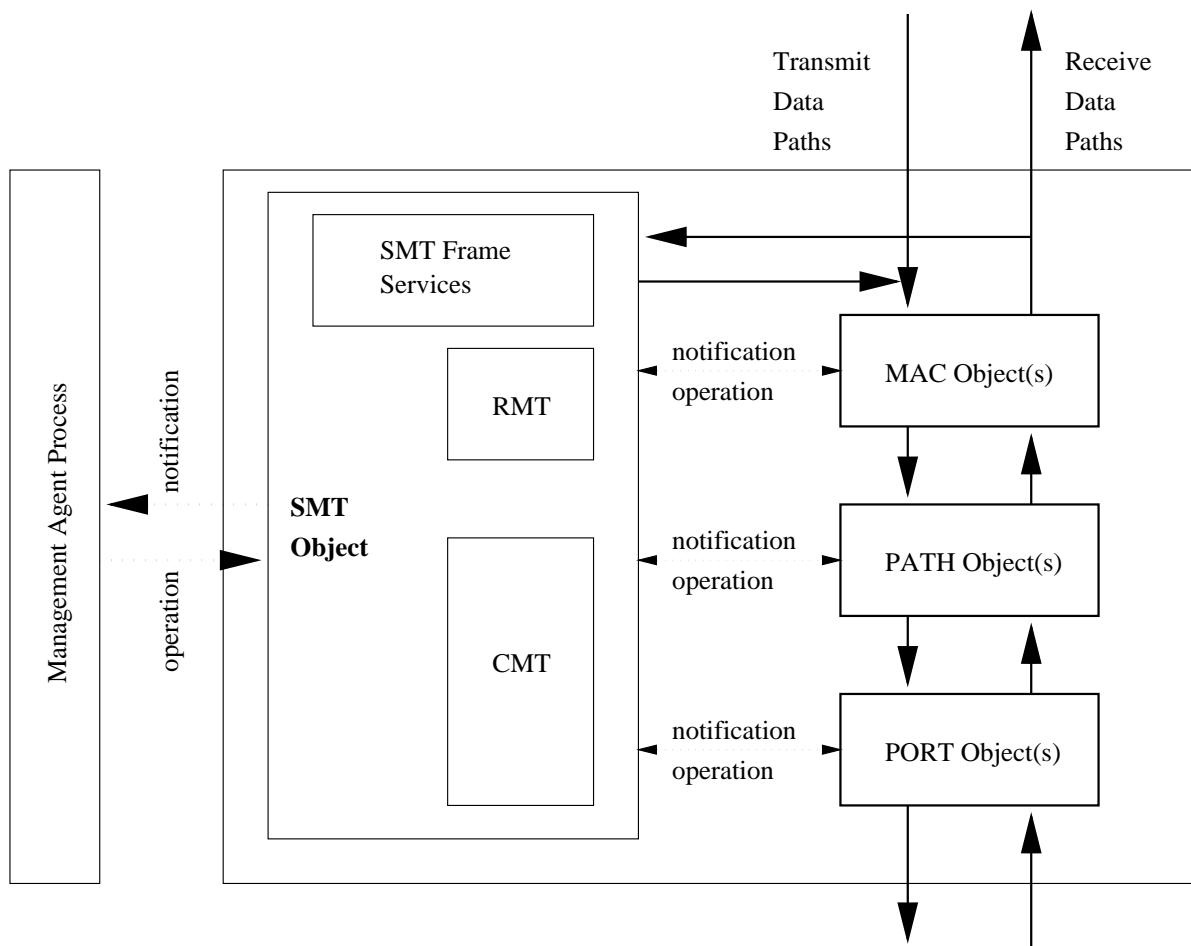


Abbildung 21. SMT Management Model

- Status: MAC meldet erkannte Fehler an SMT.
- UnitData.request: Senden einer SMT Service Data Unit zu einem oder mehreren SMT-Nachbarn oder zur Erzeugung eines Tokens.
- UnitData.indication: Meldung durch die MAC, daß ein an diese Station adressierter Rahmen angekommen ist.
- UnitData-Status.indication: Antwort auf UnitData.request.
- Token: Abfangen eines Tokens der angegebenen Klasse durch die MAC, damit die SMT wichtige, zeitkritische Daten senden kann.
- Senden von SMT-Frames.

SMT-to-PHY-Dienste Dieses Interface erlaubt dem SMT die Kontrolle über die PHY-Schicht.

- Line-State.request: Senden von Strömen von QUIET, IDLE, HALT, MASTER-Symbolen zur Stationseinbindung oder -entfernung.
- Status.indication: Benachrichtigung der SMT über einen der oben genannte Zustände.

- Control.request: Senden von Reset, Loopback-Start (und -Ende).

SMT-to-PMD-Dienste Dieses Interface verwendet die SMT, um die Funktion der PMD-Schicht zu steuern. Diese Dienste haben Vorrang vor SMT-to-PHY-Diensten.

- Control.request: Bei Transmit-disable wird die logische Null und bei -enable werden die PM-UnitData-Daten übertragen.
- bypass.request: insert bedeutet in Ring einfügen, d.h. Daten werden empfangen (bypass abgeschaltet) und deinsert, daß die Daten vom Eingang direkt zum Ausgang weitergeleitet werden (Bypass-Modus).
- Signal.indication: Liefert der SMT den empfangenen Signalpegel.

4.2 SMT-Frames

Ein SMT-Rahmen ist ein FDDI-Mechanismus, der für Peer-to-Peer- und Schichtenmanagement verwendet wird. Die Rahmen bestehen aus den drei Hauptkomponenten MAC Header, SMT Header und SMT Info (siehe Abb. 20). Im MAC Header wird das FC-Feld bei einem SMT Info-Rahmen auf 0100 0001 bzw. 0100 1111 bei einem NSA (Next Station Addressing) -Rahmen gesetzt. Quell- und Zieladresse stehen ebenfalls im MAC Header. Der NSA-FC-Wert wird für Neighbour Information Frames (NIF, request- und announcement-Frames) verwendet; der SMT-Info-FC-Wert dafür bei allen Rahmen, die nicht für das Neighbour Notification Protocol verwendet werden.

Der SMT Header enthält folgende Felder:

Frame Class	Funktion des Rahmens (NSA oder SMT-Info)
Frame Type	Bekanntgabe-, Anforderung- oder Antwort-Rahmen
Version ID	legt die Struktur des Info-Felds fest
Transaction ID	zur Zuordnung von Fragen und Antworten
Station ID	enthält implementierungsabhängigen Teil und 64-bit Universaladresse (eindeutig)
Pad	um 32-bit-Ausrichtung der Parameter im Info-Feld zu erreichen
Info Field Length	Länge des Info-Felds. Maximal 4458 Oktette

Die maximale Länge des Info-Felds berechnet sich folgendermaßen:

Max Frame - MAC Header - Min MAC Trailer - SMT Header =
 4500 - 16 - 6 - 20 = 4458 (laut Standard trotzdem ≤ 4352 bytes).

Es gibt zehn definierte Rahmenklassen:

- Neighbor Information Frames (NIF): Zur periodischen Bekanntgabe der eigenen Adresse und einer allgemeinen Stationsbeschreibung.
- Status Information Frames (SIF): Werden vom SIF-Protokoll verwendet, um Stationskonfigurations- und -betriebsinformationen bekanntzugeben und zu erfahren (potenziell zur Fehlerisolierung und Sammlung statistischer Informationen).

- Echo Frames (ECF): Für das SMT-zu-SMT-Loopback-Testen (zur Bestätigung, daß Port, MAC und SMT einer Station funktionieren und um datenabhängige Fehler zu finden).
- Resource Allocation Frames (RAF, optional): Zur Belegung von Ressourcen, im Moment nur die Belegung synchroner Bandbreite definiert.
- Request Denied Frames (RDF): Wird zur Meldung von request frame Format- und Protokoll-Fehlern verwendet.
- Extended Service Frames (ESF, optional): Zur Erweiterung und zum Testen von neuen SMT-Diensten.
- Status Report Frames (SRF): Wird von Stationen zur Bekanntgabe von Stations-Informationen verwendet, die möglicherweise für einen das Status Report Protokoll verwendenden FDDI-Ring-Manager von Interesse sind.
- Parameter Management Frames (PMF): Diese Rahmen stellen die Möglichkeit des entfernten Zugriffs auf alle in der Management Information Base verwalteten Stationsattribute über das Parameter Management Protokoll zur Verfügung.

4.3 Frame-Based Management Protocols

SMT stellt höheren Management-Schichten rahmenbasierte Dienste zur Verfügung. Dies ermöglicht jenen, Informationen zu sammeln und die Kontrolle über das angeschlossene FDDI-Netz auszuüben. Genauer ermöglichen diese Protokolle, Netzstatistiken zu sammeln; Fehler zu erkennen, zu isolieren und zu beheben; und die Konfiguration den Anwendungs- und Verbindungswünschen anzupassen.

Außer den Neighbour Information Frames (periodisch durch das Neighbour Notification Protocol erzeugt) und den Resource Allocation Frames, werden request-Rahmen nur auf Anforderung eines Management-Prozesses generiert.

Neighbor Notification Protocol: Dieses Protokoll dient zur Feststellung der MAC-Adressen der logischen Nachbarn davor- und dahinter im Ring sowie dem Feststellen doppelter Adressen. Es arbeitet nur mit einem operablen logischen Ring (es ergänzt die Erkennung doppelter Adressen des RMT, das nur mit einem nicht-operablen Ring arbeitet). Zur Erfüllung dieser Aufgabe verwendet es die NIF und sendet sie periodisch (request-response). Ist längere Zeit kein Verkehr auf dem Netz, sendet es periodisch Handshake-Rahmen zum Test der lokalen MAC Sende- und Empfangsübertragungswege.

Status Report Protocol: Eine Station gibt periodisch Statusinformationen, die für die Verwaltung eines FDDI-Rings nützlich sein können, in SR Frames bekannt.

Zustände und Ereignisse werden - ratenmindernd - periodisch an die SRF Multicast-Adresse gesendet. Ändert sich der Zustand nicht, wird er nur noch in geometrisch steigenden Abständen gesendet. Solche Zustände können beispielsweise über ein bestimmtes Limit gestiegene Rahmenfehler oder doppelte Adressen sein. Beispiel eines Ereignisses wäre eine wechselnde MAC-Nachbaradresse.

Parameter Management Protocol: Das PMP wird zur Fernverwaltung von Stationsattributen - die in der Management Information Base stehen - verwendet. Zur Konsistenzwahrung von Set-Befehlen können die Requester einen Konsistenzkontrollmechanismus verwenden, der überlappendes Setzen von Attributen durch verschiedene Requester verhindert. Der Requester erhält aber auf jeden Fall (Set- und Get-Befehl) von der Zielstation (dem Responder) ein Antwortpaket mit der Bestätigung oder der Ablehnung eines Befehls. Des weiteren unterstützt das PMP eine Zugangskontrolle für SET-Befehle auf Basis eines Authorisations-Parameters (implementationsabhängig).

Station Status Polling: Dieses Request/Response-Protokoll verwendet die SI Frames. Es ermöglicht, Konfigurationsdaten wie Verbindungs- und Konfigurationsparameter und statistische Informationen von einer entfernten Station abzufragen.

Echo Protocol: Verwendet den Echo Frame, um SMT-zu-SMT-Schleifen auf dem FDDI-Ring zu testen.

Synchronous Bandwidth Allocation: Das SBA-Protokoll unterstützt die Verwaltung der Zuteilung synchroner Bandbreite und verwendet die RA Frames. Das SBA-Protokoll unterstützt folgende Funktionen:

- Verwaltung der limitierten synchronen Bandbreite.
- Überwachung der belegten synchronen Bandbreite.
- Überwachung des Rings wegen Überbelegung der synchronen Bandbreite.
- Eingreifen bei Ringinstabilität wegen Überbelegung der synchronen Bandbreite.

Das SBA-Protokoll geht von einem zentralen Bandbreiten-Verwaltungsprozess (ZBV) aus und benutzt Anforderungs- und Antwort-Rahmen zwischen der SBA benutzenden Station und dem ZBV. Die Überwachung der belegten synchronen Bandbreite erfolgt durch den ZBV durch Abfragen jeder bandbreite-belegenden Station - durch Mithören der NI-Frames erfährt der Prozess diese - oder er merkt sich von jedem RAF die Rate und Adresse der Stationen.

Extended Service Protocol: Dieses Protokoll unterstützt erweiterte und den Test neuer auf SMT-Rahmen basierender Dienste.

4.4 CMT

Das CMT verwaltet das Ein- und Ausfügen von Ports (= PMD + PHY) in den FDDI-Ring und die Verbindung der PHY-Komponente zur MAC-Komponente. In dieser Funktion unterstützt es eine große Anzahl von verschiedenen physikalischen und logischen Topologien, denn das Anschließen an FDDI betrifft eben die Ports und die MAC-Komponente. CMT muß also das Anschließen an den Ring und den physikalischen Verbindungsaufbau zu den Ports anderer Stationen und Konzentratoren steuern. Hinzu kommt

die Verwaltung der stationsinternen Verbindungen zwischen den verschiedenen MAC- und PHY-Komponenten.

Zur Erfüllung dieser umfangreichen Aufgaben wird das CMT in weitere Komponenten unterteilt (siehe Abb. 19).

Die CMT-Komponenten arbeiten mit einer Vielzahl von Variablen, Flags, Signalen und Timern, um die Zustände der aktuellen Verbindungen (Links und Pfade) darzustellen:

- Line States: Der Quiet Line State (QLS) der PHY-Schicht wird zur Unterbrechung der Verbindung und zum Neustart der Verbindungsinitialisierungssequenz verwendet. Wichtig ist noch der Idle Line State, der zur Trennung von Informationsbits und zur Synchronisation verwendet wird.
- Link Confidence Test: Damit wird die Qualität der Verbindung getestet - ob sie also gut genug für Ring-Operationen ist. Hierzu wird als Minimalfunktionalität das Zählen von Verbindungsabschnittfehlern verlangt.
- Link Error Monitor: Der LEM untersucht die Link Error Rate einer aktiven Verbindung und meldet anderen SMT-Instanzen, daß physikalische Verbindungen nicht mehr gut genug sind. Diese Informationen können zu Fehlerdiagnosezwecken verwendet werden.
- Path Test: Diese Funktion wird dazu verwendet, festzustellen, ob das Problem des FDDI-Rings im lokalen Knoten liegt.
- Trace Function (wird durch RMT gestartet): Wird zur Lösung der stuck beacon-Bedingung auf FDDI-Ringen verwendet. Sie veranlaßt alle Stationen und Konzentratoren der vermuteten Fehlerdomain den Ring zu verlassen und in den Pfad-Test-Modus zu gehen.

Dazu kann die Qualität getestet werden und erst wenn die Fehlerrate gering genug ist, wird der Übertragungsabschnitt freigegeben. Der Pfad-Test ist in diesem Standard nicht definiert, wird aber zur Feststellung, ob Fehler im lokalen Knoten Ursache von Ring-Fehlern sind, benötigt. Zur Feststellung von Fehlern, die nicht nur einer Verbindung (link) zugeordnet werden können, gibt es den Trace-Prozess, der alle Stationen und Konzentratoren in der fehlerverdächtigten Domain zwingt, den Ring zu verlassen und in den Pfad-Test-Modus zu gehen.

Entity Coordination Management ECM Das Management des Medieninterfaces zum FDDI-Netzwerk - welches das Management aller Aktivitäten der Ports, die dieser physikalischen Anbindung zugeordnet sind, und die Kontrolle über den optischen Bypass-Schalter enthält - heißt ECM. Es gibt pro FDDI-Station nur eine einzige ECM-Instanz. Das ECM kontrolliert den optischen Bypass des PMD und signalisiert der PMD-Schicht, wann ein Medienzugriff möglich ist. Die PCM-Zustandsautomaten der Ports A und B (in Konzentratoren zusätzlich der Ports M) werden von dem ECM gestartet, wenn das optische Bypass fertig ist. Des weiteren koordiniert dieser Teil die Trace-Funktion innerhalb des Knotens, wozu das Verbreiten und Terminieren des Traces gehört. Nach

Beendigung des Traces initiiert das ECM den Pfad-Test, um den vermuteten Fehler zu lokalisieren. Der Pfad-Test soll Fehler in knoteninternen Pfaden und MACs finden.

Physical Connection Management PCM Zu den Aufgaben des PCM's gehört die Initialisierung und Beibehaltung einer minimalen Verbindung sowie Support-Unterstützung (Support-maintenance: dadurch kann das CMT den Port der Nachbar-Station manuell in einen bekannten Zustand zwingen). Es initialisiert also die Verbindung zwischen benachbarten Ports und verwaltet die Signalisierung zwischen ihnen. Dazu wurde das PCM in zwei Teile gespalten:

- Die Zustandsmaschine
Sie enthält die Zustands- und Timinginformation und unterhält einen Signalisierungskanal.
- Der Pseudo-Code
spezifiziert die bits, die von der Zustandsmaschine signalisiert werden müssen und bearbeitet die bits der PCM-Einheit am anderen Ende.

Configuration Management CFM Zur Konfiguration der Ports und MACs in einem Knoten wird das CFM verwendet. Es wurde definiert für Single- und Dual-Attachment-Stationen und -Konzentratoren. Es beschreibt die interne Konfiguration der Ports und MACs einer Station (respektive Konzentration) und ist wegen der implementationsabhängigen Anzahl von Einheiten innerhalb eines Knotens verteilt. Stationen und Konzentrationen dürfen entweder einen Port A und B oder einen Port S haben, die zur Anbindung an den Doppelring oder einen M-Port eines Konzentration verwendet werden. (Eine FFDI-Station enthält mindestens eine MAC; ein Konzentration zusätzlich einen oder mehreren M-Ports zum Anschluß von Stationen bzw. Konzentrationen.)

Das CFM erlaubt für dual attachment stations / concentrators unterschiedliche Verbindungsmöglichkeiten auszuwählen, wenn der Knoten sich in einem Konzentrationen-Baum befindet.

Der Pfad repräsentiert die Elemente des logischen Rings, die durch die Station gehen. Der primäre Pfad - nach dem Stationswissen - repräsentiert also die Segmente des primären Pfads - analog der sekundäre. Der primäre Pfad sollte in allen Knoten vorhanden sein. Ein DAC oder eine DAS sollte aber immer einen zusätzlichen Pfad zwischen dem A:PH.UNITDATA.request und dem B:PH.UNITDATA.indication haben, um sie in Durchlaßkonfiguration verbinden zu können.

Für die Pfade gibt es mehrere Zustände wie isoliert, lokal, sekundär, primär, verbunden oder Durchgang, welche anzeigen, in welchen Pfad eine MAC oder ein Port gerade eingefügt ist.

Configuration Control Element CCE Die Verbindung von Ports zu anderen Ports innerhalb einer Station oder eines Konzentration wird durch ein abstraktes Configuration Control Element (CCE) für jede Art von Port und MAC definiert. Das CCE spezifiziert die Verbindungen, die eine standardkonforme Implementation erbringen sollte.

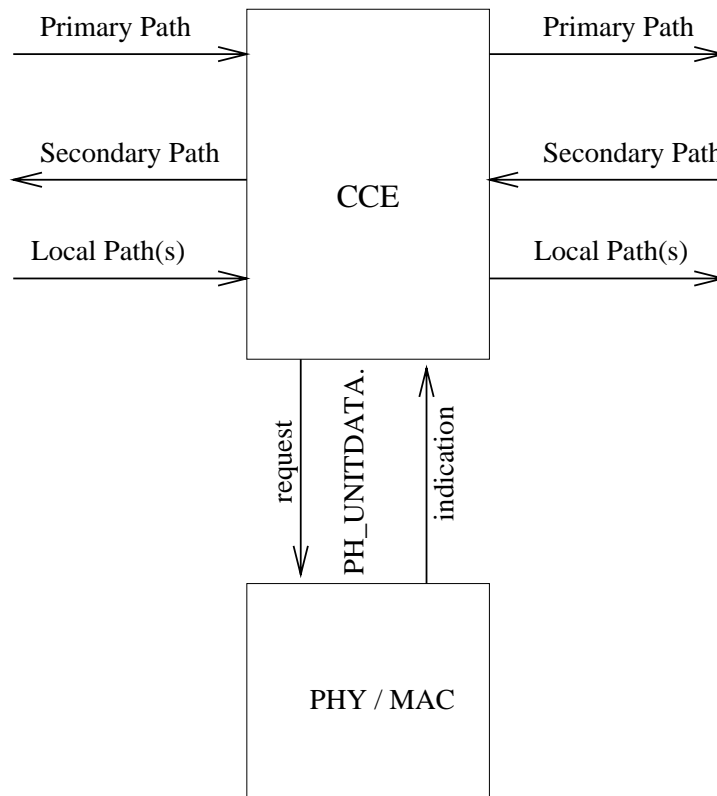


Abbildung 22. Allgemeines CCE-Interface für einen Port/MAC

In Abb. 22 wird ein allgemeines CCE gezeigt. Die Pfade gehen und kommen immer vom und zum nächsten CC-Element. Für alle A, B und S-Port-CCE's sowie mindestens eine MAC-CCE wird der Primärpfad benötigt. Der Sekundärpfad dagegen wird nur von allen A-Ports und in der Durchgangskonfiguration auch vom B-Port benötigt. Alle anderen Pfade sind optional und für zusätzliche Verbindungskonfigurationen vorgesehen.

Dem CCE muß es möglich sein, die Ressource für die benötigten Pfade zu konfigurieren oder isoliert zu bleiben. Im isolierten Zustand ist die Ressource mit keinem internen Datenpfad verbunden. Des weiteren sollten die CCE's der A- und B-Ports die Durchgangskonfiguration beherrschen. Der lokale Pfad seinerseits erlaubt die Verbindung eines gegebenen Ports zu einer MAC.

Zu den Konfigurationsmöglichkeiten zählen isolierte, lokale, gewrappte und durchgeschaltete Datenpfade je PHY und Pfad.

4.5 RMT

Das Ring Management erhält Statusinformationen von der MAC und dem Configuration Management und liefert den Status der MAC. Hierzu gehören auch das Feststellen der stuck beaconing Bedingung und das Initiieren der Trace-Funktion sowie die Auflösung von duplizierten Adressen, die den Ring nicht operabel werden lassen. Zusätzlich unterstützt das RMT Meldungen, wenn die MAC zur Datenübertragung bereit ist.

- Erkennung der Stuck Beacon-Bedingung,
- Initiierung der Trace-Funktion,

- Meldung über MAC-Bereitschaft,
- Feststellung doppelter Adressen,
- Auflösung doppelter Adressenprobleme, die Ringoperation verhindern und
- Beobachtung der Restricted Token.

Arbeitsweise

- Stuck Beacon Recovery: Nach Ablauf eines Timers und Bekanntgabe dieses Zustandes über den Ring, wird für das Connection Management das Trace-Propagation-Flag gesetzt.
- Duplicate Address Detection: Wenn My_Claim- oder My_Beacon-Pakete mit der eigenen Adresse nicht innerhalb von zweimal D_{\max} (Ringumlaufzeit) ankommen, ist das Paket definitiv von einer anderen Station mit derselben MAC-Adresse.
- Resolution of Duplicate Address Detection:
 - Entfernung aller von der MAC mit der doppelten Adresse abgesetzten PDU's und Änderung der MAC-Adresse in eine eindeutige Universaladresse.
 - Die MAC dahingehend konfigurieren, daß sie den Claim-Prozess verliert und die anderen duplizierten MAC's über die Duplizität benachrichtigt.
 - Entfernung der MAC mit der doppelten Adresse vom FDDI-Ring.
- Monitoring of restricted Token Dialogs: RMT beobachtet diese Dialoge, die ja normale asynchrone Rahmen und SMT-Rahmen verhindern. Dauern diese Dialoge zu lange, wird durch die RMT das Token recovery eingeleitet.

5 Zusammenfassung

Hier konnten nur die grundlegenden Komponenten des Station Management aufgezeigt werden, da sonst der Rahmen dieser Arbeit gesprengt worden wäre. Für ein detailliertes Aufzeigen der Interaktionen der SMT-Komponenten untereinander und mit den FDDI-Komponenten ist tiefergehendes Wissen der Funktionsweise der FDDI-Komponenten erforderlich.

Transportprotokolle mit Unterstützung von isochronem Datenverkehr

Roland Forberger

Kurzfassung

In dem vorliegenden Kapitel werden die grundlegenden Transportprotokolle TCP und TP4 besprochen, sowie das Realzeit-Transportprotokoll (RTP), die Tenet Real-Time Protocol Suite und das Xpress Transfer Protocol (XTP). Ausgehend von diesen fünf Transportprotokollen wird dann die Tauglichkeit der einzelnen Protokolle hinsichtlich der Verwendbarkeit bei isochronem Datenverkehr, wie er z.B. bei Audio- oder Video-Übertragungen anfällt, untersucht. Es wird dabei auf die einzelnen Stärken und Schwächen der vorliegenden Protokolle gezielt eingegangen, um diese anhand der Mechanismen gegenüberzustellen. Dabei zeigt sich, daß keines der Transportprotokolle vollkommen ist. Die Transportprotokolle TCP und TP4 bieten dabei die schlechteste (wenn überhaupt) Unterstützung hinsichtlich isochronem Datenverkehr. XTP bietet die umfassendsten Mechanismen an. Die verbleibenden beiden Transportprotokolle „genießen“ unterschiedliche Vor- und Nachteile hinsichtlich ihrer Verwendbarkeit bei isochronem Datenverkehr.

1 Einleitung

Der isochrome Datenverkehr, welcher vermehrt bei Multimedia-Anwendungen zum Tragen kommt, stellt völlig neue Anforderungen an die Transportprotokolle. So ist es nicht mehr vordergründig, daß die Daten unverfälscht ausgeliefert werden, sondern vielmehr das sog. Qualitätsgarantien für die jeweilige Verbindung eingehalten werden. Um diesen neuen Leistungsansprüchen bezüglich des Datentransports gerecht zu werden, müssen die Protokolle der Transportschicht entweder um die Zusicherung der jeweiligen Dienste erweitert werden oder aber völlig neue Konzepte ins Leben gerufen werden. Das nachfolgende Kapitel beschäftigt sich deshalb mit der Frage, welche existierenden Transportprotokolle überhaupt für isochronen Datenverkehr „tauglich“ sind. Dabei wird zuerst auf die grundlegenden Konzepte und Ziele der einzelnen Ansätze eingegangen, um diese dann in einem Vergleich, hinsichtlich der Stärken und Schwächen beim isochronen Datenverkehr, gegenüberzustellen. Eine abschließende Beurteilung der neuen Transportprotokollansätze für den isochronen Datenverkehr rundet dieses Kapitel ab.

2 Grundlegende Konzepte und Ziele einzelner Transportprotokolle

2.1 Das Transmission Control Protocol (TCP)

Das Transmission Control Protocol (kurz TCP) ist ein verbindungsorientiertes Transportprotokoll, welches vom amerikanischen Verteidigungsministerium (Department of

Defense) entwickelt wurde. Es sorgt neben der verbindungsorientierten Übertragung noch für deren Sicherung und überwacht die Reihenfolgentreue der Datenpakete. Um den gewünschten Dienst zu erbringen baut es auf Vermittlungsschichtebene auf das Internet Protocol (IP) auf, welches für die Wegwahl und die ungesicherte Übertragung der Daten sorgt [Krü95].

Die Hauptaufgaben von TCP sind:

- Multiplexen - mehrere Prozesse können gleichzeitig eine TCP Instanz benutzen.
- Verbindungsmanagement - Verbindungsaufbau, Aufrechterhaltung der Verbindung und Verbindungsabbau
- Datentransport (Vollduplex, Reihenfolgentreue, Flußkontrolle, Sicherheitsstufen, zeit-behaftete Daten, Prioritäten)
- Zusatzdienste - Unterstützung des schnellen Weiterreichens von wichtigen Daten
- Fehleraufzeichnung

Das TCP Paketformat beinhaltet dabei keine Felder, um irgendwelche Dienstgarantien zu geben. Lediglich ein Zeiger auf wichtige Daten innerhalb der Nutzdaten wird mittels Urgent Pointer ermöglicht [GM95]. Um dennoch gewisse Dienstmerkmale bereitzustellen, stützt TCP sich völlig auf IP, welches im „Type of Service“-Feld des IP-Paketformats Angaben über die Dienstqualität machen kann (Priorität, Verzögerung, Durchsatz und Zuverlässigkeit) [Krü95], [Krü95].

Die grundlegenden Einsatzbereiche, für welche die Internet-Protokollfamilie (TCP/UDP/IP) konzipiert wurden, beschränken sich hauptsächlich auf File-Transfer und eMail. Für aktuelle Anforderungen wie z.B. verteilte Anwendungen, synchronisierte multimediale Datenströme (Audio/Video) oder Echtzeitanwendungen ist die Internet-Protokollfamilie wegen fehlender bzw. unzureichender Mechanismen kaum geeignet, was in den nachfolgenden Abschnitten noch näher erläutert wird [Wea94].

2.2 Das ISO/OSI Transportprotokoll der Kategorie 4 (TP4)

Von ISO/OSI werden die Dienste der Vermittlungsschicht hinsichtlich ihrer Qualität in drei Kategorien (Typ A, B und C) unterteilt. Aufbauend auf diesen Vermittlungsschicht-Typen sieht die ISO/OSI fünf Transportprotokollklassen vor (Klasse 0 bis 4), von der die Klasse 4 (TP4) die aufwendigste ist [siehe Tabelle 4], da sie lediglich eine Vermittlungsschicht vom Typ C, welche bezüglich der Restfehlerrate und der angezeigten Fehler nicht den Dienstanprüchen genügt, benötigt. Somit besitzt TP4 neben den Eigenschaften der Transportprotokoll Klassen 3 auch noch Mechanismen zur Fehlererkennung und -behandlung.

TP4 nutzt für eine Transportverbindung entweder eine bereits bestehende Verbindung der Vermittlungsschicht oder baut vor dem Datentransfer eine neue Vermittlungsschichtverbindung auf. Für den Verbindungsaufbau wird dabei das Drei-Wege-Quittungsverfahren (Three-Way Handshake) verwendet, wodurch eine Verbindung erst nach Bestätigung

Transportprotokollklasse	Typ der Vermittlungsschicht	Funktionalität der einzelnen Klassen
Klasse 0	Typ A	Bildet Transportverbindung auf bestehende zuverlässige Netzwerkverbindung 1:1 ab.
Klasse 1	Typ B	Klasse 0 incl. Mechanismen zur Behandlung von Fehlern, die von der Vermittlungsschicht angezeigt werden.
Klasse 2	Typ A	Erweiterung der Klasse 0 um die Fähigkeit des Multiplexens.
Klasse 3	Typ B	Vereinigung der Mechanismen von Klasse 1 und 2.
Klasse 4	Typ C	Charakteristika der Klasse 3 mit zusätzlichen Mechanismen zur Fehlererkennung und -behebung.

Tabelle 4. ISO/OSI Transportprotokollklassen.

des Empfängers und erwidern des Senders (z.B. mittels Datenpaket) aufgebaut ist. Zudem können beim Verbindungsaufbau komplexe Aushandlungen der Dienstgüte mit der Partnerinstanz erfolgen, um somit eine Verbindung mit den geforderten Eigenschaften bereitzustellen [LKK93].

Die ISO/OSI hat für TP4 insgesamt 9 Transport-Protokolldateneinheiten (TPDU) spezifiziert. Dadurch sind die einzelnen Paketformate relativ kurz. Weiterhin unterscheidet TP4 normalen Datentransfer und einen Vorrangdatentransfer (Expedited Data), mit welchem kurze (maximal 16 Byte lange) Datenblöcke an den Empfänger gesendet werden können. Durch die Unterscheidung der Dienste/Ereignisse mittels spezieller TPDU's wird eine klare Strukturierung erzielt und somit eine Abgrenzung der einzelnen Dienste/Ereignisse untereinander (die Datenübertragung erfolgt dabei auch durch eine spezielle TSDU und nicht als Bytestrom) [Krü95].

Die ISO/OSI verfolgt mit der Standardisierung ihrer Transportprotokolle, wie auch die Internet-Protokollfamilie, hauptsächlich die Unterstützung von zuverlässigen, nicht zeitkritischen Anwendungen wie z.B. den File-Transfer [Wea94].

2.3 Das Realzeit-Transportprotokoll RTP

Die Funktionalität des Realzeit-Transportprotokoll RTP (Real-Time Transport Protocol) ist überwiegend, gemäß OSI-BRM, der Transportschicht zuzuordnen. Es stellt anwendungsorientierten Protokollen oder Anwendungen direkt Dienste zur Verfügung. Diese Dienste bzw. der Aufbau der Protokolldateneinheiten orientiert sich an den Anforderungen von interaktiven verteilten multimedialen Anwendungen. RTP bietet aber keine Garantien bezüglich Verzögerungen und Verzögerungsschwankungen. Zur Dienstleistung nutzt RTP je nach Anforderung weitere Dienste der Vermittlungsschicht oder auch der Transportschicht. Als minimale Voraussetzung ist dazu ein ungesicherter Ende-zu-Ende Transport notwendig, welcher z.B. von UDP, TCP oder TP4 erbracht werden kann. RTP kann aber auch direkt auf IP, ATM oder ST-II aufsetzen.

RTP nutzt folgende Dienste:

- Unzuverlässiger Ende-zu-Ende Transport (verlorengegangene, fehlerhafte oder reihenfolgevertauschte Pakete möglich),
- Multiplexen / Demultiplexen und
- Gruppenkommunikation (Multicast)

Andererseits stellt RTP Anwendungen folgende Funktionalitäten zur Verfügung:

- Ende-zu-Ende Transport für kontinuierliche Daten (verbindungsloser Datagramm-dienst) durch die Integration von Zeitinformation,
- Identifikation des Nutzdatentyps,
- das Kontrollprotokoll RTCP für Kontroll- und Identifikationsfunktion,
- Unterstützung von Zwischensystemen auf RTP-Ebene und
- Erweiterbarkeit und Flexibilität mittels RTP Anhangdateien.

RTP gibt dabei keine Garantien für:

- die pünktliche Auslieferung von Paketen
- andere Dienstgüte-Merkmale (z.B. Bandbreiten-Reservierung).

Konzeptionell wurde RTP als Transportprotokoll für multimediale Konferenzsysteme entwickelt, wobei es aber nicht auf diesen Bereich beschränkt ist. So bestehen Einsatzmöglichkeiten im Bereich verteilter Realzeit-Simulationen, entferntem Speicherzugriff für (multimediale) Daten, Informationsdiensten oder verteilten kooperativen Diensten. Diese Flexibilität wird durch Anhangdateien (sog. Profiles) ermöglicht, welche beim Verbindungsaufbau zwischen den Instanzen ausgehandelt werden müssen. Diese Anhangdateien legen dann die Interpretation der offenen Teile der Protokolldateneinheiten fest, so ist es z.B. möglich das Nutzdatenfeld unterschiedlich zu interpretieren (Audio, Video etc.).

Ergänzt wird RTP durch das Kontrollprotokoll RTCP (Real-Time Transport Control Protocol), welches folgende Funktionen bietet:

- Überwachung der Datenübertragung (z.B. Paketverlust und Laufzeiten) und
- Übermittlung von (Identifikations-) Information innerhalb einer Konferenz.

Das RTCP unterscheidet dabei fünf verschiedene Pakettypen. In der Spezifikation festgelegt wurden dabei Senderbericht, Empfängerbericht, Quellenbeschreibung und Abmeldung. Ein weiteres anwendungsspezifisches Paket ermöglicht zusätzlich flexible Pakete, die speziell für eine Anwendung entworfen werden können [Har95].

2.4 Die Tenet Real-Time Protocol Suite

Die Architektur der Tenet Suite beinhaltet, daß alle Schichten einer Netzwerkarchitektur in der Lage sein müssen, Dienstgarantien zu unterstützen. Folglich kann eine echtzeitfähige Verbindung nur dann entstehen, wenn die Sicherungsschicht in der Lage ist, die von der Vermittlungsschicht gewünschten Dienstqualitäten zu unterstützen (z.B. synchrones FDDI, ATM etc.). Echtzeitfähige Protokolle müssen also auf der Vermittlungs- und Transportschicht aufsetzen. Zu diesem Zweck stellt die Tenet Suite ein Netzwerkprotokoll - das Real-Time Internetwork Protocol (RTIP) - und zwei Transportprotokolle - das Real-Time Message Transport Protocol (RMTP) und das Continuous Media Transport Protocol (CMTP) - zur Verfügung [siehe Abbildung 23].

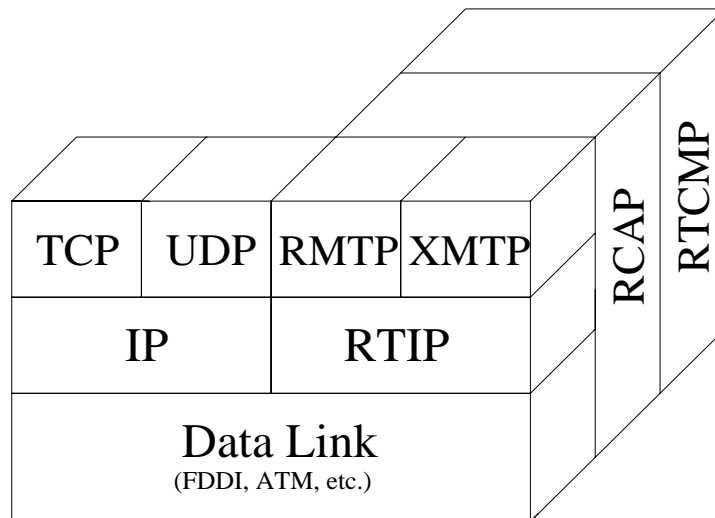


Abbildung 23. Die Tenet Real-Time Protocol Suite.

Um echtzeitfähige Verbindungen aufzubauen, werden vor dem eigentlichen Datentransfer zuerst die einzelnen Verbindungen etabliert. Dazu wird in jeder Teilstrecke versucht gemäß der geforderten Dienstqualität eine Teilverbindung bereitzustellen. Erst wenn in allen Teilstrecken eine Verbindung gemäß geforderter Dienstqualität erzielt werden konnte, kann der eigentliche Datentransfer beginnen. Zur Abwicklung des Verbindungsaufbaus dient das Real-Time Channel Administration Protocol (RCAP).

Durch die Unterstützung von Dienstgarantien entfällt die Notwendigkeit, daß unterschiedliche Echtzeitanwendungen spezielle Reservierungsprotokolle benötigen. Die beiden Transportprotokolle RMTP und CMTP unterscheiden sich dabei in der Art der Schnittstellen und deren Service. CMTP arbeitet periodisch, ist also zeitgesteuert, wobei RMTP auf dem send-receive Grundsatz aufbaut. Um ebenfalls „unkritische“ Daten zu befördern, koexistieren die Tenet Protokolle mit den Internet Protokollen. So kann ein Client „unkritische“ Daten auch mittels TCP oder UDP ausliefern und sich somit auf IP stützen. Um die verschiedenen Datenströme den einzelnen Protokollen zuordnen zu können, müssen zwischen der Vermittlungsschicht und den Netzwerktreibern zusätzliche Mechanismen zum Multiplexen/Demultiplexen eingefügt werden.

Ein weiteres Protokoll dient letztendlich zur Erkennung und Beseitigung von Fehlern

und trägt den Namen Real-Time Control Message Protocol (RTCMP). Es erstreckt sich wie das RCAP, welches ja zur Signalisierung dient, bis zur Transportschicht [BFMM94].

2.5 Das Xpress Transfer Protocol (XTP)

Das Xpress Transfer Protocol (XTP) umfaßt seit der Version 4.0 lediglich Funktionalität der Transportschicht (Version 3.6 umfaßte noch zusätzlich Funktionen der Vermittlungsschicht), was eine Änderung der Paketformate und der Multicast-Dienste nach sich zog. Die Protokollfunktionen blieben aber von der Änderung weitgehend verschont [SH95].

Durch den festen Paketaufbau und der festen Länge der einzelnen Felder ist das XTP gut für Hochgeschwindigkeitsnetze einsetzbar. Weiterhin ist das Protokoll durch eine Menge endlicher Automaten beschreibbar, wodurch eine Realisierung in Hardware und die Parallelisierbarkeit des Protokollablaufs möglich ist [GM95], [SW], [SH95].

Die Entwurfsprinzipien von XTP sind folgende:

- Trennung von Daten- und Kontrollfunktionen
- Kontrollfunktionen sind abhängig vom Dienst auswählbar und parametrisierbar
- Unterstützung von Multicast
- Unterstützung von Prioritäten
- Paketfelder fester Länge

Als Dienste sieht XTP einen Datagramm-Dienst vor, welcher entweder bestätigt oder unbestätigt (entspricht UDP-Dienst) sein kann. Einen Transaction-Dienst, welcher ein Datenaustausch in 3 Phasen für Unicast-Verbindungen bereitstellt. Einen zuverlässigen, verbindungsorientierten Unicast-Dienst (entspricht TCP-Dienst), sowie einen verbindungsorientierten Multicast-Dienst [SH95].

Zur Kommunikation zwischen den Endsystemen baut XTP zwei Unicast-Assoziationen auf, welche in einem Kontext (Context) münden, der die Statusinformationen der aktiven Verbindung zwischen den XTP-Endpunkten verwaltet. Alle Kontexte eines Endsystems werden innerhalb XTP von einem Context Manager bereitgestellt und verwaltet. Durch die Zuordnung von Kontexten zu jeder Verbindung können die Datenströme einzeln verwaltet werden und somit den jeweiligen Dienstansprüchen gerecht werden [SH95].

3 Übersicht von „gewünschten“ Protokollmechanismen für den isochronen Datenverkehr

Einführend erfolgt nun eine stichwortartige Aufzählung derjenigen Mechanismen, die für den isochronen Datenverkehr von (mehr oder weniger) entscheidender Bedeutung sind. Es soll dabei eher ein Überblick an gewünschten Mechanismen gegeben werden, als speziell ein bestimmtes Transportprotokoll dem jeweiligen Mechanismus zuzuordnen. Im anschließenden Abschnitt wird dann gezielt auf bestimmte Mechanismen und deren Verfügbarkeit in den vorgestellten Transportprotokollen eingegangen.

- Festes Paketformat, mit Auslegung auf beispielsweise 4-Byte-Grenzen für schnelle Verarbeitung (dadurch auch effiziente Realisierung in Hardware möglich).
- Adressierung mittels logischer/virtueller Kanäle, um einerseits eine schnelle Kanalzuordnung zu ermöglichen, andererseits den Protokolloverhead zu reduzieren.
- Flexible Mechanismen für unterschiedliche Dienstanforderungen (Möglichkeiten der Vorgabe des Jitters, des Durchsatzes, der Verzögerung, der Burstlänge, der Bitfehler-rate etc.).
- Synchronisation mehrerer Datenströme (z.B. für getrennte Übertragung von Video- und Audiokanälen).
- Bereitstellung von Puffern (zum „Auffangen“ eines Bursts) und Reservierung von Ressourcen über die einzelnen Teilstrecken der Verbindung.
- Unterstützung von verschiedenen Prioritäten.
- Verwerfen von fehlerhaften Paketen (keinesfalls „go-back-n“ Verfahren).
- Bedarfsgemäßes Quittungsverfahren für Datenpakete (z.B. Quittierung der empfangenen Datenpakete nur nach expliziter Aufforderung durch den Sender).
- Bereitstellen von einfachen, effizienten Mechanismen zur Flußkontrolle mit ausreichend großem Kredit.
- Unterstützung von Multicast-Verbindungen (Aufbau, Verwaltung und Abbau), falls z.B. eine Konferenzschaltung aufgebaut werden soll.
- Getrennter Kontrollfluß (Signalisierungskanal).
- Anpassungsmöglichkeiten des QoS (Quality of Service) einer bestehenden Verbindung, um beispielsweise kurzzeitig die Übertragungsqualität zu erhöhen.
- Einfacher Verbindungsauf- und Abbau (wichtig für schnelle, kurzlebige Verbindungen).

4 Stärken und Schwächen der Transportprotokolle bei isochronem Datenverkehr

4.1 Bereitstellung geforderter Dienstmerkmale

TCP unterstützt mittels IP nur rudimentäre Eigenschaften bezüglich Prioritäten oder Dienstanforderungen. Beides sind aber grundlegende Voraussetzungen für den Betrieb von Echtzeitanwendungen [Wea94]. Zudem werden die von IP unterstützten Prioritäten oftmals von den darunterliegenden Schichten nicht interpretiert. TP4 bietet beim Verbindungsaufbau zwar die Möglichkeit zur Aushandlung verschiedener Dienstparameter mit

der Gegenstelle, stellt aber keine Übertragungs-/Empfangspuffer bereit oder aber speziell für den isochronen Datenverkehr benötigte Dienstparameter (Jitter, Verzögerung etc.). XTP hingegen bietet umfassende Mechanismen bei der Prioritätenunterstützung oder des Verkehrsflusses (NOFLOW, RES, SREQ), welche sich auch beim Betrieb einer Verbindung teilweise noch ändern lassen [SH95]. Auch RTP bietet dank seines zusätzlichen Kontrollprotokolls (RTCP) Überwachungsmöglichkeiten der Datenübertragung bzw. Aushandlungsmöglichkeiten von Dienstmerkmalen für den zu unterstützenden Dienst [Har95]. Der Ansatz der Tenet Suite unterstützt durch das RCAP (Real-Time Channel Administration Protocol) komplexe Funktionen zur Reservierung der geforderten Dienstmerkmale [BFMM94].

4.2 Unterstützung von Multicast-Verbindungen

Die Internet-Protokollfamilie mit TCP/UDP und das ISO/OSI TP4 basieren auf einer Punkt-zu-Punkt Verbindung. Eine Multicast-Verbindung können Sie somit nur durch einzelne Unicast-Verbindungen [siehe Abbildung 24] bewerkstelligen, da grundlegende Konzepte für eine Multicast-Verbindung gänzlich fehlen (lediglich IP hat die Möglichkeit einzelne Datagramme über eine Klasse-D Adresse an mehrere Empfänger zu verschicken) [Wea94]. Die anderen vorgestellten Transportprotokolle wie XTP oder RTP unterstützen von sich aus Multicast-Verbindungen und erleichtern somit dem Umgang selbiger [siehe Abbildung 25]. Will ein Sender beispielsweise unter Verwendung von TCP zehn Empfänger erreichen, so muß dieser für jeden Empfänger eine unabhängige Unicast-Verbindung bereitstellen und verwalten.

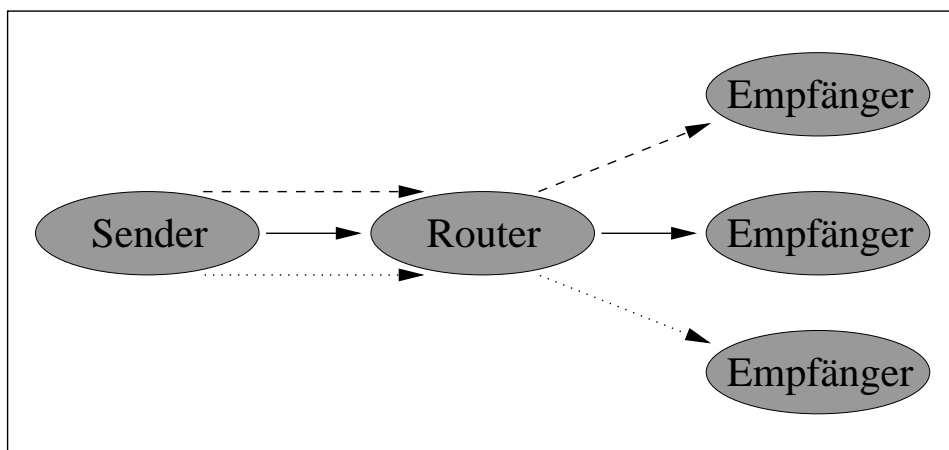


Abbildung 24. Ohne Multicast-Unterstützung.

4.3 Behandlung von Fehlern

TCP und TP4 verwenden beim Erkennen eines Fehlers das „go-back-n“ Verfahren [Wea94], wodurch alle nachfolgenden Pakete ebenfalls verworfen werden. Bei einem Burst mit entsprechender Fenstergröße der zu sendenden Pakete, kann der Durchsatz erheblich zurückgehen, ganz abgesehen von der zusätzlichen Netzbelastung. Mittels UDP unterstützt die

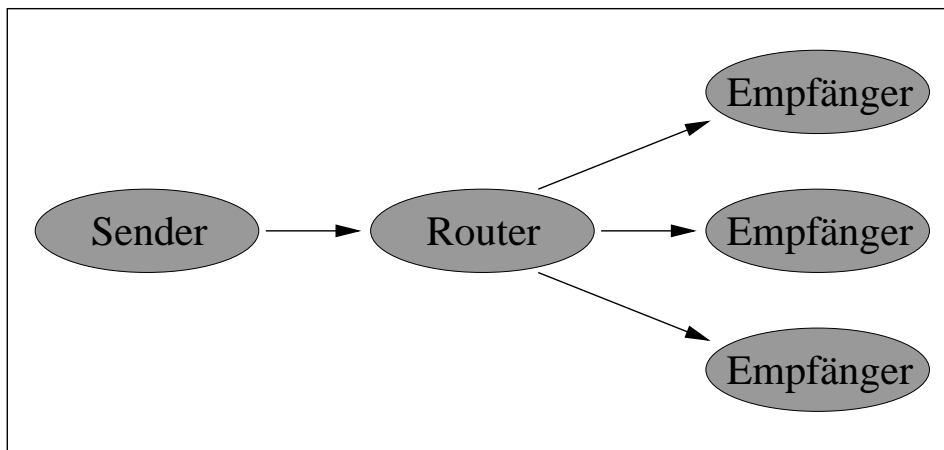


Abbildung 25. Mit Multicast-Unterstützung.

Internet-Protokollfamilie noch einen Datagrammdienst, welcher bei fehlerhaften Paketen diese einfach verwirft (was bei Video-Daten durchaus gewünscht wird). XTP seinerseits unterstützt zusätzlich zu dem go-back-n Verfahren noch ein selektives Verfahren [SW], bei dem nur diejenigen Pakete angefordert werden, die fehlerhaft waren. Ein Datagrammdienst besteht ebenfalls durch setzen der Pakettypen-Flags RCLOSE, WCLOSE [SH95]. Da das RTP bzw. die Tenet RTP Suite auf UDP aufbaut oder diese ergänzt, kann bei beiden Protokollen für den isochronen Datenverkehr ein Datagrammdienst verwendet werden [Har95], [BFMM94].

4.4 Synchronisierung mehrerer Datenströme

TCP/UDP und TP4 bieten keine Mechanismen um unterschiedliche Datenströme, wie z.B. getrennte Video- und Audiokanäle, zu synchronisieren. RTP hingegen besitzt eine sog. Synchronisationsquelle (synchronization source, SSRC), mit welcher mehrere Beitragsquellen (contributing source, CSRC) auf einem Kanal, unter Berücksichtigung der jeweiligen Dienstqualität, übertragen werden können [Har95]. XTP kann mittels synchronisierendem Handshake Kontrollpakete zum Statusabgleich zwischen Sender und Empfänger austauschen und somit einem Abgleich der Datenströme erzielen [SH95].

4.5 Aufbau des Paketformats und Adressierung

TP4 ist durch die Vielzahl von unterschiedlichen Formaten (insgesamt 9 unterschiedliche TSDUs) schwer für Hochleistungsnetze handhabbar, dafür bieten sie aber einen kompakten Aufbau (was den Protokolloverhead reduziert). TCP besitzt nur ein einziges Paketformat mit sehr unterschiedlichen Längen der einzelnen Felder [Krü95], [Krü95]. Zusätzlich erschwert der variable Options-Teil die Protokollbearbeitung, da dadurch die Länge des Paketkopfes schwanken kann. Die Adresse des Senders/Empfängers wird bei TCP immer in Form der Portadresse komplett mitgeführt. XTP besitzt eine feste Headerlänge von 32 Bytes und eine Felderaufteilung, welche auf eine Hardwareoptimierung ausgelegt ist [SH95]. Zur Kanalzuordnung werden eindeutige Referenzen festgelegt, wodurch eine schnelle Zuordnung ermöglicht wird. RTP besitzt einen variablen Paketkopf,

da dieser mittels Anhangdateien individuell erweitert werden kann. Der Paketkopf ist aber immer auf 4-Byte-Grenzen ausgerichtet. Die Adresszuordnung erfolgt bei RTP mittels SSRC-Feld (synchronization source), welche zufällig ermittelt wird und global eindeutig ist [Har95].

5 Beurteilung der vorgestellten Transportprotokolle hinsichtlich isochronem Datenverkehr

Dadurch, daß die Internet-Protokollfamilie, wie auch das ISO/OSI TP4 hauptsächlich für den sicheren File-Transfer und weniger für Hochgeschwindigkeitsdaten ausgelegt sind, versteht es sich von selbst, daß diese beiden Transportprotokolle nicht den geforderten Ansprüchen des isochronen Datenverkehrs entsprechen können. Dies liegt hauptsächlich auch daran, daß diese Transportprotokolle den neuen Anforderungen z.B. von glasfaserbasierten Netzen nicht mehr gerecht werden, da TCP und TP4 vermehrt von Bitfehlern ausgehen, bei den modernen Netzwerken aber häufig Pufferengpässe beim Empfänger, Verstopfungen (und somit Datenverluste) bei den Routern oder übertragungsbedingte Bündelfehler auftreten [Wea94]. Dies hat zur Folge, daß die beiden Transportprotokolle in dieser abschließenden Beurteilung außen vor bleiben.

5.1 Der Ansatz der Tenet Real-Time Protocol Suite

Dadurch, daß der Tenet-Ansatz die bestehende Internetfamilie (TCP/UDP/IP) um echtzeitfähige Protokolle (RMTP/CMTP) erweitert, ist eine Migration - basierend auf einem bereits bestehenden System unter TCP - zu einem Hochleistungsnetzwerk mit Echtzeitfähigkeiten möglich, zumal TCP/UDP unverändert als Transportprotokoll weiterverwendet werden können. Das schichtenübergreifende Signalisierungsprotokoll RCAP [vergl. dazu Abbildung 23] garantiert einen Verbindungsaufbau gemäß der geforderten Dienstqualität, was für Echtzeitanwendungen unerlässlich ist. Messungen haben gezeigt, daß unter Verwendung von RMTP/RTIP bei Videoübertragungen - unter zusätzlicher Netzwerkbelastung - ein konstanter Durchsatz gehalten werden kann (bei UDP/IP gibt es hingegen Durchsatz-Einbrüche). Die durchschnittliche Paketverzögerung lag dabei um 7,37 ms [BFMM94]. Diese Tests bestätigen dem Tenet Real-Time Protocol die Tauglichkeit für isochronen Datenverkehr.

5.2 Der Ansatz des Realzeit-Transportprotokolls (RTP)

Anders als der Tenet-Ansatz geht das RTP bei seiner Transportprotokollimplementierung ja von einer bereits bestehenden Ende-zu-Ende Transportverbindung aus, setzt also auf ein bestehendes Transportprotokoll auf, um dieses funktional zu erweitern. Dies hat zur Folge, daß völlig neue Funktionen entstanden sind. So unterscheidet RTP zusätzlich sog. *Zwischensysteme*, welche RTP-Pakete bearbeiten (ändern von Nutzdaten oder des Paketkopfs). Diese Zwischensysteme können einerseits verschiedene Datenströme bündeln (geschieht in einem Mixer) oder die Nutzdaten umkodieren (Funktionalität

des sog. Übersetzers). Durch den Mixer können beim Übergang in ein leistungsärmeres Netzwerk mehrere Audioströme verschiedener Quellen gebündelt werden, wodurch zwar die Qualität sinkt, die isochrone Übertragung jedoch erhalten bleibt. Da RTP den Anwendungen einen Ende-zu-Ende Transport für kontinuierliche Daten bereitstellt, kann RTP für die Übertragung isochroner Daten verwendet werden. Garantien für bestimmte Dienstgüte-Merkmale (z.B. der Bandbreiten-Reservierung) sind aber nicht Bestandteil von RTP, was speziell unter Lastsituationen zum Einbruch einer bestehenden Verbindung führen kann. Dies wirkt sich besonders stark bei isochronem Datenverkehr aus, wodurch RTP um die Funktionalität der Dienstgüten-Garantien erweitert werden muß.

5.3 Der Ansatz des Xpress Transfer Protocols (XTP)

Der wohl umfassendste Ansatz zur Unterstützung der unterschiedlichsten Transportdienste wurde mit dem Xpress Transfer Protocol erzielt. Durch seine unterschiedlichen Pakettypen, welche mittels cmd-Feld (Command-Feld) im XTP-Header verschlüsselt sind, kann XTP die verschiedensten Dienstanforderungen - durch setzen der jeweiligen Optionen - unterstützen (um nur einige zu nennen: Aktivieren/Deaktivieren der Fehlerkontrolle, Aktivieren/Deaktivieren der Flußkontrolle, Anzeigen von Fehlen, Schließen der Sende/Empfangsrichtung, Anfordern eines Kontrollpakets etc.). Mit der Möglichkeit zur Aushandlung der Verkehrscharakteristiken, können beim Verbindungsaufbau gezielt die zu unterstützenden Dienstqualitäten angegeben werden. Es ist dabei auch möglich, sich Qualitätsverletzungen anzeigen zu lassen, sowie genaue Angaben zur geforderten Verkehrscharakteristik zu machen (z.B. gewünschter Durchsatz, minimal geforderter Durchsatz, höchste akzeptable Verzögerung). Insgesamt läßt sich festhalten, daß der Ansatz des Xpress Transfer Protocols der wohl umfassendste Entwurf - auch hinsichtlich der Unterstützung des isochronen Datenverkehrs - zur Implementierung eines Transportprotokolls darstellt. Wie umfassend die jeweilige XTP-Implementierung in der Realität dann ist, liegt in der Hand des Programmierers, da die XTP-Spezifikation nur einige Pakettypen (wie z.B. das „FIRST“-Paket zum Verbindungsaufbau) fest vorschreibt, andere der Entscheidung des Programmierers überläßt.

5.4 Abschließende Beurteilung

Zusammenfassend läßt sich feststellen, daß keines der vorgestellten Transportprotokolle perfekt ist. Jedes hat seine ganz besonderen Vorzüge aber auch Nachteile. So ist die Koexistenz von TCP/UDP bei dem Ansatz der Tenet Suite sicherlich für denjenigen interessant, der zweigleisig fahren will und bestehende TCP/UDP-Dienste unverändert nutzen will. Diejenigen, die RTP einsetzen, werden die zusätzlichen Eigenschaften der Zwischensysteme zu schätzen wissen und zuletzt diejenigen die sich für XTP entscheiden, dessen Flexibilität hinsichtlich der unterstützten Dienste. Es läßt sich abschließend aber dennoch sagen, daß derzeit der Ansatz des Xpress Transfer Protocols den wohl am weitreichendsten Dienst hinsichtlich isochronem Datenverkehr bereitstellt.

Mobile Datenkommunikation in GSM-Netzen

Martin Kemmler

Kurzfassung

Mobile Datenkommunikation hat in den letzten Jahren sehr stark an Bedeutung gewonnen. Die nachfolgende Ausarbeitung soll einen Einblick in die Infrastruktur eines GSM-Netzes geben, sowie die darin angebotenen Dienste aufzeigen. Dabei werden zuerst in Kapitel 1 allgemeine Eigenschaften eines GSM-Netzes erläutert. Anschließend erfolgt in Kapitel 2 die Beschreibung der Infrastruktur sowie die technischen Grundeigenschaften und -spezifikationen. Kapitel 3 beschreibt die in einem GSM-Netz angebotenen Dienste, vornehmlich dem Sprachdienst als Basisdienst. Zum Abschluß werden in Kapitel 4 mögliche zukünftige Entwicklungen als Ausblick noch gestreift.

1 Einführung

Das digitale Mobilfunk-Zeitalter begann in der Bundesrepublik Deutschland am 1.7.1992. Nach nahezu 10 Jahren Entwicklung erfolgte dies im Rahmen eines europäischen Standards - GSM (Global System for Mobile Communication). In Deutschland ist GSM besser bekannt unter der Bezeichnung D-Netz. Mit der Einführung des D-Netzes wurden in der Bundesrepublik erstmals zwei Lizenzen vergeben:

D1, betrieben durch die Deutsche Telekom AG

D2, betrieben durch ein Firmenkonsortium, vornehmlich Mannesmann AG.

Ein weiteres digitales Mobilfunknetz, das E-Netz, wird seit geraumer Zeit zusätzlich in Deutschland betrieben. Dieses Netz wird vom E-plus-Konsortium (VEBA AG, Thyssen AG, ...) betrieben und arbeitet nach einem ähnlichen Prinzip wie GSM. Einige Unterschiede werden später noch erläutert.

Momentan werden in Deutschland ca. zwei Millionen digitale Mobilfunkgeräte betrieben und Schätzungen gehen davon aus, daß es bis zum Jahr 2000 ungefähr 10 Millionen sein werden. In Westeuropa sollen dann über 45 Millionen Geräte in Betrieb sein, hauptsächlich in Großbritannien, Deutschland und Italien.

GSM hat sich inzwischen zu einem weltweiten Standard mit Ausnahme von den USA und Japan entwickelt. Der Weltmarktanteil liegt momentan bei ca. 85% und soll im Jahr 2000 trotz zunehmender Bedeutung der Nachfolgetechnologien wie DCS-1800 (PCN, E-Netz), D-AMPS und E-TDMA (Time Division Multiple Access) immer noch beachtliche 50% betragen [Boo95], [Gro95].

Funknetze sind ein wichtiger Faktor in Regionen, in denen die Festnetz-Infrastruktur nicht sehr gut entwickelt ist. Das ist in vielen Teilen Osteuropas der Fall, aber auch in

Ostdeutschland haben sich die Funknetze, bedingt durch die mangelhafte Festnetzverbreitung und -qualität, schnell etabliert.

Durch die Schaffung eines einheitlichen Standards -GSM- wurde es möglich, für die Benutzer einen europaweiten Netzzugang zu verwirklichen. Jeder Mobilfunkteilnehmer ist permanent erreichbar und kann unter seiner „persönlichen Rufnummer“ ständig und überall kommunizieren. Die „persönliche Kommunikation“ (personenbezogene Rufnummer) erlaubt, daß die Benutzerstation temporär wechselnde Rufnummern in Form von auswechselbaren Karten erhält, was jedoch eine spezielle Registrierung im Netz erfordert. In Anbetracht der Tatsache, daß die Ortsgebundenheit der Mobilteilnehmer nicht mehr gegeben ist, läßt sich die tatsächliche Entfernung nicht mehr aus der Rufnummer ablesen.

GSM erlaubt den Zugriff auf gewohnte Telekommunikationsdienste auch in fremden Netzen. Diese sehr praktische Eigenschaft ist jedoch nur für die Grunddienste gewährleistet. Die weiterführenden Dienste sind in den jeweiligen Ländern unterschiedlich weit entwickelt. Zusätzlich erlaubt GSM auch die Verschlüsselung der Übertragung, was ein Abhören unmöglich macht.

Der GSM-Standard erlaubt einen gewissen Spielraum bei der praktischen Realisierung der einzelnen inneren Netzkomponenten. Es muß jedoch gewährleistet sein, daß der innere Aufbau des Netzes keine Rückwirkungen auf die Schnittstellen zu den Partnernetzen hat und daß die Funkschnittstelle zwischen der mobilen Benutzerstation und dem Netz einheitlich ist [FBJ⁺93].

Das Netz seinerseits sorgt für den ständigen Kontakt zu der Mobilstation (MS, Mobile Station), was als kontinuierliche Mobilität bezeichnet wird. So erhält das Netz laufend Kenntnis über den Aufenthaltsort der MS und kann eine Umwertung der ortsunabhängigen logischen Rufnummer des Teilnehmers in eine physikalische Netzadresse des Aufenthaltsortes durchführen. Dies führt zur Trennung zwischen Rufbehandlung (call control) und eigentlicher Nutzverbindung (connection control).

Zudem hat das Netz die Fähigkeit, bestehende Verbindungen entsprechend der Bewegung der MS möglichst unterbrechungsfrei nachzuführen.

2 Aufbau / Organisation

2.1 Infrastruktur

Das GSM-Netz läßt sich in drei Subnetze unterteilen, die jeweils unterschiedliche Aufgaben erfüllen:

Funknetz,
intelligentes Mobilvermittlungsnetz und
Verwaltungsnetz.

Sende-/Empfangsstation, zentrale Steuerungseinrichtung und Transcodiereinrichtung werden in einer Station zum Funkfeststationssystem (BSS, Base Station System) zusammengefaßt. Das BSS faßt Funktionen zur Bereitstellung mobiler Funkanschlüsse zusammen.

Funknetz

Das Funknetz übernimmt die Funkversorgung im geographischen Bereich eines GSM-Netzes. Es umfaßt alle zur Funkkommunikation erforderlichen Sende- und Empfangseinrichtungen. Die einzelnen Elemente des Funknetzes sind:

Sende-/Empfangsstation (BTS, Base Transceiver System),
zentralen Steuerungseinrichtung (BSC, Base Station Center) und
Transcodiereinrichtung (TCE), sowie
A-Schnittstelle und
A_{bis}-Schnittstelle.

- **Sende-/Empfangsstation (BTS):** Eine BTS versorgt eine Funkzelle und enthält die dafür erforderlichen Funkeinrichtungen und Antennen. Die Übertragung zwischen BTS und MS wird auch als Um-Schnittstelle bezeichnet.
- **zentrale Steuerungseinrichtung (BSC):** Die zentrale Steuerungseinrichtung BSC ist für das Netzwerkmanagement zuständig und übernimmt die übergreifenden Steuerungsfunktionen für mehrere BTS. Er lokalisiert den Netzkunden, regelt das Weiterreichen von Funkzelle zu Funkzelle und überwacht die jeweilige Signalstärke.
- **Transcodiereinrichtung (TCE):** Die TCE wandelt Nutzkanäle der Funkschnittstelle mit Übertragungskapazitäten ≤ 13 KBit/s in 64 KBit/s-Kanäle des Festnetzes und umgekehrt. Sie gehört funktional zur BSC. Die TCE kann entweder bei einer BSC oder bei einer Mobilvermittlungseinrichtung (MSC, s.u.) befinden, wobei im zweiten Fall die 64 KBit/s-Kanäle im Vierfach-Multiplex zwischen BSC und MSC genutzt werden.
- **A-Schnittstelle:** Zwischen Funknetz und Mobilvermittlungsnetz wird die offen standardisierte A-Schnittstelle verwendet, welche auf 2 MBit/s-Übertragungssystemen basiert. Die 2 MBit/s werden in 31 64 KBit/s-Kanäle unterteilt, wobei 30 davon als Nutzkanäle und ein Kanal als zentraler Zeichengabekanal zur Verfügung stehen.
- **A_{bis}-Schnittstelle:** Die ebenfalls offen standardisierte A_{bis}-Schnittstelle ermöglicht die vierfache Ausnutzung von 64 KBit/s-Kanälen. Jeder dieser 16 KBit/s-Nutzkanäle überträgt die Nettoinformation eines Funkkanals (13 KBit/s Sprache, 12, 6 oder 3.6 KBit/s Daten) und spezielle Steuerungsinformationen zur Steuerung der TCE [FBJ+93].

intelligentes Mobilvermittlungsnetz

Das intelligente Mobilvermittlungsnetz besteht aus ISDN-fähigen Vermittlungseinrichtungen, die auf Funkübertragung und Weiterreichung erweitert wurden, sowie neuen Netzelementen, die die persönliche Mobilität und das Weiterreichen von Gesprächen unterstützen. Die Aufgaben des intelligenten Mobilvermittlungsnetzes sind sehr weitreichend. Es sorgt für die ständige Erreichbarkeit des Mobilteilnehmers unter dessen persönlicher Rufnummer. Weiterhin stellt es Kommunikationsverbindungen zwischen mobilen Benutzern und einem oder mehreren Kommunikationspartnern im GSM-Netz

oder in Partnernetzen her. Eine weitere Aufgabe ist die Bereitstellung der angebotenen Dienste (s. Kapitel 3) für die Mobilfunkteilnehmer. Das intelligente Mobilvermittlungsnetz wird aus folgenden Einheiten gebildet [FBJ⁺93]:

Heimatregister (HLR, Home Location Register),
Aufenthaltsregister (VLR, Visitor Location Register)
Mobilvermittlungseinrichtung (MSC, Mobile Switching Center),
Authentikationszentrale (AC, Authentication Center),
Geräteidentifizierungsregister (EIR, Equipment Identification Register) und
Signalisierungsnetz (ZZK-Netz).

- **Heimatregister (HLR):** Das HLR bildet die obere Ebene einer zweistufigen Datenbank mit zentralisierter Steuerungsfunktion. In ihm werden Benutzernummer, Benutzerberechtigungen, Basisdienste, ergänzende Dienstmerkmale und Angaben zu Mehrwertdiensten verwaltet. Im HLR wird außerdem die Adresse des VLR gespeichert, in dem sich der mobile Benutzer aufhält bzw. aktualisiert, wenn der VLR-Bereich gewechselt wird. Zusätzlich führt das HLR eine Umsetzung der persönlichen Rufnummer in eine physikalische Zieladresse durch. Somit unterstützt das HLR die persönliche Mobilität des Benutzers. Bestimmte Dienstmerkmale wie z.B. unbedingte Anrufweitschaltung werden ebenfalls durch das HLR durchgeführt. Ein Heimatregister speichert die Informationen für typisch 100.000 - 500.000 Benutzer.
- **Aufenthaltsregister (VLR):** Das VLR bildet entsprechend der Hierarchie die untere Ebene der Datenbank. Es führt für alle Benutzer, die sich in seinem Bereich aufhalten, eine Kopie der HLR-Informationen. Zusätzlich werden noch weitere Informationen wie der Aufenthaltsort des Benutzers im Funknetz, die Aktualisierung dieser Information bei Bewegungen der Mobilstation und der Erreichbarkeitszustand des Mobilteilnehmers gespeichert. Auch hier werden bestimmte Dienstmerkmale durchgeführt wie z.B. bedingte Anrufweitschaltung für den Fall, daß der Benutzer funktechnisch nicht erreichbar ist.

Bei einem Wechsel des Aufenthaltsbereichs einer Mobilstation innerhalb eines VLR-Bereichs (s.u.) erfolgt keine Änderung im HLR. Allgemein läßt sich sagen, daß ein Zugriff auf das HLR nur dann erforderlich ist, wenn der Benutzer den VLR-Bereich wechselt, bestimmte gespeicherte Informationen vom Benutzer geändert werden, oder ein ankommender Verbindungswunsch ansteht. Selbst bei der gewählten Struktur sind so noch einige 100 Zugriffe in der Sekunde auf die Datenbank erforderlich. Inkonsistenzen zwischen HLR und VLR werden dadurch vermieden, daß Aktualisierungen, z.B. ein neues Ziel für die Rufumleitung, grundsätzlich zuerst im HLR durchgeführt werden, bevor eine neue Kopie an das VLR geht.

- **Mobilvermittlungseinrichtung (MSC):** Die MSC bildet die Schnittstelle zum Funknetz und zu den jeweiligen Festnetzen. Sie ist eine Erweiterung ISDN-fähiger Vermittlungseinrichtungen und verarbeitet Ergebnisse aus den Datenbanzugriffen, gibt Funk- und Netzverbindungen frei und führt die Rufbehandlung einschließlich Vergütung durch.

- **Authentikationszentrale (AC):** Die AC enthält alle Mittel, die zur Erzeugung sicherheitsrelevanter Parameter erforderlich sind, wie z.B. teilnehmerindividuelle Schlüssel. Die erzeugten Parameter werden mit Hilfe eines geschützten Filetransfers zu anderen Netzknoten übertragen.
- **Geräteidentifizierungsregister (EIR):** Das EIR verwaltet die mobilen Benutzersstationen. Es speichert Mobilstationsspezifische Parameter wie Gerätekennungen und HW-/SW-Versionen der Geräte. Außerdem wird eine Liste aller gestohlenen bzw. nicht zugelassenen Endgeräte geführt.
- **Signalisierungsnetz:** Für den Austausch von Steuerungsinformationen zwischen HLR, VLR und MSC wird das Signalisierungsnetz verwendet. Es besteht aus zentralen Zeichenkanälen (ZZK-Netz).

MSC, HLR, EIR und VLR steuern den gesamten Gesprächsverkehr über umfangreiche Softwareprogramme.

Verwaltungsnetz

Das Verwaltungsnetz sorgt dafür, daß Netzkomponenten unterschiedlicher Hersteller mit einheitlichen Steuerungsprozeduren betrieben werden können. Außerdem führt das Verwaltungsnetz Betriebs- und Wartungsoperationen innerhalb des Funk- und Vermittlungsnetzes durch.

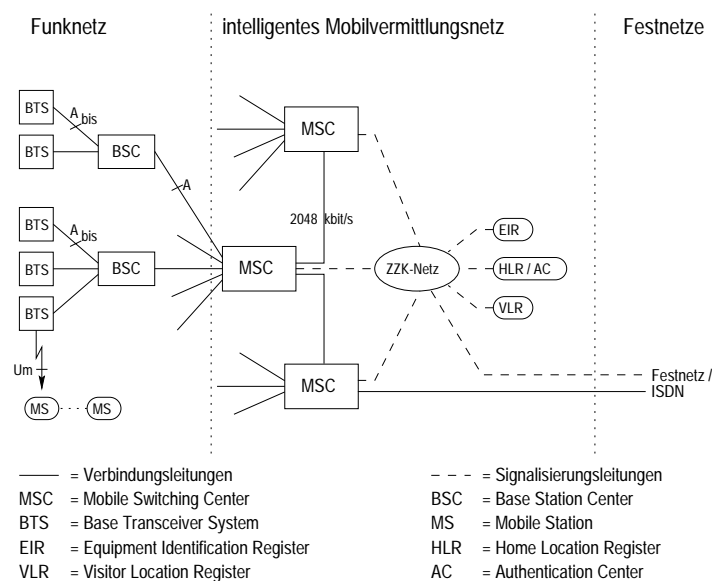


Abbildung 26. Architektur eines GSM-Netzes

2.2 Geographische Organisation

Ein GSM-Netz ist in unterschiedliche geographische Bereiche gegliedert:

Funkzelle,

Aufenthaltsbereich (LA, Location Area),
MSC-Bereich und
VLR-Bereich.

Funkzelle

Die kleinste funktionale Einheit ist die Funkzelle. Ihre Größe wird durch die Ausbreitungseigenschaften der Funkwellen bestimmt. Das GSM-Netz ist als Kleinzellen-Netz konzipiert. Es besteht aus über 2500 Zellen mit Sende- und Empfangsanlagen. Eine Funkzelle kann einen Radius zwischen 500 m und 35 km abdecken. Er richtet sich u.a. nach der Topographie und dem zu erwartenden Gesprächsaufkommen pro Stunde. Die Topographie spielt deshalb eine Rolle, da sich die Funkwellen bei dieser Frequenz quasi-optisch, also „so weit das Auge reicht“, ausbreiten. Deshalb sind Funkzellen in hügeligen Gebieten oder mit Hochhäusern, welche die Ausbreitung ebenfalls beschränken, kleiner als im Flachland, wo nur wenige Hindernisse die Funkwellenausbreitung stören. Das Gesprächsaufkommen beeinflusst den Zellradius wie folgt: Je mehr Gespräche bewältigt werden müssen, desto kleiner sind die Zellen und dementsprechend dichter liegen ihre Zentren. Dieser „engmaschige“ Fall tritt z.B. an vielbefahrenen Autobahnkreuzen und in Großstädten auf [MMU95]. Bei E-Netz-Funkzellen ist der Radius erheblich kleiner, was auf die höhere Trägerfrequenz zurückzuführen ist (<200 m Radius).

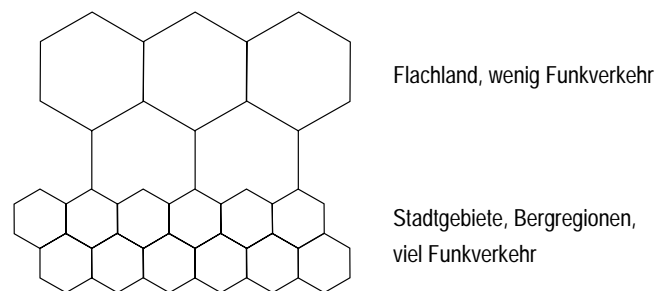


Abbildung 27. geographische Anordnung der Funkzellen

Aufenthaltsbereich (LA)

Mehrere benachbarte Funkzellen können zu einem Zellenverband, dem Aufenthaltsbereich, zusammengefaßt werden. Bei ankommendem Ruf wird der Benutzer in allen Funkzellen des LA mittels Rundsendeverfahren gesucht. Wechselt die Mobilstation den Aufenthaltsbereich, muß die Aufenthaltsinformation des Benutzers im Netz aktualisiert werden.

MSC-Bereich

Eine oder mehrere LA bilden wiederum einen MSC-Bereich. Dieser unterstützt die persönliche Mobilität des Benutzers und bildet die Schnittstelle zum Funknetz und zu den jeweiligen Festnetzen. Ein zusätzlicher Dienst des MSC-Bereichs ist die Anzeige laufender Gebühren während einer Verbindung.

VLR-Bereich

Eine oder mehrere MSC-Bereiche können dann noch zu einem VLR-Bereich zusammengefaßt werden.

Es gibt nun verschiedene Möglichkeiten der praktischen geographischen Realisierung. So können mehrere Ebenen an einem Ort zusammengefaßt werden, was den Signalisierungsaufwand verringert, den Verwaltungsaufwand und die Zahl der Zugriffe jedoch erhöht. Folgende Möglichkeiten stehen zur Verfügung:

zentralisierte VLR/HLR-Realisierung,
kombinierte Lösung MSC mit integriertem VLR und
MSC mit integriertem VLR und HLR.

- **zentralisierte VLR/HLR-Realisierung:** Bei der zentralisierten VLR/HLR-Realisierung sind HLR und VLR getrennt von der MSC ausgeführt. Dabei sind HLR und VLR für mehrere MSC zuständig.
- **kombinierte Lösung MSC mit integriertem VLR:** Bei der physikalisch kombinierten Lösung von MSC mit integriertem VLR hingegen verringert sich der Signalisierungsaufwand. Das HLR wird mit Rücksicht auf eine effiziente Teilnehmerverwaltung getrennt realisiert. So ist nur eine externe Signalisierung zwischen MSC/VLR und HLR zu realisieren. Diese Lösung wird z.B. bei der Telekom im D1-Netz verwendet.
- **MSC mit integriertem VLR und HLR:** Eine weitere Möglichkeit ist die Realisierung einer MSC mit integriertem VLR und HLR. Bei dieser Lösung entfällt die redundante Speicherung von Daten im VLR für Teilnehmer, die sich im Heimat-MSC aufhalten. Diese Kombination wird im C-Netz der Telekom angewandt.

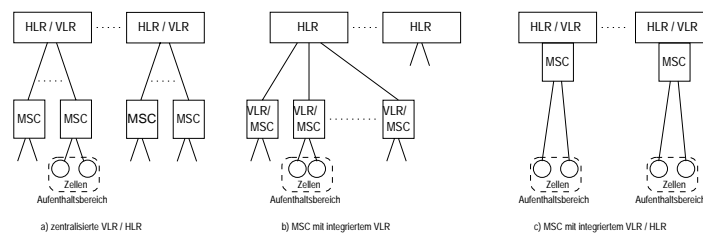


Abbildung 28. mögliche Netzkonfigurationen

2.3 Verkehrslenkung und Rufaufbau

Die persönliche Mobilität der Benutzer verhindert eine Zuordnung der Rufnummer zum aktuellen Aufenthaltsort. Die Rufnummer dient vielmehr zur Initiierung einer Datenbankabfrage über den aktuellen Aufenthaltsort. Wegen administrativer Grenzen erfolgt diese Abfrage in der Regel erst im Heimatnetz. So geht z.B. ein Anruf aus Dänemark an einen deutschen Benutzer, der sich in Schweden aufhält, erst nach Deutschland, wo festgestellt wird, daß sich der Benutzer in Schweden aufhält. Von Deutschland erfolgt dann

der entsprechende Rufaufbau. Diese Beschränkung bezieht sich auf den Rufaufbau, der eigentliche Verbindungsaufbau kann eventuell auch direkt erfolgen. Die Trennung zwischen Rufaufbau (call control) und Verbindungsaufbau (connection control) ist für eine optimale Nutzung der Funknetze sinnvoll und erfordert in jedem Fall eine Feststellung des Aufenthaltsortes vor der eigentlichen Gesprächsverbindung.

Wie funktioniert nun der Rufaufbau bei einem Anruf aus dem Festnetz?

Zuerst erhält eine beliebige MSC des GSM die Adressnachricht aus dem Festnetz. Anhand der Rufnummer stellt die MSC fest, daß es sich um einen Mobilteilnehmer des eigenen Netzes handelt. Dann startet die MSC eine Abfrage der Aufenthaltsadresse bei dem der Rufnummer zugeordneten HLR und übermittelt gleichzeitig die erhaltene Übermittlungsdienstkennung und die Rufnummer. Das HLR prüft die Teilnehmerberechtigung und fordert dann vom VLR, in dem der Teilnehmer zuletzt registriert wurde, eine Adresse für die Verkehrslenkung ab (MSRN, mobile subscriber routing number). Gleichzeitig werden bestimmte Daten des Rufes wie Teilnehmernummer und der damit verbundene Dienst übermittelt. Die MSRN dient später der Zuordnung dieser Daten und zum Routing des Gesprächs. Nun reicht das HLR die erhaltene MSRN an die MSC weiter, welche die Teilnehmerrufnummer durch die MSRN ersetzt und baut den Ruf zur Ziel-MSC auf. Erhält nun die (Ziel-)MSC einen ankommenden Ruf mit MSRN als Rufnummer, startet sie zum VLR die Behandlung eines ankommenden Rufes. Aufgrund der MSRN kann das VLR die vorher erhaltenen Rufdaten zuordnen und startet für den zuletzt eingetragenen Aufenthaltsbereich einen Selektivruf. Bei Erfolg informiert die MSC das VLR über den Zugriff der MS auf das Netz und weist der MS einen Kanal zu.

Diese Art des Verbindungsaufbaus ist sehr aufwendig. Hält sich der Mobilteilnehmer gar im Ausland auf, wird diese Prozedur noch aufwendiger, da der Ruf zunächst im Heimatnetz aufgebaut wird. Von dort wird dann die Signalisierung zum VLR des Aufenthaltsnetzes durchgeführt und der Ruf dann weiter in das Aufenthaltsnetz aufgebaut. Erste Erfahrungen mit dieser Art des Rufaufbaus haben gezeigt, daß die Signalisierung beim internationalen Roaming bei D1 recht schnell ist. So dauert eine Abfrage der Routinginformation von Deutschland nach Finnland für einen Teilnehmer, der sich wiederum in Deutschland aufhält, ungefähr 1.2 Sekunden. Andererseits sind die Übertragungszeiten auf dem internationalen Funkweg sehr lang, besonders die Zeiten für den (mehrfachen) Selektivruf des Teilnehmers und der Verbindungsaufbau, der in Extremfällen über eine Minute (!) erfordern kann.

Bewegt sich nun ein Teilnehmer während eines Gesprächs über die Zellengrenzen hinweg, muß die Verbindung an die neue Funkzelle weitergereicht werden. Dieser Vorgang ist nicht wie bei den analogen Vorgängernetzen im Netz zentralisiert, sondern auf die MS und das Netz verteilt. Das Funknetz kann einen Wechsel eigenständig vornehmen, die zuständige MSC wird dann lediglich informiert. Befindet sich die MS im Überlappungsbereich mehrerer Zellen, kann eventuell eine Weiterreichung aufgrund der günstigeren Leistungsbilanz der anderen Zelle oder einer gleichmäßigeren Auslastung der Zellen erfolgen [FBJ⁺93].

Sehr wichtig zur Aufrechterhaltung eines optimalen Netzbetriebes ist eine geringe Ausfallwahrscheinlichkeit der einzelnen Netzkomponenten. So dürfen beispielsweise die Knoten

zur Dienststeuerung maximal 10 bis 20 Stunden pro Jahr ausfallen. Die Ausfallzeit einer Signalisierungsverbindung zwischen zwei Knoten darf maximal 10 Minuten und für Knoten zum Durchschalten von Signalisierungsverbindungen maximal 3 Minuten pro Jahr betragen. Zusätzlich sind auch Zeitbeschränkungen der einzelnen Netzkomponenten nötig, um die Gesamtaufbauzeiten für den Kunden in akzeptablen Grenzen zu halten. So wird verlangt, daß die Datenbankabfrage zur Feststellung des Aufenthaltsortes nicht mehr als drei Sekunden benötigt und das Durchschalten einer Signalisierungsverbindung innerhalb von 20 bis 100 ms erfolgt.

2.4 Technische Daten

Die GSM-Spezifikation belegt das Frequenzband zwischen 890 MHz und 960 MHz mit 124 festen Frequenzpaaren. Dabei ist der Frequenzbereich in zwei Teilbänder mit einem Duplexabstand von 45 MHz unterteilt. Der Bereich von 890 MHz - 915 MHz wird für die Übertragung zwischen MS und Netz verwendet, was auch als Uplink bezeichnet wird. Der Downlink-Bereich, der die Übertragung vom Netz zur MS realisiert, liegt zwischen 935 MHz und 960 MHz. Der Abstand zwischen den einzelnen Unterbändern beträgt 200 kHz. So liegt z.B. die Frequenznummer 10 in Uplink-Richtung bei 892 MHz und in Downlink-Richtung bei 937 MHz. Im Vergleich dazu belegt das E-Netz ein Frequenzspektrum von 1710 MHz bis 1785 MHz (Uplink) und von 1805 MHz bis 1880 MHz (Downlink).

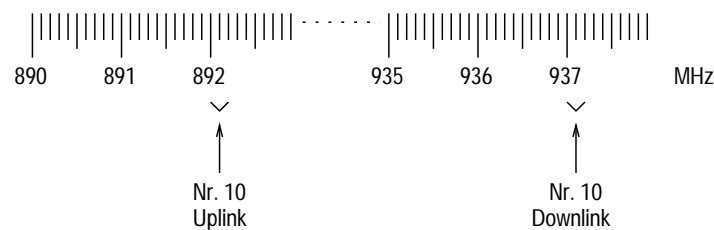


Abbildung 29. Frequenzaufteilung

Die beiden Teilbänder werden durch Mehrfachzugriffsverfahren auf mehrere Übertragungskanäle aufgeteilt. Es wird eine Kombination aus 3 Verfahren angewendet:

Raummultiplex (SDMA, Space Division Multiple Access),
 Frequenzmultiplex (FDMA, Frequency Division Multiple Access) und
 Zeitmultiplex (TDMA, Time Division Multiple Access).

Raummultiplex (SDMA)

Bei SDMA werden gleiche Frequenzen in ausreichend räumlichem Abstand wiederverwendet.

Frequenzmultiplex (FDMA)

Durch FDMA kommen verschiedene Frequenzlagen mit einem Abstand von 200 kHz zur Anwendung.

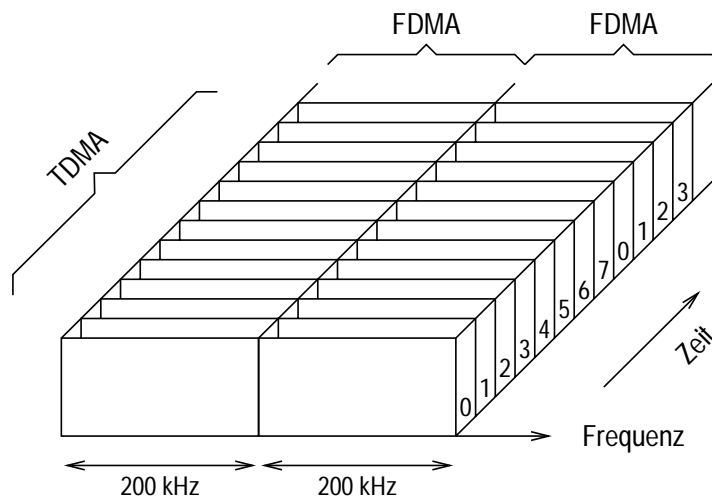


Abbildung 30. Mehrfachzugriff auf Zeit- und Frequenzebene

Zeitmultiplex (TDMA)

Das Zeitmultiplexverfahren wurde aufgrund praktischer Tests ausgewählt, bei denen auch andere TDMA-Varianten, ein symmetrisches Breitband-CDMA (Codemultiplex-Verfahren) und ein Hybridsystem getestet wurden. Die Zeitschlitzlänge eines Nutzkanals beträgt $577 \mu s$ ($=156.25 \text{ Bit}$), davon sind $30,5 \mu s$ ($=8.25 \text{ Bit}$) als Schutzzeit gegen Störungen wie Überlappung von verschieden entfernten MS in der BTS vorgesehen. Eine Übertragung innerhalb eines Zeitschlitzes wird als Burst bezeichnet und umfaßt 148 Bit. Das Übertragungsprotokoll dieser 148 Bit sieht wie folgt aus: 3 Startbits, 57 Datenbits, 1 Trennbit, 26 Korrelationsbits, 1 Trennbit, 57 Datenbits und 3 Stopbits. Daraus ergibt sich in Summe die Übertragung von 114 Datenbits, 26 Korrelationsbits und 6 Anfang-/Endebits.

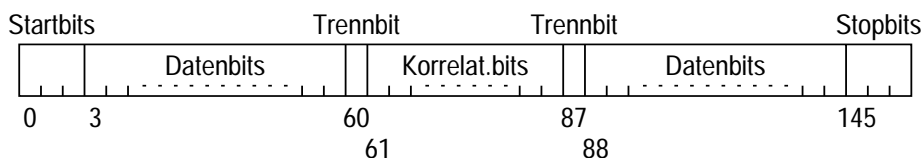


Abbildung 31. Burst

Acht Zeitschlitzte ergeben einen TDMA-Frame, dessen Dauer somit $4,615 \text{ ms}$ beträgt. Jedem Teilnehmer wird ein Timeslot zugeordnet, demzufolge darf die MS nach sieben Timeslots wieder senden, wodurch die Station ungefähr 217 Mal pro Sekunde sendet. Die acht Timeslots wiederum bilden einen physikalischen Kanal. Daher stehen pro Unterband acht Kanäle zur Verfügung. Die Bruttoübertragungskapazität auf einem Träger beträgt ungefähr 270 KBit/s incl. den Signalisierungskanälen.

GSM unterscheidet in seiner Struktur prinzipiell zwischen Verkehrs- und Signalisierungskanälen. Bei GSM sind Voll- und Halbratenkanäle möglich, wobei bei einem Halbratenkanal dann zwei logische Kanäle zur Verfügung stehen.

Ein Verkehrskanal hat eine Brutto-Übertragungskapazität von 22.8 KBit/s , die in 13

KBit/s für den Verkehrskanal und 9.8 KBit/s für Fehlererkennung und -korrektur aufgeteilt werden.

Die Signalisierungskanäle sind wie folgt unterteilt:

Broadcast Control Channels,
Dedicated Control Channels und
Common Control Channels.

Broadcast-Kanäle dienen zur Synchronisierung und zur Übertragung zellspezifischer Parameter vom Netz an die MS.

Dedicated Control Channels (DCC) sind reine Signalisierungskanäle und übertragen Kontroll- und Meßparameter zwischen Netz und MS.

Die Common Control Channels sind dazu da, einen physikalischen Kanal zwischen MS und Netz einzurichten. Unter Umständen können diese Kanäle auch „gestohlen“ werden und als DCCs verwendet werden.

Die physikalischen Kanäle lassen sich in logische Kanäle aufteilen. Diese beschreiben die Funktion, die ein physikalischer Kanal zu einem bestimmten Zeitpunkt hat. Es wird zwischen zwei Kategorien unterschieden:

Mehrpunktkanäle und
Punkt-zu-Punkt-Kanäle.

Mehrpunktkanäle werden benutzt, solange keine Verbindung zwischen MS und Netz aktiv ist. Sie sind für die Verteilung von Information zur Frequenzsynchronisation der MS nötig. Sie dienen auch zur Verteilung funkspezifischer Organisationsinformationen wie Funkzellen- und Netzkennung. Sehr wichtig ist auch die Lokalisierung der MS, die ebenfalls über Mehrpunktkanäle erfolgt.

Punkt-zu-Punkt-Kanäle (PPK) werden für die Übertragung von Steuerungsinformationen und für Kurzmitteilungen verwendet. Es gibt auch einen sogenannten Rundsendekanal, der an alle Benutzer einer Funkzelle gerichtet ist. Der wichtigste PPK ist der Verkehrskanal für die Übertragung. Um eine optimale Aufrechterhaltung der Funkverbindung zu gewährleisten, werden über PPKs auch Steuerungsinformationen unquittiert ausgetauscht.

Es ist noch anzumerken, daß die Übertragung mit einer Versetzung um 3 Bursts zwischen Uplink und Downlink erfolgt (d.h. Senden, 3 Bursts Zeit, Empfangen, 3 Bursts Zeit, Senden, ...). Somit ist nur ein Synthesizer für Senden und Empfangen erforderlich, der dann auf die jeweilige Frequenz umstellt (Uplink-Frequenz und Downlink-Frequenz). Das TDMA-Verfahren stellt hohe Anforderungen an den Sender. Innerhalb von 28 μ s muß die Sendeleistung bei der MS oder auf Netzseite um 70 dB erhöht werden. Für den Fall, daß sich bei langsam bewegenden MS durch Mehrwegausbreitung ein Schwundloch bildet, kann optional ein Frequenzsprungverfahren (slow frequency hopping) verwendet werden. Dabei wird zwischen zwei aufeinanderfolgenden Zeitschlitten eines Funkkanals die Trägerfrequenz gewechselt. Da der Empfang um über 30 dB schwanken kann, sind verschiedene Maßnahmen zur Fehlerkorrektur wie systematische Redundanz zur

Vorwärts-Fehler-Korrektur (FEC, Forward Error Control), Frequenzsprungverfahren und Verwürfelung der Daten auf verschiedene Zeitschlitze erforderlich [WD93], [Kau94].

2.5 Übertragung

Bei der Übertragung zwischen MS und Netz stehen zwei Trägerdienste zur Verfügung:

transparent und
non-transparent.

transparente Übertragung

Bei einer transparenten Übertragung ist der Durchsatz und die Verzögerungsdauer konstant. Zur Verbesserung der Übertragungssicherheit werden ausschließlich fehlerkorrigierende Kodierungsverfahren (FEC) verwendet. Es kommen keine ARQ-Verfahren (ARQ, Automatic Repeat Request) zur Anwendung.

non-transparente Übertragung

Die non-transparente Übertragung leitet sich von der transparenten Übertragung ab, benutzt jedoch zusätzlich ein vom HDLC-Protokoll abgeleitetes, auf Fehlererkennung basierendes ARQ-Protokoll, welches auf die besonderen Eigenschaften bei der Funkübertragung, wie Abschattungen und Schwund durch Mehrwegeausbreitung hin optimiert wurde. Dieses Protokoll heißt Radio Link Protokoll (RLP). Es kommt der asynchrone gleichberechtigte Betriebsmodus zur Anwendung. Zurückweisungen, Wiederholung ab einer bestimmten Blocknummer und die wiederholte Anforderungen einzelner Blöcke sind zulässig. Abweichend vom HDLC-Protokoll werden die Steuerkommandos auch im Informationsblock übertragen. Die maximale Fenstergröße beträgt 62, die Blocks werden also modulo 62 durchnummeriert. Weitere Besonderheiten sind z.B. die feste Länge der Protokolldateneinheit (16 Bit Kopf, 200 Bit Information, 24 Bit FCS) und eine Parameterverhandlung.

Das RLP reduziert die Restfehlerwahrscheinlichkeit auf $<10^{-9}$, bedingt jedoch einen schwankenden Datendurchsatz sowie schwankende Verzögerungszeiten. Dafür gleicht es Unterbrechungen des Übertragungsweges beim Weiterreichen aus. Bei gehäuften Wiederholungen kann eine Reduktion des Durchsatzes auftreten, welche teilweise durch das Weglassen der Start- und Stopbits ausgeglichen werden kann. Auch bei niedriger Datenrate wird auf der Funkschnittstelle stets mit der höchsten Bitrate übertragen, wodurch zusätzliche Kapazität für Wiederholungen zur Verfügung steht.

3 Dienste

Die angebotenen Dienste im GSM-Netz sind sehr vielfältig. Es gibt als Grundlage Basisdienste, die in jedem Netz zur Verfügung stehen, oder in absehbarer Zeit eingerichtet werden. Zusätzlich sind einige Dienste geplant, die in den in der Entwicklung führenden Ländern wie Deutschland und Großbritannien bald realisiert werden, in einigen anderen

Ländern jedoch noch einige Zeit auf sich warten lassen dürften. Als Benutzer solcher Dienste muß man sich im Klaren sein, daß Roaming, also das Weiterreichen in fremde GSM-Partnernetze, nur bei gleicher Dienstanbietung der Netzbetreiber möglich ist. Dies wird z. Zt. nur für die Basisdienste garantiert.

Für die Netzbetreiber gilt der Grundsatz, die angebotenen Dienste so übersichtlich wie möglich zu gestalten, sodaß der Benutzer diese möglichst ohne zusätzliches Wissen wie in einem Festnetz bedienen kann.

Allgemein ist ein Übergang in feste Partnernetze wie das Fernsprechnet, ISDN und auch verschiedene Datennetze vorgesehen, wobei jedoch gewisse Einschränkungen zu machen sind. So haben Festnetze eine niedrige Bitfehlerrate bei kurzen Signallaufzeiten, während das Mobilfunknetz eine hohe Bitfehlerrate mit veränderlichen Laufzeiten aufweist. Diese Eigenschaften führen dazu, daß Festnetze eine hohe Übertragungskapazität besitzen, während bei Mobilfunknetzen die Kapazität aufgrund des begrenzten Funkfrequenzspektrums knapp ist.

Im Bezug auf die angebotenen Dienste kann man zwischen Übermittlungsdiensten und Telediensten unterscheiden. Übermittlungsdienste sind Telekommunikationsdienste mit der Fähigkeit, Informationen codetransparent und anwendungsabhängig zu übermitteln. Sie lassen sich in transparente und non-transparente Dienste unterteilen.

Teledienste umfassen zusätzlich zu den untergelagerten Übermittlungsdiensten alle für die direkte Benutzer-Benutzer-Kommunikation erforderlichen Kommunikationsfunktionen nach einheitlichen Protokollen. Prinzipiell sollte man zwischen

Sprachdienst und
nicht - Sprachdiensten

unterscheiden.

3.1 Sprachdienst

Dieser wohl wichtigste Dienst gehört zu den Telediensten. Er wird selbstverständlich von allen Netzbetreibern zuerst eingeführt und unterstützt. Dazu gehört auch eine europaweit einheitliche Notrufnummer. Mit der Notrufnummer 112 wird der Mobilfunkteilnehmer automatisch zur lokal zuständigen Rettungsleitstelle verbunden. Für die Übertragung von Sprachdaten wird ein Vollraten-Sprachcodec verwendet, der die analogen Sprachdaten in digitale Daten mittels „Regular Pulse Excitation with Long Term Prediction Loop“ (RPE-LTP) transformiert. Dabei werden die Sprachdaten in einem speziellen konvolutionellen Kodierer (Konvolutionelle Kodierung = Algorithmus zur Fehlerkorrektur von Daten, auch Faltungskodierung) geschickt, ergänzt durch Parity-Bits. Sprachdaten sind unterschiedlich groß geblockt, wobei jedem Block eine andere Gewichtung bezüglich der Fehlererkennung und -behebung zugeordnet wird. Dieses Verfahren ist speziell auf Sprache und deren Übertragung über eine Funkschnittstelle hin optimiert [Kau94].

3.2 nicht - Sprachdienste

Die nicht - Sprachdienste können nun ihrerseits weiter unterteilt werden. Folgende Dienste werden aktuell im GSM-Netz angeboten, bzw. in Zukunft angeboten werden:

- Datenübertragung (non-transparent),
- Fax-Gruppe III (transparent),
- Kurznachrichtendienst,
- Kurzmitteilungs-Verteildienst,
- Anrufweitschaltung / Rufumleitung,
- Mobilbox,
- Anrufsperrung,
- automatischer Rückruf,
- Gesprächskonferenz sowie
- geschlossene Benutzergruppen.

Zusätzlich werden noch weitere kleinere Dienste eingeführt, auf die aber hier nicht eingegangen werden soll [Mey95a], [WD93].

Seit Anfang 1995 wird im D-Netz die Daten- und Faxkommunikation als Dienst im Regelbetrieb angeboten. Ein Grund dafür war, daß Modem und Fax für das Festnetz im Ausland häufig nicht richtig einsetzbar sind, da entweder andere Stecker nötig, überhaupt keine Anschlüsse vorhanden oder die Hotel-Telefongebühren übermäßig teuer sind. Man muß dazu aber anmerken, daß die Daten- und Faxkommunikation noch nicht offiziell über BRD-Grenzen hinaus funktioniert, was jedoch nur eine Frage der Zeit ist. Für die Daten- und Faxkommunikation ist es erforderlich, daß der Netzbetreiber in den MSC Modems installiert, da dort die Verbindung zu den Festnetzen umgesetzt wird. Es müssen also digitale GSM-Signale in analoge Frequenzen moduliert werden.

Man unterscheidet Verbindungen auch nach dem Ausgangspunkt, von welchem diese aufgebaut wurden. Wenn eine MS ein Telefonat aufbaut, wird dies als Mobile Originated (MO) bezeichnet. Die entgegengesetzte Richtung ist dementsprechend Mobile Terminated (MT). Die MO-Richtung ist einfacher zu realisieren als die MT-Richtung. Initiiert der Benutzer den Anruf, kann er selber bestimmen, um welche Art der Kommunikation es sich handelt und ggf. die notwendigen Geräte bereits installiert haben. Bei MT muß der Netzbetreiber neue Rufnummernstrukturen aufbauen, da das Handy nicht erkennen kann, ob Sprache, Daten oder ein Fax ankommt. Diese Realisierung ist also nur über zusätzliche Rufnummern möglich. Geplant sind bis zu 5 Rufnummern, pro Typ eine. Die hier möglichen Typen sind Sprache, Fax, Daten 1200, Daten 2400 und Daten mit 9600 KBit/s. Durch die Vermittlungszentrale wird dann dem Handy aufgrund der angewählten Rufnummer signalisiert, um welche Art von Anruf es sich handelt. Auf diese Art kann das Handy dann umgeschaltet werden. Ein weiteres Plus dieser Multi-Numbering-Strategie ist die Möglichkeit, ankommende Daten evtl. nicht über Sprachleitungen an das Handy zu senden, sondern über internationale Datenleitungen. Dies erfordert jedoch weiterreichende Roamingabkommen.

Um nun ein Überblick über die zusätzlich entstehenden Kosten zu bekommen, sollen hier ein paar Preise (Stand: Frühjahr 1995) aufgeführt werden. So liegen die zusätzlichen

monatlichen Grundgebühren für eine zusätzliche Faxnummer bei 11,50 DM und für eine zusätzliche Datennummer bei 25,30 DM. Für den Fall, daß beide Nummern gewählt werden, ist ein Summenpreis von 33,35 DM zu bezahlen (alle Preise bei D2). Zusätzlich ist jedoch noch eine passende PCMCIA-Karte zu erwerben. Diese Karten passen jedoch meistens nur für einen bestimmten Handy-Typ und können nicht beliebig verwendet werden, da die interne Verarbeitungslogik bei den Herstellern nicht standardisiert ist. Diese PCMCIA-Karten kosten momentan noch zwischen 1.000,- und 1.600,- DM, sollen jedoch in Kürze auch für unter 500,- DM zu haben sein. Bei den Herstellern der PCMCIA-Karten haben sich zwei Lager gebildet. Die einen nutzen ausschließlich das RLP-Protokoll, was eine non-transparente Übertragung zur Folge hat. Die anderen Hersteller führen eine Ende-zu-Ende-Kontrolle mit Protokollen wie MNP4 und MNP5 durch. Diese Übertragung erfolgt transparent.

Datenübertragung (non-transparent)

Da die Hauptnutzung der Mobilfunknetze in der Sprachübertragung liegen, wurde für die Ausgestaltung des Funkübertragungsverfahrens ein geeigneter redundanzreduzierender Fernsprechanal zugrunde gelegt. Zwar ist die technische Struktur, die ähnlich der ISDN-Struktur realisiert wurde, auch auf Datenübertragung ausgelegt, aber aufgrund des gewählten RPE-LTP-Verfahrens (s.o.) können Daten nicht so ohne weiteres übertragen werden. Schickt man Nicht-Sprachdaten in dieses Kodiersystem, kommen diese auf der anderen Seite nicht in ihrer Originalform an. Aus diesem Grunde wird ein spezielles Verfahren für die Datenübertragung verwendet. Dieses Verfahren ermöglicht es, auf der Um-Schnittstelle verlorengegangene Bursts zu rekonstruieren. Das im folgenden vorgestellte Verfahren bezieht sich auf eine Übertragungsrate von 9.600 Bit/s, wobei die Verfahren für eine andere Übertragungsgeschwindigkeit sich nur durch eine andere Blockung unterscheiden. Ansonsten bleibt das Prinzip erhalten.

Der Datenstrom wird in einzelne 60 Bit-Blöcke unterteilt. 4 Blöcke werden zusammen mit einem Block, bestehend aus 4 Steuerbit, zu einem Block mit 244 Bit zusammengefaßt. Nun erfolgt eine Faltungskodierung, bei der jeweils 5 Bit-Blöcke betrachtet werden. Aufgrund der außerordentlich redundanten Kodierung entsteht nun ein neuer Block mit 488 Bit. Diese so entstandene Bitfolge wird um 32 Bit nach einer feststehenden Methode gekürzt (=456 Bit) und in 4 einzelne 114 Bit-Blöcke unterteilt. Die so kodierten 456 Bit werden nun über 19 unterschiedliche Blöcke zu 24 Bit geschachtelt. Nachdem dieses Interleaving erfolgt ist, werden sie in 4 aufeinanderfolgende Bursts aufgeteilt und versandt. Beim Einpacken in die Bursts werden Gruppierungen aus geraden und ungeraden Bits gebildet. Für den Fall, daß nun ein Burst auf der Um-Schnittstelle (Luftschnittstelle) verlorengeht, lassen sich die Daten immer noch aufgrund obiger Kodierung vollständig rekonstruieren [Kau94].

Die theoretisch möglichen Übertragungsraten bei GSM sind 600 Bit/s, 1200 Bit/s, 2400 Bit/s, 4800 Bit/s und 9600 Bit/s. Eine mögliche exotische Variante ist auch 1200/75 für BTX, wobei hier die Übertragung vom Netz zum Benutzer mit 1200 Bit/s erfolgt, während die MS zum Netz mit 75 Bit/s überträgt. Diese Möglichkeit wird nicht von allen Netzbetreibern unterstützt.

Da die Übertragungsrate auf der Um-Schnittstelle auf 13,6 KBit/s beschränkt ist, ist es

klar, daß eine transparente 64 KBit/s ISDN-Verbindung nicht möglich ist. Die Möglichkeit einer non-transparenten ISDN-Übertragung besteht theoretisch, es ist jedoch noch nicht untersucht, welche Konsequenzen diese permanent langsamere Kommunikation mit den ISDN-Komponenten bewirken kann.

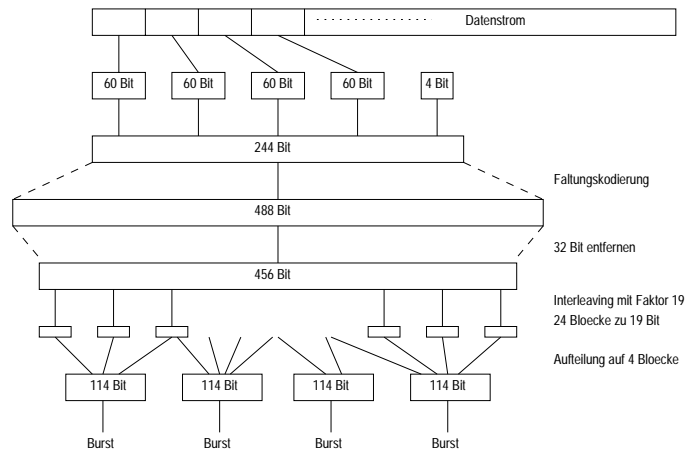


Abbildung 32. Kodierung von Nicht-Sprachdaten

Fax-Gruppe III (transparent)

Die Faxkommunikation gehört zu den Telediensten. Es ist eine Kommunikation mit Geräten der Fax-Gruppe III spezifiziert. Theoretisch sind sowohl transparente als auch non-transparente Übertragungen möglich. Aufgrund der räumlichen Trennung zwischen MS und Modem, sowie der durch die Luftschnittstelle auftretenden zeitlichen Verzögerungen, erwies sich die Spezifikation als schwierig. Daher unterliegt die Nutzung von Zusatzfunktionen gewissen Einschränkungen. Beide Verfahren benötigen eine gesonderte Protokollbehandlung. Die transparente Fax-Kommunikation wurde als unabdingbar eingestuft und primär realisiert. In der Signalisierungsphase wird nur eine Übertragungsgeschwindigkeit von 300 Bit/s verwendet. Die einzelnen Bits werden aus Redundanzgründen wiederholt. Durch die transparente Übertragungsart können bei einem Weiterreichen bis zu 11 mm einer Seite verloren gehen. Die non-transparente Übertragungsart verwendet auch das RLP-Protokoll. Zur Zeit wird dies jedoch noch nicht von den Netzbetreibern angeboten. Zusätzlich wären auch noch spezielle Adapter nötig, die die schwankende Übertragungsrate ausgleichen. Der Vorteil liegt aber eindeutig in der verbesserten Dienstgüte.

Der Fax-Dienst unterliegt jedoch auch einigen Restriktionen, die noch nicht zur endgültigen Zufriedenheit gelöst sind. So kann es z.B. vorkommen, daß das Handy bei ankommendem Fax nicht auf Empfang steht. Ein weiteres Problem entsteht, wenn das Handy zwar auf Empfang, jedoch nicht mit einem Laptop verbunden ist, oder in diesem die entsprechende Software nicht aufgerufen wurde. Dies alles macht die Verwendung einer Mailbox sinnvoll, die ankommende Faxe speichert und eine entsprechende Meldung an die MS ausgibt. Der Benutzer kann dann das Fax zu einem beliebigen Zeitpunkt abrufen. Die Mailbox muß jedoch vom Netzbetreiber eingerichtet und verwaltet werden.

Kurznachrichtendienst

Der Kurznachrichtendienst (SMS, short message service), ebenfalls ein Teledienst, erlaubt die Übertragung von maximal 160 Zeichen pro Nachricht. Diese werden dann auf dem Display des Handy angezeigt, oder über eine V.24-Schnittstelle an ein externes Gerät weitergeleitet. Eine logische Verkettung von mehreren Nachrichten hintereinander ist möglich. Die erfolgreiche Übermittlung der Nachricht wird quittiert, also auf dem Display des Senders angezeigt.

Das Versenden von Kurznachrichten über das Handy ist wegen der komplizierten Eingabe sehr umständlich. Eine einfachere Möglichkeit des Versendens ist die Benutzung eines Operators, der in der Dienstzentrale die Nachricht per Telefon erhält und diese dann als Text versendet. Durch die Verbindung des Handy über die PCMCIA-Schnittstelle zu einem Laptop kann der Text auch am Rechner eingegeben und über eine entsprechende Software versandt werden. Die Übertragung der Kurznachricht erfolgt über einen Signalisierungskanal, wodurch eine bessere Ausnutzung der Funkstrecke ermöglicht wird. Durch diese Art der Übertragung können Verzögerungen im Sekundenbereich auftreten. Eine Kurznachricht kann ohne vorherigen Aufbau einer Netzverbindung erfolgen.

Erfolgt nun eine die Versendung einer Kurznachricht durch eine MS, so wird diese an die MSC übermittelt. Die MSC übergibt diese Nachricht dem Service Center (SC), welches eine Quittung an die sendende MS zustellt. Das SC übermittelt die Kurznachricht mit Zieladresse, Zeitstempel und Nutzerdaten an eine generelle MSC. Dort erfolgt die Abfrage der Aufenthaltsadresse der gewünschten Ziel-MS. Ist die Adresse ermittelt, erfolgt der Versand an die zuständige MSC, welche die Nachricht an die MS übermittelt. Der ordnungsgemäße Empfang bei der MS wird von dieser bestätigt.

Kurzmitteilungs-Verteildienst

Der Kurzmitteilungs-Verteildienst erlaubt das zyklische Aussenden von Kurzmitteilungen von einem Verteilerzentrum (CBC, Cell Broadcast Center) an die Mobilstationen über einen speziellen Cell-Broadcast-Channel. Dieser Dienst ist kein Bestandteil des Netzes, sondern eine spezielle Zusatzfunktion. Ein CBC kann mit mehreren BSC verbunden sein, jedoch kann jede BSC nur mit einem CBC verbunden sein. Die Nachrichten können von verschiedenen Quellen stammen. Jede Nachricht kann 15 Kurzmitteilungen mit gleicher Quellenangabe, gleicher Seriennummer und bis zu 82 Zeichen pro Kurzmitteilung umfassen. Eine Nachricht kann auf Wunsch auch komplett ignoriert werden.

Anrufweitchaltung / Rufumleitung

Die Anrufweiterleitung oder -umleitung kann für mehrere Fälle ausgewählt werden. So ist es möglich, für den Fall, daß ein Benutzer im Augenblick nicht erreicht werden will, eine sofortige Umleitung einzustellen. Alternativ können auch Weiterleitungen im Besetztfall, wenn keine Antwort erfolgt oder bei Nichterreichbarkeit eingestellt werden. Die Umleitung kann zu jeder beliebigen anderen Rufnummer durchgeführt werden. Diese kann innerhalb des eigenen Funknetzes, in anderen Funknetzen oder in Festnetzen liegen. Abhängig von der Zieladresse fallen u.U. noch zusätzliche Gebühren für weitergeleitete Gespräche an.

Mobilbox

Die Mobilbox funktioniert im Prinzip wie ein persönlicher Anrufbeantworter, der vom Netzbetreiber zur Verfügung gestellt wird. Er erlaubt den Anrufern, ihre Nachrichten mündlich zu hinterlassen, welche dann vom eigentlichen Empfänger zu einem beliebigen Zeitpunkt abgerufen werden können. Ein Benutzer kann seine eigene Telefonnummer auf die Mobilbox umleiten und so quasi einschalten.

Anrufsperrung

Der Benutzer hat mehrere Möglichkeiten, ankommende oder abgehende Gespräche abhängig von der Rufnummer zu sperren (Ausnahme: Notruf). Dieser Dienst bezieht sich immer auf die D-Karte, nicht auf das Handy. Der Benutzer kann z.B. alle abgehenden Anrufe, alle abgehenden Auslandsanrufe oder alle ankommenden Anrufe bei Aufenthalt im Ausland sperren.

weitere Dienste

Die Netzbetreiber stellen zusätzlich noch weitere Dienste wie automatischer Rückruf, Gesprächskonferenz und geschlossene Benutzergruppen zur Verfügung. Zusätzlich werden von Fremdbetreibern noch Dienste wie z.B. Hotelreservierungen angeboten. Diese Dienste sollen hier aber nicht näher erläutert werden.

4 Ausblick

Zukünftigen Entwicklungen werden sich auf zwei Schwerpunkte konzentrieren: Weiterentwicklung von GSM und - parallel dazu - die globale Vernetzung mittels digitaler Funknetze.

Innerhalb von GSM sind Multimediaanwendungen und Videokonferenzen in (ferner?) Zukunft im Gespräch. In näherer Zukunft sind sicherlich half-rate-coded Geräte zu erwarten, die unterschiedliche Übertragungsraten erlauben und dies auch eventuell für Sprache ermöglichen [Mey95a]. Der Vorteil liegt aus Benutzersicht in günstigeren Übertragungspreisen und aus Netzbetreibersicht in besserer und größerer Auslastung des Netzes.

Die andere Sehensweise bezieht sich auf eine globale Vernetzung mittels digitaler Funknetze. Hier ist neben erweiterten Funktionalitäten bestehender Netze auch ein einheitliches System im Gespräch, welches momentan als „Universal Mobile Telecommunication System“ (UMTS) bezeichnet wird. Das UMTS soll einen weltweit kompatiblen mobilen Zugang zu Telekommunikationsdiensten ermöglichen und damit die persönliche Kommunikation erweitern. Die Frequenzbänder sollen bei 1885-2025 MHz (Bandbreite 140 MHz) und 2110-2200 MHz (90 MHz) liegen, wobei diese Bänder auch für andere Mobilkommunikationssysteme benutzt werden können, z.B. DECT 1880-1900 MHz. UMTS soll ein Dienstspektrum bis zu 2Mbit/s bereitstellen und möglichst weitgehende Kompatibilität zu B-ISDN (ATM) bieten. Ferner soll UMTS auch satellitengestützte Komponenten aufweisen, welche für die Überbrückung größerer Strecken sinnvoll sind. Die Frage der Übertragungsverfahren ist noch nicht abschließend geklärt, wobei die Entscheidung

zwischen TDMA und dem momentan in den USA verwendeten CDMA (Code Division Multiple Access = Codemultiplex) fallen wird [Mey95b], [FBJ⁺93].

Routingverfahren in Mobil-Netzen

Simon Oser

Kurzfassung

Die Entwicklung von immer kleineren Rechnern führte zu den heute weit verbreiteten Laptops. Da manche Benutzer dieser Geräte auch auf Reisen ununterbrochen mit dem Internet verbunden sein wollen, wurden in letzter Zeit mehrere neue Konzepte zur Anbindung von mobilen Rechnern an das Internet entwickelt. In diesem Beitrag sollen deshalb neue Routing-Ansätze vorgestellt werden, die im Hinblick auf den Einsatz in Mobil-Netzen entwickelt wurden. Dabei dürfen Probleme wie effizientes Routing und Netzwerksicherheit nicht unbetrachtet bleiben.

1 Einleitung

Mit der Entwicklung von immer leistungsfähigeren Laptops und dem stetig anwachsenden Internet mit seinen Diensten, wie z.B. E-Mail, FTP und WWW, wird der Wunsch der Benutzer dieser Geräte immer größer auch auf das Internet zugreifen zu können. Durch die Entwicklung und Einführung von Funk-Netzen steht dieser Idee fast nichts mehr im Wege. Ein mobiler Rechner braucht also nicht mehr mit einem Kabel direkt an ein bestehendes Netz angeschlossen werden, sondern er kann überall, sowohl daheim als auch auf Reisen, z.B. im Zug, benutzt werden.

Aufgrund dieser neuen Funktionalität wird ein neues Protokoll benötigt, das die Einbindung der Rechner ins Internet ermöglicht. Da das Internet geradezu von TCP/IP dominiert wird, sollte das neue Protokoll also eine Erweiterung des IP-Protokolls sein.

Mit der Einbindung ins Internet ergeben sich aber auch einige Probleme. Bisher ging man davon aus, daß ein Rechner immer seinen physikalischen Standort beibehält und somit stets am selben Netz angeschlossen ist. Dies spiegelt sich auch in der Vergabe von Netzwerkadressen wieder. Eine IP-Adresse besteht laut Definition nach [PC93] aus zwei Teilen, und zwar aus einem Teil, der das Netzwerk angibt, in dem sich der Rechner (Host) befindet, und einem zweiten Teil, der den Host identifiziert. Würde sich ein Rechner also aus dem (Sub-)Netzwerk hinausbewegen müsste sich folglich seine IP-Adresse ändern.

Allen hier vorgestellten Protokollen liegt die gleiche Idee zu Grunde. Man benutzt nicht mehr wie bisher im stationären Fall eine IP-Adresse, sondern ab sofort zwei IP-Adressen, die aber beide gültige IP-Adressen sind. Die erste Adresse ist eine konstante Adresse und wird dazu benutzt den Rechner eindeutig im Netz zu identifizieren. Diese Adresse wird auch als Home-Address bezeichnet. Wenn man Pakete an einen mobilen Rechner schicken will, benutzt man diese Adresse. Die zweite, variable Adresse, die auch als Care-of-Address bezeichnet wird, wird dazu benutzt den Rechner im Netz zu lokalisieren. Bewegt sich ein Rechner so müssen alle Pakete an diese Adresse umgeleitet werden.

Daraus folgt, daß es einen Mechanismus geben muß, der die konstante Home-Address auf die variable Care-of-Address abbildet und Pakete umleitet.

Alle Daten, die zum Abbilden der Adressen benötigt werden, sollten in einer sogenannten Routing-Information-Base abgelegt werden. Die Informationen, die zur Abbildung benötigt werden, müssen nicht zentral an einem Ort abgelegt sein, sondern können durchaus über das ganze Internet verteilt sein.

Rechner, die dieses neue Protokoll nicht implementiert haben, sollten trotzdem problemlos mit mobilen Rechnern kommunizieren können.

Des weiteren sollten aber auch keine Veränderungen an Endsystemen und Zwischensystemen (Routern) vorgenommen werden müssen, die keine speziellen Funktionen für mobile Rechner bieten.

Wie diese Probleme gelöst werden soll zuerst einmal am Beispiel eines einfachen Schemas, des sogenannten Basis-Schemas, erläutert werden. Daraufhin werden noch weitere Ansätze vorgestellt, die sowohl eine Routen-Optimierung von Paketen ermöglichen, als auch ein Umleiten von Paketen verhindern sollen.

Annahme: Es soll ab sofort angenommen werden, daß die Verbindung zwischen dem mobilem Rechner und dem Netzwerk drahtlos geschieht, auch wenn es die Möglichkeit gibt einen mobilen Rechner direkt ans Netz anzuschließen.

2 Basis-Schema

Die vorgestellten Konzepte sind alle sehr ähnlich, darum soll hier das sogenannte Basis-Schema vorgestellt werden, das in [CM95] definiert wurde und auch in den anderen hier vorgestellten Konzepten in abgewandelter Form zu finden ist.

2.1 Einleitung

Ein mobiles Netzwerk besteht aus einer Menge von mobilen Rechnern (Mobile-Hosts), deren Heimat-Stationen (Home-Agents) und weiteren Stationen mit speziellen Funktionen. Hier folgen einige Definitionen:

Mobile-Host: Ein Mobile-Host ist ein Rechner, der seinen physikalischen Standort wechseln kann. Wenn er seinen Anschlußpunkt wechselt, wechselt er aber zugleich auch seine physikalische IP-Adresse. Damit er jederzeit erreichbar ist benötigt er aber eine konstante Adresse. Diese wird auch „Heimat-Anschrift“ (Home-Address) genannt und gehört zu seinem Heimat-Netzwerk (Home-Network). Außerdem wird eine zweite Adresse benötigt, die seinen momentanen Standort angibt. Diese zweite Adresse wird auch nach [MJP95] Care-of-Address genannt und ist die Adresse des Foreign-Agents (siehe unten). Die Zusammenfassung von Care-of-Address und Heimat-Adresse soll hier als Bindung bezeichnet werden.

Home-Agent: Ein Home-Agent ist ein stationärer Rechner, der fest an das Netzwerk angeschlossen ist. Jeder Mobile-Host hat einen Home-Agent. Die IP-Adresse des Mobile-Host wird mit dem Home-Agent identifiziert, Pakete, die an den Mobile-Host adressiert sind, werden also zuerst dem Home-Agent ausgeliefert, der sie dann an den

Mobile-Host weiterleitet. Abstrakt gesehen hat der Mobile-Host beim Home-Agent sein zu Hause.

Ein Home-Agent besitzt eine Home-List in der seine eigenen Mobile-Hosts eingetragen sind.

Foreign-Agent: Ist ein Mobile-Host nicht mehr in der Reichweite seines Home-Agent, muß er einen anderen Anschlußpunkt finden. Foreign-Agents bieten Mobile-Hosts genau so einen Anschlußpunkt, d.h. eine drahtlose Verbindung zwischen dem Mobile-Host und ihm selbst, und ermöglichen somit eine Verbindung mit dem Netzwerk. Ein Mobile-Host benutzt die Adresse des Foreign-Agents als Care-of-Address.

Ein Foreign-Agent besitzt eine Visitor-List in der alle Mobile-Hosts eingetragen werden, für die er momentan als Anschlußpunkt dient.

Source-Host: Ein Rechner, der Pakete an einen Mobile-Host schickt, soll hier als Source-Host bezeichnet werden.

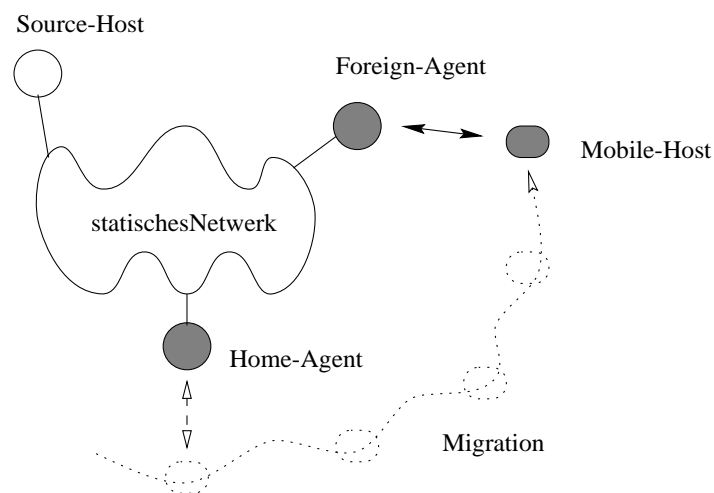


Abbildung 33. Mobiles Netzwerk.

Mobility-Agent: Home-Agents als auch Foreign-Agents werden als Mobility-Agents bezeichnet. Jeder Mobility-Agent besitzt eine Forwarding-List. Diese wird dazu benutzt, um ankommende Pakete an den Mobile-Host umzuleiten.

Wie gelangen nun Pakete an einen Mobile-Host?

2.2 Idee des Basis-Schemas

Benötigt ein Mobile-Host einen Anschluß an das Netzwerk, so sucht er einen Mobility-Agent und läßt sich bei ihm registrieren, wobei dem Home-Agent natürlich immer der Aufenthaltsort des Mobile-Host, d.h. die Care-of-Address, bekannt sein muß, falls der Mobile-Host nicht zu Hause ist.

Erhält ein Home-Agent ein Paket von einem Source-Host für einen seiner Mobile-Hosts und ist dieser gerade „zu Hause“, liefert er ihm das Paket aus. Ist er jedoch bei einem

Foreign-Agent registriert, müssen alle Pakete durch einen sogenannten Tunnel zur Care-of-Address, d.h. zur Adresse des aktuellen Foreign-Hosts, geschickt werden. Dieser reicht dann die Pakete an den Mobile-Host weiter, wenn er noch als sein Agent dient. Ist dies nicht der Fall werden die Pakete an den Home-Agent zurückgeschickt. Es gibt auch die Möglichkeit, daß er die Pakete an den neuen Foreign-Agent weitertunnelt, wenn ihm die Adresse bekannt ist.

Das Weiterleiten des Paketes wird auch als Tunneln bezeichnet. Dabei wird das Paket, das für einen Mobile-Host bestimmt ist, vom Home-Agent eingekapselt und mit Hilfe des normalen Routings an den Foreign-Agent weitergeleitet. Bevor der Foreign-Agent das Paket ausliefert muß es wieder in seinen ursprünglichen Zustand zurückverwandelt werden. Wie genau getunnelt wird regelt ein sogenanntes Einkapselungs-Protokoll (Encapsulation-Protocol). Näheres hierzu siehe auch Kapitel 5.3.

Jedesmal wenn sich ein Mobile-Host zu einem neuen Ort bewegt, ruft er eine Registrierungs-Prozedur auf, um die existierenden Bindungen zu erneuern. Der Home-Agent wird daraufhin benachrichtigt, daß der Mobile-Host seinen Standort gewechselt hat. Erhält er diese Nachricht so erneuert er seine Forwarding-List. Des weiteren muß der alte Foreign-Host über Standortwechsel informiert werden, damit er den Mobile-Host aus seiner Visitor-List löscht. Und schließlich muß auch der neue Foreign-Host den Mobile-Host in seine Visitor-List eintragen.

Ob die neue Adresse dem alten Foreign Host bekannt gegeben wird oder ob der Mobile-Host einfach nur aus der Visitor-List gelöscht wird ist von Protokoll zu Protokoll verschieden.

2.3 Beispiel

Wie oben erwähnt kommt ein von einem Source-Host stammendes Paket an einen Mobile-Host bei seinem Home-Agent an. Dieser tunnelt daraufhin das Paket an den Foreign-Agent, der dann das Paket dem Mobile-Host ausliefert. Siehe hierzu Abbildung 34.

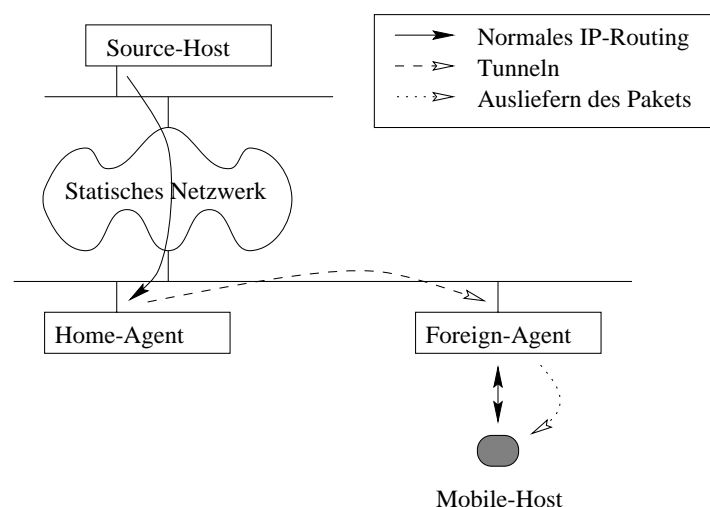


Abbildung 34. Basis-Schema.

3 Local-Region-Control und Patron-Control

Dieser Abschnitt beruht auf dem Artikel von Gihwan Cho und Lindsay F. Marshall [CM95]. Er beschreibt wie mit Hilfe von Local-Regions und Patrons ein optimaleres Routing und eine bessere und effizientere Verteilung der Routinginformation etabliert werden kann. Was Local-Regions und Patrons sind soll hier erklärt werden.

3.1 Motivation

Der Schlüssel zur Kommunikation mit mobilen Rechnern ist, wie oben beschrieben, die Bildung eines Tunnels zwischen dem Home-Agent und dem Foreign-Agent.

Angenommen Herr X. befindet sich gerade in den USA und sein Home-Agent ist in Karlsruhe an das Netzwerk angeschlossen, so müssen in diesem Fall alle Pakete an ihn, laut Basis-Schema, durch den Home-Agent getunnelt werden. Dies ist nicht gravierend, wenn der Source-Host auch in Karlsruhe ist. Ist der Source-Host aber in seiner unmittelbaren Umgebung, so müssen alle Pakete zuerst einmal über den großen Teich nach Europa und dann wieder zurück in die USA.

Im schlimmsten Fall muß also das IP-Paket zum Home-Agent geschickt werden, von wo aus es zum Mobile-Host getunnelt wird. Diese Art von Routing nennt man auch Triangle-Routing und kann sehr ineffizient sein. Siehe auch Abbildung 35.

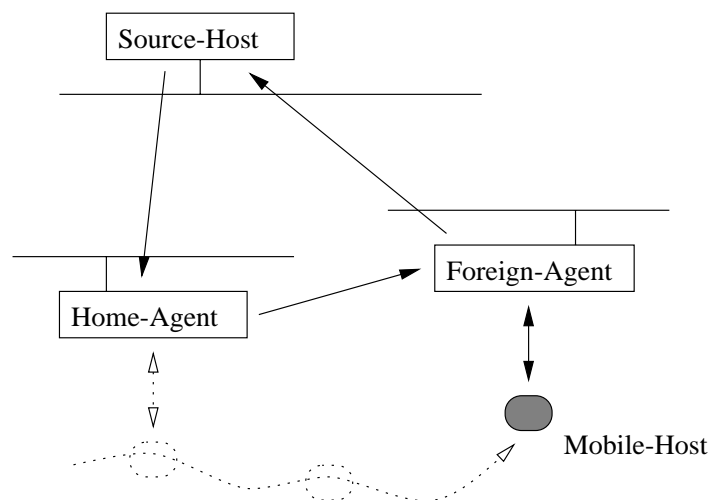


Abbildung 35. Triangle-Routing.

Weiß aber der Source-Host oder ein auf dem Weg liegender Router wo sich der Mobile-Host befindet, so kann dieser natürlich selbst das Paket an den Foreign-Agent tunneln, d.h. ohne Umweg über den Home-Agent, wie wir beim Redirection-Tunneling in Kapitel 3.3 sehen werden.

Deshalb hängt die Effizienz entscheidend von der Verteilung der Routing-Information im Netz ab, wobei man natürlich beachten muß, daß eine unbedachte Verteilung der Informationen eventuell Netzwerkressourcen verschwendet.

Was hier vorgestellt werden soll ist ein Konzept, das sowohl optimales Routing als auch so wenig Updates der Routinginformationen benötigt wie möglich.

Anmerkung: Man sollte beachten, daß Pakete, die der Mobile-Host an einen stationären Rechner verschickt, immer optimal geroutet werden.

3.2 Lokalität in Mobil-Netzen

Wenn sich ein Mobile-Host in Bewegung befindet, dann wechselt er des öfteren seinen Mobility-Agent und produziert damit eine große Menge Updateinformationen.

Intuitiv kann eine lokale Region (Local-Region) für einen Mobile-Host definiert werden, als diejenigen Subnetworks, in denen er sich oft bewegt, wobei die Subnetworks weglassen werden in denen er selten oder nie anzutreffen ist.

Für jeden Mobile-Host existiert also eine andere lokale Region, die sich nach seinen Aufenthaltsorten richtet. Die lokale Region eines Mobile-Host kann entweder dadurch festgelegt werden, daß man seine Bewegungen aufzeichnet und seine lokale Region rechnergestützt bestimmt, oder indem man den Benutzer fragt, wo er sich oft aufhält, und dann die lokale Region manuell eingibt.

Ein Host kommuniziert meistens nur mit einer beschränkten Anzahl Source-Hosts. Wird nur an die Source-Hosts ein Routing-Update verschickt, die mit dem Mobile-Host zur Zeit am meisten kommunizieren, so verringert sich die Anzahl Updates und somit die Verschwendung von Netzwerk-Ressourcen. Aus diesem Grund ist es praktisch wenn man sogenannte Patron-Hosts einführt. Patron-Hosts sind Source-Hosts von denen die meisten Pakete an einen bestimmten Mobile-Host ausgehen und die damit sehr wahrscheinlich auch weiterhin die meisten Pakete an diesen verschicken. Diesen soll immer der genaue Standort des Mobile-Host mitgeteilt werden, damit sie ihre Pakete an den Mobile-Host selbst tunneln können.

Es ist offensichtlich, daß man sich nur die Source-Hosts als Patron-Hosts merkt, die nicht innerhalb der lokalen Region befinden. Warum ?

3.3 Locale-Region-Control und Redirection-Tunneling

Wird ein Mobile-Host das erste Mal an das Netzwerk angeschlossen, wird die lokale Region, in der er sich am meisten aufhalten wird, festgelegt.

Es wird dann eine hierarchische Beziehung zwischen den Mobility-Agents und den Mobility-Routern, die die lokale Region ausmachen, vorgegeben. Diese Beziehung wird als ein sogenannter Redirection-Tree dargestellt.

Ein Mobility-Router besitzt die gleiche Funktionalität wie ein Mobility-Agent, er hat die Fähigkeit Pakete an einen Mobile-Host, der innerhalb seines Netzwerks registriert ist, zu tunneln.

Die Wurzel dieser Hierarchie bildet ein Mobility-Router, der die spezielle Aufgabe hat Pakete von oder zur lokalen Region umzuleiten d.h. zu tunneln. Dieser Mobility-Router

wird als Redirection-Agent bezeichnet. Der Redirection-Agent wird von einem Netzwerkadministrator bestimmt, der mit der Netzwerktopologie einer bestimmten Domain vertraut ist.

Der Redirection-Agent besitzt eine Liste, die Redirection-List, in der alle Standorte der Mobile-Hosts vermerkt sind, die innerhalb seines Netzwerks ihren Home-Agent haben. Jedesmal wenn ein Mobile-Host seinen Standort wechselt, muß er eine zusätzliche Bindung, sowohl in seinem Redirection-Agent als auch, wie im Basis-Schema beschrieben, in seinem Home-Agent und seinem vorhergehenden Foreign-Agent, erzeugen. Der Redirection-Agent benutzt die Information in der Redirection-List, um Pakete, die ihn passieren, zu den aktuellen Standorten der Mobile-Hosts umzuleiten.

Die Rolle des Redirection-Agents innerhalb der lokalen Region kann am besten mit Hilfe des Redirection-Trees beschrieben werden.

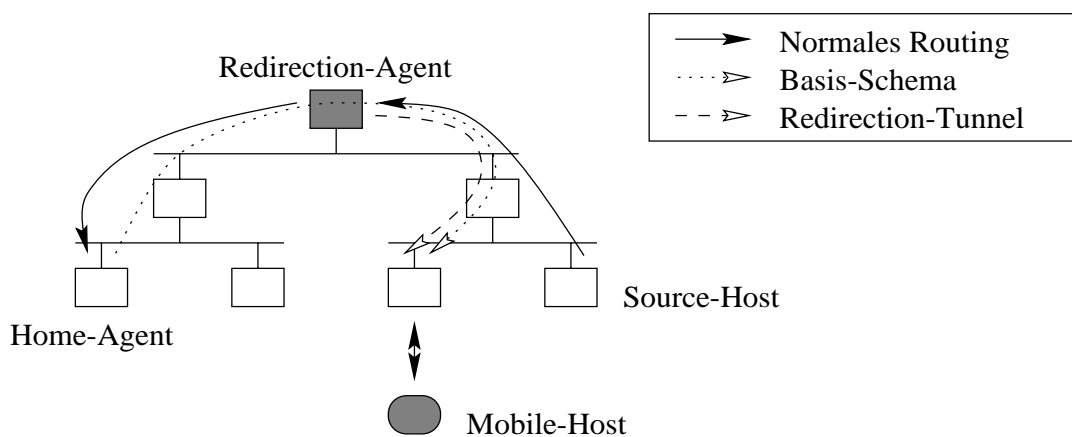


Abbildung 36. Redirection-Tunneling.

Falls der Source-Host im gleichen Unterbaum wie der Destination-Host (Mobile-Host) ist, und der Home-Agent sich aber im anderen Unterbaum befindet, dann werden die Pakete vom Redirection-Agent abgefangen und direkt an den Mobile-Host umgeleitet, ohne daß diese den Home-Agent passieren müssen. (Siehe Abbildung 36.)

Da das Schema noch diese zweite Bindung aufrechterhält, ist der zurückgelegte Weg hier kürzer als beim Basis-Schema.

Falls der Mobile-Host sich innerhalb der lokalen Region befindet, fängt der Redirection-Agent alle Pakete, die von Source-Hosts außerhalb der lokalen Region stammen, ab und leitet diese direkt um zum Mobile-Host. Da es den Redirection-Agent gibt, kann sich ein Mobile-Host innerhalb der lokalen Region frei bewegen ohne seinen genauen Standort Source-Hosts außerhalb bekanntgeben zu müssen.

Es ist offensichtlich, daß sich eine Geschwindigkeitssteigerung ergibt als auch, daß das lokale Netz nicht mehr unnötig belastet wird.

3.4 Patron-Hosts

Betrachtet man jetzt aber den in der Motivation beschriebenen Fall, d.h. Mr. X hält sich in den USA auf und sein Kommunikationspartner befindet sich direkt neben ihm, dann bewirkt ein Redirection-Agent relativ gesehen keine wesentliche Verbesserung. Das lokale Netz wird zwar entlastet, aber dennoch müssen alle Pakete, die von seinem Partner ausgehen, einen enormen Umweg benutzen. (Erinnerung: Triangle-Routing)

Deshalb werden die oben schon vorgestellten Patron-Hosts eingeführt.

Ist der Mobile-Host innerhalb der Locale-Region muß dem Patron-Host der genaue Standort des Mobile-Host nicht bekannt sein, da alle Pakete vom Redirection-Agent abgefangen und umgeleitet werden. Dieser Weg ist schon optimal!

Ist ein Mobile-Host hingegen außerhalb der lokalen Region so kann ein Patron-Host mit dem Mobile-Host direkt kommunizieren, ohne Umweg über den Redirection-Agent. Siehe auch Abbildung 37.

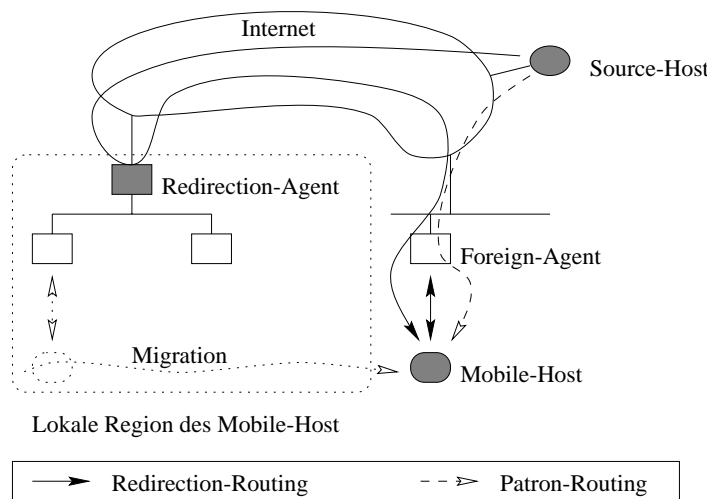


Abbildung 37. Patron-Host-Konzept.

Dies geschieht dadurch, daß der Patron-Host, wie schon oben erwähnt, seine Pakete selbst an den Mobile-Host tunnelt.

Source-Hosts, die mit einem Mobile-Host regelmäßig kommunizieren, sollten immer den neusten Standort mitgeteilt bekommen, wenn sich der Mobile-Host außerhalb der lokalen Region befindet.

Wenn ein Mobile-Host seine lokale Region verläßt oder betritt registriert er sich bei einem Mobility-Agent und sendet Registrierungspakete zum Home-Agent, zum alten Foreign-Agent und natürlich auch zum Redirection-Agent. Der Redirection-Agent kann feststellen, ob sich der Mobile-Host noch in der lokalen Region befindet. Tut er es nicht, verständigt er den Mobile-Host, der dann seinen aktuellen Standort allen Hosts in der Patron Liste bekanntgibt.

Da sich die Patron-List aber auch ändern kann, muß auch auf die Konsistenz der Location-Caches der Patrons geachtet werden. Ein Mobile-Host darf seine Patron-List nur dann

verändern, wenn er sich in der lokalen Region befindet. Beachtet man diese Regel nicht und löscht einen Patron-Host aus der Liste, so nimmt der Location-Cache des Patrons einen inkonsistenten Zustand an.

3.5 Forwarding-Pointer

Verläßt ein Mobile-Host die lokale Region, dann sendet er an alle Patrons seinen Standort. Ab diesem Zeitpunkt werden Forwarding-Pointers benutzt, um den Host von seinem ersten Foreign-Agent, der außerhalb der lokalen Region liegt, zu seinem aktuellen zu verfolgen. Dies kann durchaus ineffizient sein wie Abbildung 38 zeigt.

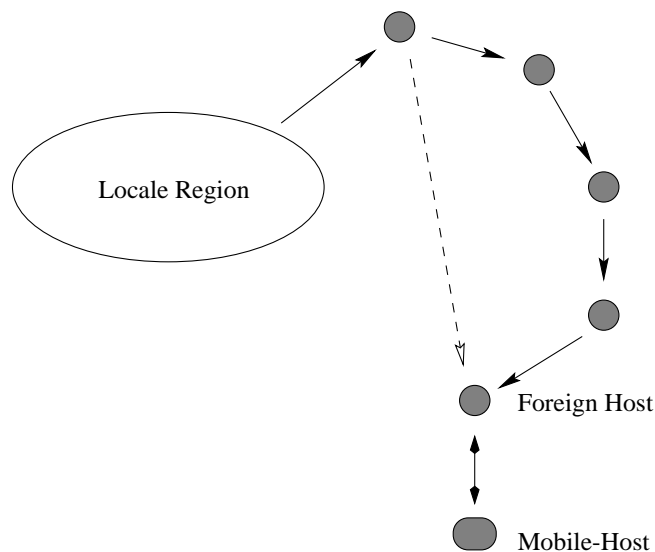


Abbildung 38. Erneuerung der Forwarding-Pointer.

4 Friend-Networks

Dieser Teil betrachtet die in [Yua95] vorgestellten Friend-Networks.

Wie im Basis-Schema oben schon vorgestellt, besteht hier ein mobiles Netzwerk auch aus einer Menge von mobilen Rechnern, den Mobile-Hosts, und einer Menge von statischen Rechnern, die hier als Mobility-Support-Router und den Base-Stationen bezeichnet werden. Diese sollen es den mobilen Rechnern ermöglichen mit Rechnern zu kommunizieren, ohne das bestehende Netzwerk grundlegend zu verändern.

Ein Mobility-Support-Router erfüllt hier die gleiche Funktionalität wie ein Foreign-Agent und Home-Agents werden hier als Base-Stationen bezeichnet.

4.1 Anforderungen an die Infrastruktur

Da die existierende Infrastruktur nicht verändert werden kann, muß die Unterstützung durch die Mobile-Hosts, die Mobility-Support-Router und die Base-Stationen übernommen werden.

Hier sollen die benötigten Funktionalitäten nach [Yua95] aufgelistet werden. Was benötigt wird sind also:

- Ein sogenanntes Beaconing-Protocol (Leuchtfeuer-Protokoll), das es dem Mobile-Host ermöglicht eine Verbindung zu einer beliebigen Base-Station aufzubauen.
- Ein Einkapselungsprotokoll, das es ermöglicht Pakete für einen Mobile-Host zum momentanen Mobility-Support-Router zu tunneln.
- Ein Routing-Information-Update-Protokoll, um die Standortinformation der Mobile-Hosts zwischen den Mobility-Support-Routern auszutauschen
- Ein Sender-Umleitungs-Mechanismus (Sender-Redirect-Mechanism), der es dem Mobility-Support-Router erlaubt den Source-Host die neue Adresse des Mobile-Hosts mitzuteilen.

Diese Anforderungen sind natürlich einleuchtend, wobei die letzte Anforderung nur dann funktioniert wenn der Source-Host diese Funktionalität unterstützt. Wie genau das Beaconing-Protocol funktioniert wurde hier leider nicht näher erläutert. Man kann hier gut sehen, wie die Funktionalität in mehrere Teil-Protokolle aufgegliedert wird.

4.2 Definition eines „Friend-Networks“

Gihwan Cho und Lindsay F. Marshall haben in [CM95] versucht die Routen dadurch zu optimieren, daß sie den Patron-Hosts auch noch zusätzlich mitteilten, wo sich der Mobile-Host aufhält. Ruixa Yuan geht in [Yua95] einen anderen aber dennoch ähnlichen Weg. Er sagt, die Netzwerkbelastung kann einfach verringert werden indem man Updates nur an Router verschickt, die diese wirklich dringend benötigen.

Deswegen werden Rechner aus einem Teil-Netzwerk, die mit einem Mobile-Host kommunizieren, logisch zusammengefaßt und als Friend-Network bezeichnet.

Beim Konzept der Friend-Networks werden Updates nur an Mobility-Support-Router verschickt, die zum Friend-Network gehören.

Definition: Ein Teil-Netzwerk soll als (Teil-) Friend-Network bezeichnet werden, falls gilt:

1. es besteht ein großes Verkehrsaufkommen zwischen dem Netzwerk und dem Mobile-Host und
2. das Netzwerk ist ein „vertrauenswürdiges“ Netzwerk und besitzt einen Mobility-Support-Router.

Wobei ein Netzwerk als vertrauenswürdig eingestuft wird, wenn von ihm keine Gefahr ausgeht, d.h. wenn kein Mißbrauch stattfindet. Ein Mobility-Support-Router liefert natürlich auch indirekt Informationen über den Aufenthaltsort von Personen, wenn man ihren Rechner kennt. Herr Clinton wird also das Netzwerk von Herrn Jelzin nicht als vertrauenswürdig bezeichnen.

Durch die Friend-Networks kann somit die effiziente Verbreitung der Routinginformation und die Netzwerksicherheit ermöglicht werden.

4.3 Routing Updates

Falls ein Mobile-Host sich in ein neues Netzwerk bewegt, ermöglicht ihm das Beaconing-Protocol die Netzwerkanbindung mit einem Mobility-Support-Router. Der neue Mobility-Support-Router dient dann als Agent des Mobile-Host.

Das Erneuern der Routinginformation erfolgt dann folgendermaßen:

- Der neue Mobility-Support-Router informiert den Home-Mobility-Support-Router über den neuen Standort.
- Der Home-Mobility-Support-Router verschickt ein Routing-Update an alle Mobility-Support-Router der Friend-Networks.
- Der Home-Mobility-Support-Router verschickt ein Update an alle Hosts mit denen der Mobile-Host gerade kommuniziert.

4.4 Mobile-Traffic-Routing

Ein Paket, das für einen Mobile-Host bestimmt ist, wird an seine Heimatadresse verschickt. Trifft das Paket auf einen Mobility-Support-Router, wird es von diesem eingepackt und an den aktuellen Mobility-Support-Router geschickt. Dieser reicht dann das Paket an den Mobile-Host weiter. Siehe auch Abbildung 39. Falls dieser schon nicht mehr für ihn zuständig ist, wird das Paket an den Home-Mobility-Support-Router weitergeleitet.

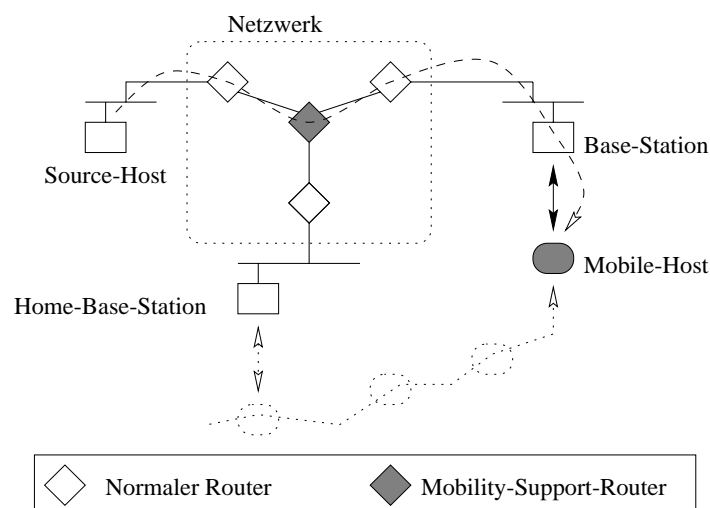


Abbildung 39. Mobility-Support-Router-Konzept.

5 Das Internet-Mobile-Host-Protocol

Das Internet-Mobile-Host-Protocol (IMHP), das in [MJP95] vorgestellt wird, ist kompatibel zum TCP/IP-Protokoll und ermöglicht sowohl eine Routenoptimierung als auch eine

Beglaubigung (Authentication, dt. Authentizität) aller Managementpakete, um eventuellen Attacken, wie z.B. illegales Umleiten, vorzubeugen.

5.1 Infrastruktur und Routing

Die IMHP-Architektur besteht aus Mobile-Hosts, Local-Agents, Cache-Agents und Home-Agents.

Der Mobile-Host ist hier genauso wie oben definiert. Er soll aber trotzdem noch einmal hier aufgeführt werden:

Mobile-Host: Ein Mobile-Host kann sich mit Hilfe zusätzlicher Software frei bewegen. Er besitzt eine konstante Heimatadresse, die zu seinem Home-Network gehört, wie bei jedem anderen Host auch. Andere können diese verwenden, um dem Mobile-Host Pakete zu schicken unabhängig davon wo dieser sich gerade befindet. Außerdem besitzt er eine zweite variable Adresse' die seinen physikalischen Standort angibt.

Local-Agent: Wenn ein Mobile-Host Verbindung mit einem Netzwerk aufnimmt, dann muß er sich bei einem Local-Agent registrieren.

Jeder Local-Agent besitzt eine Visitor-List in der alle Mobile-Hosts eingetragen sind, die sich gerade bei ihm befinden. Der Mobile-Host benutzt diese Liste, um Pakete an Mobile-Hosts, die sich gerade bei ihm registriert haben, weiterzugeben.

Die Adresse des Local-Agents dient als Standort-Adresse des Mobile-Host und wird Care-of-Address genannt.

Ein Local-Agent löscht einen Mobile-Host nach einer bestimmten Zeitdauer aus seiner Visitor-List, wenn dieser sich nicht nach einer bestimmten Zeitdauer wieder registriert hat.

Immer wenn sich ein Mobile-Host bei einem Local Agent registriert, muß er alle Local-Agents angeben bei denen er evtl. noch in der Visitor-List steht. Diese Bekanntmachung wird dazu benutzt den Mobile-Host aus der Visitor-List alter Local-Agents zu löschen.

Cache-Agent: Ein Cache-Agent lernt Bindungen durch das IMHP-Binding-Management-Protocol. Falls er ein Paket erhält, für das er eine Bindung in seinem Cache hat, tunnelt er es an die Care-ofAddress.

Ein Eintrag wird nach einer bestimmten Zeit wieder gelöscht, d.h. er muß in zyklischen Abständen erneuert werden. Ein Eintrag wird aber auch gelöscht, wenn der Mobile-Host sich zu Hause befindet.

Jeder Knoten, der die Kommunikation mit Mobile-Hosts optimieren will, sollte also als Cache-Agent agieren. Viele Local-Agents werden auch die Funktionalität eines Cache-Agents besitzen. Falls ein solcher Cache-Agent benachrichtigt wird, daß ein Mobile-Host aus seiner Visitor-List jetzt bei einem neuen Local-Agent ist, wird er ihn aus seiner Visitor-List löschen und einen Eintrag in seinem Location-Cache erzeugen, um evtl. bei ihm verspätet ankommende Pakete weiterzuleiten.

Home-Agent: Jeder Mobile-Host besitzt einen Home-Agent, der mit dem Home-Network verbunden ist. Der Home-Agent besitzt eine Home-List, in der alle seine Mobile-Hosts eingetragen sind. Für diese dient er als Cache-Agent und für alle anderen zusätzlich als Local Agent.

Wenn sich ein Mobile-Host bei einem Local-Agent registriert, muß er sich auch bei seinem Home-Agent melden, der dann seine Care-of Address in seinem Location-Cache einträgt.

Jeder Eintrag im Location-Cache oder in der Visitor-List wird nach einer bestimmten Zeit wieder gelöscht. Das Zeitintervall wird bei der Registrierung durch Verhandlung festgelegt. Der Mobile-Host muß sich also immerwieder registrieren, am besten noch bevor das Zeitintervall abgelaufen ist. Ist das Zeitintervall abgelaufen und hat der Mobile-Host sich nicht erneut bei einem Local-Agent registriert, wird angenommen, daß er sich bei seinem Home-Agent befindet.

5.2 Authentizität der Managementpakete

Der Hauptziel dieses Protokolls ist aber nicht nur eine Wegoptimierung, sondern die Authentizität aller Managementpakete, d.h. es muß sichergestellt werden, daß ein Mobile-Host, der sich als ein bestimmter Mobile-Host ausgibt, auch wirklich dieser ist. Ohne Authentizität könnten Pakete einfach umgeleitet werden, ohne daß der Host, der Pakete abfangen will, sich auf dem direkten Weg zwischen Quelle und Ziel befinden muß.

Fügt man Authentizität zu einem mobilen Netzwerkprotokoll hinzu, dann muß jeder Knoten seine Bindungen, für die er Einträge besitzt, beglaubigen lassen. Dazu wird allerdings ein System benötigt, das mit einem Public-Key-Verfahren funktioniert. Da es aber momentan noch politische und rechtliche Schwierigkeiten gibt, kann man solch ein System noch nicht benutzen.

Das IMHP wurde deshalb so konzipiert, daß es von einem solchen System in Zukunft Gebrauch machen kann. Solange dies nicht möglich ist, benutzt es ein einfaches System, das unter anderem auf Zufallszahlen beruht und sich auf den momentanen Sicherheitsvorkehrungen des Internets abstützt.

Die genaue Funktionsweise dieses Systems soll hier nicht weiter betrachtet werden.

5.3 Einbindung in das IP-Protocol

Um ein Paket zu tunneln wird ein kleiner Tunneling-Header zwischen den IP-Header und den Transport-Header eingefügt. Siehe auch Abbildung 40.

Die Originaldaten des alten IP-Headers werden in den IMHP-Tunneling-Header kopiert und der IP-Header wird so abgeändert, daß das Paket an den Local-Agent geschickt wird. Das Paket benutzt dann normales IP-Routing, um den Local-Agent zu erreichen. Der Local-Agent entfernt den IP-Header bevor er das Paket ausliefert. Das IMHP-Protocol fügt nur 8 oder 12 Bytes Overhead zu jedem Paket hinzu, um das Paket nicht unnötig zu vergrößern.

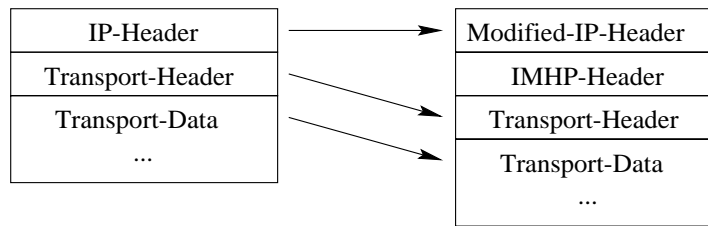


Abbildung 40. IMHP-Tunneling-Header.

6 IP-Mobility-Support-Protocol

6.1 Einleitung

Das in [Per95] vorgestellte IP-Mobility-Support-Protocol ist ebenfalls ein TCP/IP kompatibles Protokoll. Es weist große Ähnlichkeit zu dem in [MJP95] vorgestellten Internet-Mobile-Host-Protocol. Trotz der Ähnlichkeiten, soll hier auf Unterschiede und Gemeinsamkeiten eingegangen werden.

Ein Mobile-Host wird hier als Mobile-Node bezeichnet. Home-Agent und Foreign-Agent haben die gleiche Bedeutung wie oben.

6.2 Gemeinsamkeiten

Das IP-Mobility-Support Protocol benutzt wie das IMHP auch eine Authentikation der Managementpakete, um sich vor Attacken zu schützen. Des weiteren benutzt es einen Timer. D.h. ein Mobile-Node muß sich noch bevor dieser abgelaufen ist neu registrieren. Dieser ist notwendig damit alte, falsche Einträge gelöscht werden. Außerdem wird ein Einkapselungs-Protokoll, das dem des IMHP recht ähnlich ist.

6.3 Unterschiede

Ein Unterschied zum IMHP und auch zu allen anderen bereits vorgestellten Protokollen ist, daß ein Mobile-Host sich nicht nur über einen Foreign-Agent an das Netzwerk anschließen kann, sondern auch direkt. Wie eine solche direkte Verbindung aussieht wird dabei nicht erwähnt, es wird nur gesagt, daß es diese Möglichkeit auch gibt.

Die Care-of Address gibt den momentanen Aufenthaltsort des Mobile-Node an, wobei diese entweder mit einem Foreign-Host in Verbindung gebracht wird oder dynamisch an den Mobile-Host vergeben wird.

6.4 Annahmen und Anforderungen

Eine Organisation, die einen Mobile-Node besitzt weist diesem eine IP-Adresse zu. Diese kann unabhängig von seinem Standort benutzt werden.

Des weiteren wird angenommen, daß ein Mobile-Node nicht öfter als ein Mal pro Sekunde seinen Agent wechselt. Dies ist notwendig, um den Verkehr zu begrenzen.

7 Bewertung und Ausblick

Die Vielzahl ähnlicher Ansätze für mobile Rechner hat Bewegung ins Internet gebracht. Mit dem IP-Mobility-Support-Protocol gibt es zwar schon einen weit entwickelten und durchdachten Internet-Draft, aber wie auch alle anderen Ansätze deckt er noch nicht alles ab. So behandelt das IP-Mobility-Support-Protocol noch keine Routen-Optimierung, was bedeutet, daß Pakete in der Praxis große Umwege in Kauf nehmen müßten. Aber auch die anderen Ansätze sind noch nicht so ausgereift, daß sie in der Internet-Welt bestehen könnten. Wie man sieht, ist noch viel zu tun und alles momentan noch im Fluß. Dennoch bleiben Fragen offen. Sind die vorgestellten Protokolle wirklich so kompatibel zum IP-Protocol wie vorgesehen oder entstehen vielleicht doch nicht vorhersehbare Risiken dadurch, daß ein Router das Mobilitätsprotokoll nicht spricht? Müssen wir vielleicht sogar befürchten, daß Pakete in Zukunft einfacher abgefangen und umgeleitet werden können als bisher? Wir werden sehen.

Multicast-Routing im Internet

Frank Schadt

Kurzfassung

Im Internet besteht zur Zeit nur die Möglichkeit, Daten an einen bestimmten Empfänger (Unicast) oder an alle angeschlossenen Rechner (Broadcast) zu übertragen. Aufgrund neuerer Anwendungen (z.B. Videokonferenzen) besteht jedoch ein großer Bedarf Nachrichten an eine Gruppe von Empfängern (Multicast) zu senden. Für diese neue Kommunikationsform Multicast müssen die bestehenden Routingprotokolle angepaßt bzw. neue Protokollspezifikationen entworfen werden. Die beiden Protokolle DVMRP und MOSPF entstanden durch Erweitern von bestehenden Protokollen. Sie werden heute auch zur Multicast-Übertragung eingesetzt. Neben diesen zwei Protokollen gibt es weitere Protokollspezifikationen, die gegenüber DVMRP und MOSPF einige Vorteile aufweisen. Hier sind besonders CBT und PIM zu nennen. In dieser Arbeit werden die MC-Protokolle vorgestellt.

1 Einführung

Durch die steigende Anzahl von Internet-Benutzern und den Bedarf nach neuen Anwendungen (z.B. Videokonferenzen), gewinnt die Gruppenkommunikation zunehmend an Bedeutung. In dieser Einführung wird die Kommunikationsform Multicast und ihre Vorteile gegenüber Unicast dargestellt. Danach wird auf die Architektur des Internet eingegangen und das Multicast-Backbone vorgestellt.

1.1 Kommunikationsformen

Bislang bieten die meisten Netzwerkprotokolle den höheren Schichten nur einen Übertragungsdienst an, durch den genau zwei Netzknoten miteinander Daten austauschen können. Jeder Knoten in einem solchen System kann zu einem bestimmten Zeitpunkt eine Nachricht nur an genau einen anderen Knoten senden. Diese Kommunikationsform, bei der genau ein Sender an genau einen Empfänger senden kann, nennt man Unicast (Bild 41).

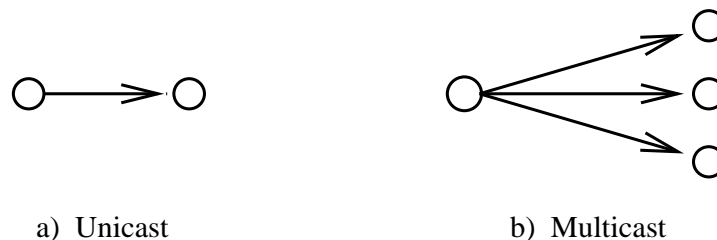


Abbildung 41. Kommunikations-Formen

In der Protokollfamilie TCP/IP, die im Internet verwendet wird, ist ursprünglich nur der Unicast-Übertragungsdienst spezifiziert. Neue Kommunikationsformen (z.B. Video-

Konferenzen) erfordern es jedoch, daß eine Nachricht von einem Sender an eine Gruppe von Empfängern gesendet wird. Diese Art von Kommunikationsbeziehung nennt man Multicast (Bild 41). Man kann einen Multicast-Dienst auch durch Nutzung mehrerer Unicast-Dienste nachbilden, indem das gleiche Paket n-mal kopiert wird und jeweils eine Kopie an unterschiedliche Gruppenmitglieder gesendet wird. Die effizientere Methode ist die Nutzung eines Multicast Dienstes, bei dem genau ein Paket an alle n Gruppenmitglieder gesendet wird [Byn]. Der Multicast Dienst hat den Vorteil, daß über jede Netzverbindung nur ein Paket pro Gruppenaufruf versendet wird. Im Gegensatz dazu werden bei mehrfacher Nutzung von Unicast über eine Verbindung bis zu n Pakete gesendet. Die Netzlast läßt sich beim Multicast - je nach Topologie des Netzes - um ein Vielfaches verringern. Ein weiterer Vorteil des Multicast-Dienstes gegenüber dem herkömmlichen Unicast-Dienst besteht in der Tatsache, daß viele Transportmedien im LAN-Bereich Multicast bzw. Broadcast auf Hardware- und Medienzugriffsebene unterstützen. Wenn z.B. bei einer Übertragung n Gruppenmitglieder durch ein Ethernet-LAN verbunden sind, wird bei Nutzung eines Unicast-Dienstes die Nachricht n-mal über das LAN übertragen. Bei Multicast-Übertragung würde nur ein Paket gesendet werden. Auch beim Versenden einer Nachricht über das Internet ist die Netzbelastung beim Multicast viel geringer als bei einem mehrfach angewandten Unicast-Dienst (Bild 42).

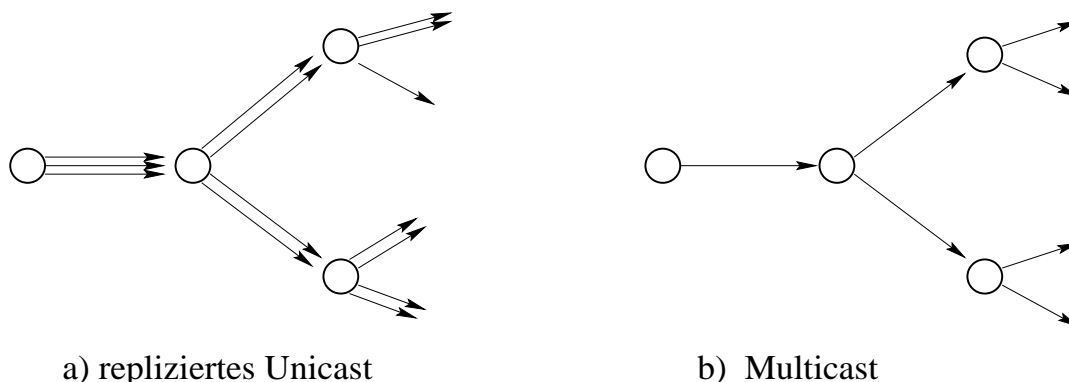


Abbildung 42. Netzauslastung

1.2 Architektur des Internet

Bestandteile des Internet:

Hosts: Als Host wird jedes, in das Internet integrierte Computersystem bezeichnet, welches über TCP/IP Protokolle kommuniziert. Auf einem Host laufen Anwendungsprogramme, die miteinander kommunizieren.

Netzwerke: Ein Netzwerk setzt sich aus mehreren Hosts zusammen. Im Internet gibt es die unterschiedlichsten Arten von Netzwerken. TCP/IP ist von der Netzwerktechnologie unabhängig.

Router: Ein Router verbindet mehrere Netzwerke miteinander zu einem Internet.

Das Internet verbindet tausende von Netzwerken miteinander. Somit ist es millionen von Computern möglich Informationen auszutauschen. Aufgrund seiner Größe und des ständig andauernden Wachstums ist das Internet in hierarchische Ebenen gegliedert. Auf der untersten Ebene werden die Hosts üblicherweise zu LANs zusammengeschlossen. Die LANs werden wiederum zu größeren Netzwerken zusammengeschlossen, die von Universitäten, staatl. Organisationen oder privaten Gesellschaften betrieben werden. Diese Subnetze werden auch autonome Systeme (AS) genannt. In einem AS wird gewöhnlich genau ein Routing-Protokoll benutzt. In verschiedenen autonomen Systemen können unterschiedliche Protokolle mit unterschiedlichen Adreßformaten, Leistungseigenschaften und Sicherheitsaspekten zum Einsatz kommen.

1.3 Das Multicast-Backbone (MBone)

Eine flächendeckende Unterstützung für Multicast-Dienste sollte so schnell wie möglich im Internet eingeführt werden. Das Problem ist jedoch, daß die meisten IP-Router nur unicast-fähig sind. Damit ein Router Multicast Daten übertragen kann, braucht er den Multicast-Daemon "mrouted", sowie eine Multicast-Unterstützung im Kernel des Rechners. Um trotzdem im Internet Multicast-Daten versenden zu können, errichtete man ein "virtuelles" Netz "über" dem Internet, das Multicast Backbone. Das MBone besteht nur aus multicastfähigen IP-Routern. Um nun die Multicast-Pakete von einem multicastfähigen Router (MC-Router) zum nächsten zu senden, müssen die nicht multicastfähigen Router überbrückt werden. Das wird im MBone durch das sogenannte "Tunneln" ermöglicht [Fis]. Dazu werden die MC-Pakete in IP-Datagramme eingepackt und zum nächsten MC-Router geschickt. Die dazwischen liegenden UC-Router können diese "Tunnel"-Pakete von anderen IP-Paketen nicht unterscheiden und leiten alle Pakete über Unicastverbindungen weiter. Die MC-Router erkennen am Header des IP-Pakets ob es sich um ein eingepacktes MC-Paket handelt (Bild 43).

Das MBone ist nur als experimentelles Netzwerk zu sehen, das in Zukunft im Internet aufgehen wird.

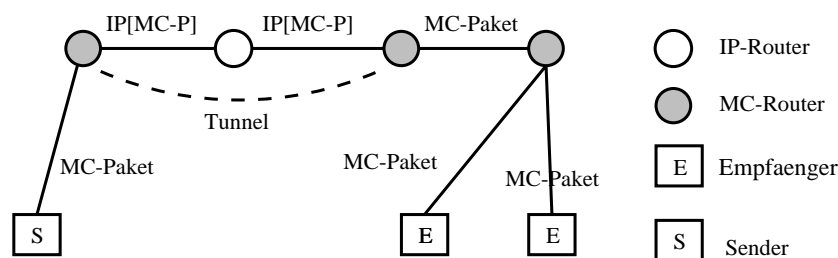


Abbildung 43. MBone und Tunneln

2 Grundlegende Mechanismen des Multicast-Routing

In diesem Abschnitt werden die grundlegenden Mechanismen des Multicast-Routing vorgestellt. Im folgenden Abschnitt werden dann spezielle MC-Protokolle betrachtet, die auf diesen Mechanismen aufbauen.

2.1 Fluten

Die einfachste Art und Weise eine Nachricht an eine Gruppe von Empfängern zu senden ist, die Nachricht über das gesamte Netz zu fluten. Wenn eine Quelle Q eine Nachricht an eine Gruppe G versenden will, wird diese Nachricht an alle Knoten im Netz gesendet. Bei diesem Verfahren ist sichergestellt, daß jeder potentielle Empfänger das Nachrichtenpaket erhält. Wenn ein Knoten ein Paket mit Multicast-Adresse empfängt, prüft er zuerst, ob er das Paket schon zuvor empfangen hat. Falls das Paket das erste Mal bei ihm vorbeikommt, leitet er es in allen Richtungen weiter, ausgenommen der Richtung, aus der er das Paket empfangen hat. Andernfalls verwirft er es, um ein Kreisen des Pakets zu vermeiden. Somit wird sichergestellt, daß das Nachrichtenpaket von allen Knoten im Netz empfangen wird. Dieser Mechanismus ist leicht zu implementieren. Die einzige Schwierigkeit besteht darin, festzustellen, ob ein Paket schon einmal empfangen wurde. Dazu kann jeder Router eine Liste mit den kürzlich empfangenen Paketen führen. Diese Liste kann bei einem Hochgeschwindigkeitsnetz recht lang werden.

Der Fluten-Algorithmus eignet sich besonders, wenn ein robustes Verfahren erforderlich ist. Die Nachteile des Verfahrens sind der grosse Speicherplatzaufwand und der Verbrauch von Netzressourcen.

2.2 Spannende Bäume

Im letzten Abschnitt wurde ein einfacher und robuster Algorithmus vorgestellt, der jedoch recht ineffizient mit Netzwerkressourcen umgeht. Nun wird eine Lösung vorgestellt, welche die Ressourcen besser ausnutzt als dies beim Fluten-Algorithmus der Fall war. Eine Topologie aus Routern und Subnetzen wird in eine schleifenfreie Anordnung gebracht, indem einige Router manche ihrer Verbindungen "ignorieren". Einige Verbindungen des Netzes werden benutzt, die anderen nicht. Es wird sozusagen ein "Overlay"-Netzwerk gebildet, welches aus allen Knoten und den nicht ignorierten Verbindungen des Netzes zusammengesetzt ist. Die Menge der gewählten Verbindungen bildet einen zyklusfreien Graphen, also einen Baum. Dieser Baum enthält alle Knoten des Netzwerks und spannt somit das Netzwerk auf. Aus diesem Grund nennt man ihn spannenden Baum (Bild 44).

Alle Router im Netz berechnen den spannenden Baum. Eine Multicast-Nachricht wird nun entlang des Baumes versendet. Jeder Router sendet ein ankommendes Paket auf allen seinen zum Baum gehörenden Anschlüssen weiter, mit Ausnahme des Anschlusses auf dem das Paket empfangen wurde. Das Netz wird somit nicht so stark belastet wie beim Fluten.

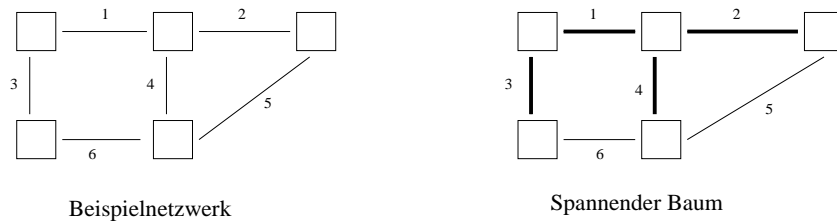


Abbildung 44. Beispielnetz und Spannender Baum

Trotzdem hat diese Technik zwei wesentliche Nachteile: Durch das Versenden der Pakete entlang eines Baumes wird der gesamte Verkehr auf die zum Baum gehörenden Verbindungen konzentriert. Weiter wird keinerlei Gruppenzugehörigkeit unterstützt. Ein Paket das auf diese Weise an eine Gruppe verschickt wird, kommt bei allen Knoten des Netzes an. Wegen dieser Nachteile werden Algorithmen bevorzugt, die gruppenspezifische Graphen berechnen und effizienter mit den Netzressourcen umgehen.

2.3 Reverse-Path Forwarding

Der Reverse-Path Forwarding (RPF) Algorithmus wird derzeit im Mbone verwendet. Jeder Router im Netz benutzt eine Routing Tabelle, in der die zum Weiterleiten von Paketen notwendige Information steht. Mit Hilfe der Routing-Tabelle berechnet der Router einen spannenden Baum pro Netzwerkquelle. Falls n Quellen eine Nachricht an die gleiche Gruppe senden, müssen für diese Gruppe n Bäume erstellt werden [Hui95]. In der einfachsten Form arbeitet RPF wie folgt:

Ein Router empfängt ein Multicast-Paket mit Quelle Q am Anschluß A . Nun überprüft der Router, ob er, aufgrund der Information aus seiner Routing-Tabelle, Pakete zu Q über Anschluß A senden würde. Anschluß A gehört in diesem Fall zum kürzesten Pfad in Richtung Quelle Q . In diesem Fall versendet er das Paket auf allen seinen Anschlüssen, mit Ausnahme von A . Andernfalls wird das Paket verworfen, weil der Router die Pakete der Quelle Q von dem Nachbarrouter erhält, der auf dem kürzesten Pfad liegt. Zu beachten ist bei diesem Verfahren, daß der kürzeste Weg vom Router zur Quelle Q berechnet wird. Dieser muß nicht mit dem kürzesten Weg von der Quelle Q zum Router übereinstimmen. Wenn das Netzwerk nicht symmetrisch ist, unterscheiden sich diese beiden Wege. Darum heißt das Verfahren auch Reverse Path Forwarding. RPF benutzt als einzige Ressourcen die Unicast-Routing-Tabelle. Um die Effizienz des Verfahrens zu steigern, sollte man jedoch nicht die Unicast-Routing-Tabelle verwenden. Besser wäre eine Berechnung des kürzesten Pfades von Quelle Q zum aktuellen Knoten.

Eine weitere Verbesserung des Verfahrens ist möglich, indem man "einen Schritt" vorausschaut. Der Router müßte wissen, ob er auf dem kürzesten Pfad zwischen einem Nachbarn und der Quelle Q liegt. Falls das nicht der Fall ist, braucht er das Paket nicht an den Nachbarn zu versenden, weil der es nach obigem Algorithmus sowieso verwirft. Weil RPF für jede Quelle einen spannenden Baum berechnet, werden Pakete über verschiedene Verbindungen verteilt. Dadurch wird das Netz gleichmäßiger ausgenutzt. Die einfache Form des RPF hat jedoch wie die bisher vorgestellten Verfahren den Nach-

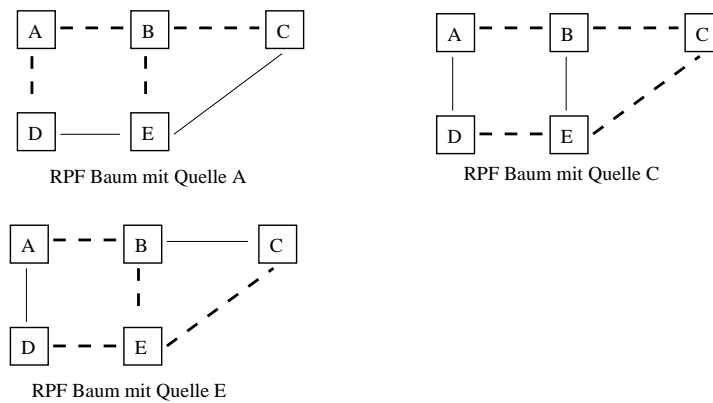


Abbildung 45. RPF Bäume einer Gruppe mit verschiedenen Quellen

teil, daß beim Aufbau des Baumes die Gruppenzugehörigkeit nicht beachtet wird und Pakete somit ins gesamte Netz gesendet werden. Durch Benutzung des Internet Group Managment Protokolls (IGMP) kann ein Router erkennen, ob in einem ihm angeschlossenen Subnetz Gruppenmitglieder vorkommen. Pakete werden vom Router nur in dieses Subnetz gesendet, wenn sich dort ein Gruppenmitglied befindet. Diese Vorgehensweise nennt man “truncated broadcast”.

2.4 RPF und Kürzen

Durch den Aufbau eines spannenden Baumes pro Quelle, werden durch RPF alle Pakete von jeder Quelle aus an alle im Netz vorhandenen Knoten versendet. Der Basis-RPF-Algorithmus wird jetzt erweitert, indem die Gruppenzugehörigkeit in den Routern gespeichert wird. Pakete werden dann nur den Baum entlang gesendet, falls ein Gruppenmitglied in diesem Ast des Baumes liegt. Das Ergebnis dieser Vorgehensweise sind minimale Bäume, die jedoch ständig aktualisiert werden müssen. Diese Version des RPF ist auch bekannt unter dem Namen “flood and prune”.

Alle Endknoten im Netz, die im RPF-Baum keine weiteren Knoten hinter sich haben, heißen Blätter. Falls sich in den Subnetzen, die an den Blattknoten angeschlossen sind, keine Gruppenmitglieder der Gruppe G befinden, sollte der Knoten auch keine Pakete empfangen, die an die Gruppe G gesendet werden. Er schickt eine “prune“-Nachricht an den Router, der ihm das Paket geschickt hat. Der Router speichert die Nachricht und sendet an diesen Knoten keine MC-Pakete mehr, die für die Gruppe G bestimmt sind. Falls er auf allen Anschlüssen eine “prune“-Nachricht für die gleiche Nachricht der Quelle Q an Gruppe G erhalten hat, will er selbst ebenfalls keine Pakete der Quelle Q an die Gruppe G empfangen. Auf diese Weise wird die “prune“-Nachricht den Baum “hinaufgereicht” (Bild 46).

Ein Nachteil des Verfahrens ist das erstmalige Versenden eines Paketes an alle Knoten im Netz (“fluten”). Weiter müssen die Router die “prune“-Nachrichten pro Gruppe und Quelle speichern. Dies kann bei großen Netzen und vielen Gruppen zu größeren Problemen führen.

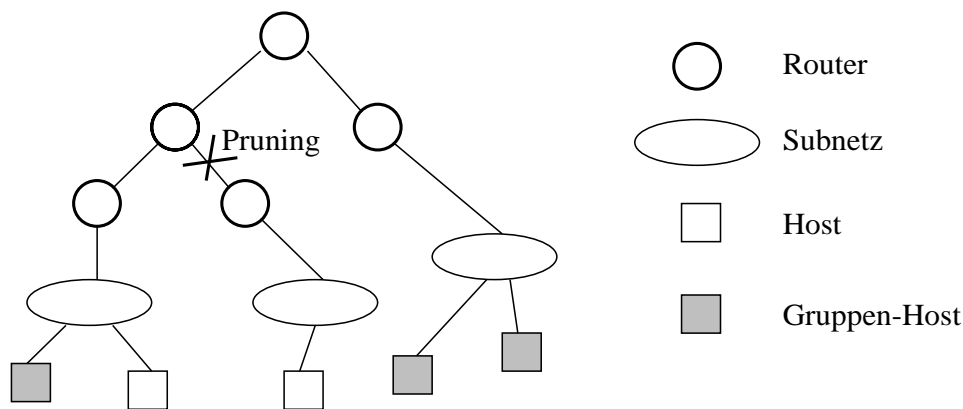


Abbildung 46. Pruning eines Baumes

3 Multicast Protokolle

In diesem Abschnitt werden einige Multicast-Protokolle vorgestellt. Sie verwenden die im vorigen Abschnitt vorgestellten grundlegenden Mechanismen. Zuerst werden zwei Protokolle besprochen, die zur Zeit schon im Einsatz sind: DVMRP und MOSPF. Danach betrachten wir mit CBT und PIM zwei neuere Protokollspezifikationen, welche noch nicht in einem größeren Rahmen benutzt werden.

3.1 DVMRP

DVMRP (Distance Vector Multicast Routing Protocol) ist aus dem Unicast Routingprotokoll RIP (Routing Information Protocol) abgeleitet. RIP wurde im Internet über Jahre hinweg für das Routing in Routing-Domänen eingesetzt. DVMRP ist eines der ersten Protokolle, die für Multicast-Routing eingesetzt wurden. Es wird im MBone benutzt und ist ausschließlich für das Routing von Multicast-Datenpaketen zuständig. Unicast-Pakete werden weiterhin von RIP bearbeitet. Deshalb muß in einem Router neben DVMRP auch RIP vorhanden sein [ZS95].

DVMRP basiert auf dem im letzten Kapitel vorgestellten Reverse-Path Forwarding (RPF) Algorithmus. Es wird ein Baum mit kürzesten Wegen von der Quelle zu den Zielen aufgebaut, wobei die Nachrichten-Quelle nicht zur Gruppe gehören muß. Die ersten Pakete werden mittels "truncated broadcast" übertragen. Danach wird das Fehlen von Gruppen-Mitgliedern in Subnetzen im Baum in Richtung Quelle zurückpropagiert ("pruning"). Dieser Vorgang wurde im Kapitel über RPF genauer beschrieben. Es besteht auch die Möglichkeit, das Abschneiden eines Zweigs wieder rückgängig zu machen: Durch das Verschicken einer "graft"-Nachricht wird die "prune"-Nachricht gelöscht.

Weil im Internet zur Zeit nicht alle Router multicastfähig sind, muß eine Möglichkeit gefunden werden, die nicht multicastfähigen Router zu überbrücken. Im DVMRP Protokoll wird dazu die Technik des Tunnelns verwendet. Dabei wird das Multicast-Paket komplett in ein IP-Paket eingepackt. Diese Technik nennt man auch Einkapseln. Im Abschnitt über das MBone wurde diese Technik näher erläutert.

Der größte Vorteil von DVMRP ist seine Einfachheit. Als größter Nachteil ist die Skalierbarkeit zu nennen. DVMRP baut pro Quelle und Gruppe einen Baum auf. Dies führt

bei einer wachsenden Anzahl von Gruppen und Sendequellen zu einem unakzeptablen Verbrauch von Speicher. Weiter müssen auch Router die nicht auf einem Baum liegen, "prune"- und "graft"-Nachrichten speichern. Für den Einsatz in großen Netzen ist das DVMRP somit nicht geeignet.

3.2 MOSPF

MOSPF (Multicast Extension of Open Shortest Path First) ist aus dem Unicast Routing-Protokoll OSPF entstanden. Es basiert wie OSPF auf Link-State-Routing. Beim Link-State-Routing erkundet jeder Router seine Nachbarschaft, d.h. er stellt fest mit welchen anderen Routern er direkt verbunden ist. Diese Verbindungs-Zustands Information (daher der Name Link-State Routing) sendet er in gewissen Abständen an alle Router im Netz. Dadurch erhält jeder Router Kenntnis über die Netz-Verbindungsstruktur. Weil jeder einzelne Router eine komplette Information über die Topologie des Netzes hat, kann er den kürzesten Pfad von jeder Quelle zu jeder Gruppe berechnen. Dies geschieht mit dem Algorithmus von Dijkstra [HAU83] zur Berechnung der kürzesten Wege in einem Graphen.

Die Vorteile eines Link-State Algorithmus im Vergleich zum Distanz-Vektor Algorithmus liegen in der schnelleren Konvergenz des Verfahrens. Dafür muß jedoch die Gruppenmitgliedschaft durch das gesamte Netz gesendet werden. Aus diesem Grund ist MOSPF nur für das Routing in autonomen Systemen von begrenzter Größe geeignet. Es baut für jedes Paar (Quelladresse, Zieladresse), wobei die Zieladresse eine Gruppenadresse sein kann, einen MC-Baum mit der Quelle als Wurzel auf. Die Erstellung dieser Bäume erfolgt nur nach Bedarf, d.h. beim ersten Senden einer Dateneinheit mit dem bestimmten Adresspaar. Danach werden die Bäume eine bestimmte Zeit im Speicher gehalten. Es wäre viel zu aufwendig die Bäume für alle Gruppen und Quellen zu ermitteln.

Im Gegensatz zu OSPF werden von MOSPF keine alternativen Wege unterstützt. Auch die Einträge in den Routern zum Weiterleiten der Daten unterscheiden sich. Bei OSPF bezog sich das Weiterleiten der Pakete nur auf die Zieladresse, bei MOSPF muß zusätzlich noch die Quelladresse berücksichtigt werden. Man braucht pro (Quelladresse, MC-Zieladresse)-Paar einen Eintrag in der Routingtabelle.

Im Gegensatz zu DVMRP ist bei MOSPF die Technik des Tunnelns nicht vorgesehen, da in nächster Zeit damit zu rechnen ist, daß jeder IP-Router MC-fähig ist.

Ein Nachteil dieses Verfahrens ist der große Speicheraufwand, da in jedem Router pro Adresspaar ein MC-Baum erstellt wird.

3.3 CBT

Die zwei bisher vorgestellten Verfahren DVMRP und MOSPF haben beide den Nachteil, daß sie für jede Gruppe und jede Nachrichten-Quelle einen MC-Baum berechnen und abspeichern müssen. In diesem Abschnitt wird eine Protollarchitektur vorgestellt, die pro MC-Gruppe nur einen Baum aufbaut und damit in bezug auf Skalierbarkeit einen wesentlichen Vorteil gegenüber den bisher vorgestellten Verfahren bietet. Ein weiterer Vorteil dieser neuen Architektur ist die Unabhängigkeit vom darunterliegenden Unicast-Routing-Algorithmus [BFC93].

CBT-Architektur Dieses Verfahren basiert auf den sogenannten Core-Based Trees (CBT). In einem CBT gibt es einen ausgezeichneten Router, der als Core des Baumes bezeichnet wird. Der Core-Router muß nicht topologisch in der Mitte der Knotenmenge liegen. Alle Zweige des Baumes gehen von diesem ausgezeichneten Router aus. Sie bestehen aus den sog. non-core Routern, die einen kürzesten Pfad zwischen den “End“-Routern des Baumes und dem Core bilden. Im Gegensatz zu den bisherigen Protokollen wird pro Gruppe nur ein Baum gebildet, den sich alle Quellen einer Gruppe teilen. Charakteristisch für dieses Verfahren ist der empfängerinitiierte Aufbau des Baumes.

Beim CBT-Routing lassen sich zwei Phasen unterscheiden:

1. Senden der MC-Pakete zum Baum durch den Unicast-Dienst.
2. Senden der MC-Pakete über den Baum

Vorteile der CBT-Architektur im Gegensatz zu quellenbasierten Routingbäumen:

Skalierbarkeit: Jeder Router muß nur Information über jede Gruppe speichern und nicht über jedes (Quelle,Gruppe)-Paar. Aus diesem Grund wurde CBT entwickelt.

Baum-Aufbau: Nur Router, die zum Baum einer Gruppe gehören, müssen Baum-Information speichern.

Unabhängigkeit vom UC: Der Baumaufbau sowie das Versenden der MC-Pakete erfolgt unabhängig vom UC-Routing-Protokoll. Die gesamte MC-Baum-Information kann allein aus der UC-Tabelle eines Routers gewonnen werden.

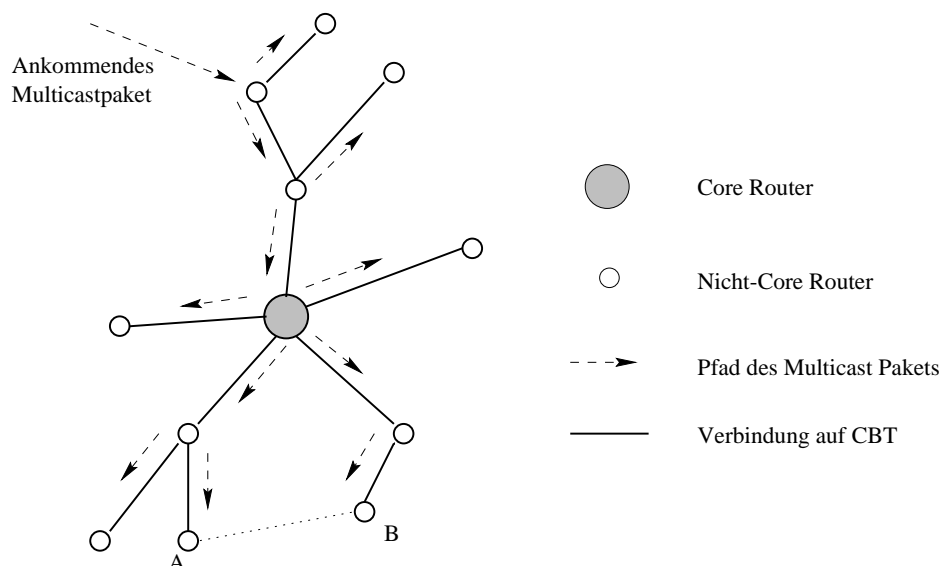


Abbildung 47. CBT Baum

Nachteile:

Pfadauswahl: Beim CBT-Routing wird nicht zwingend der optimale Pfad zwischen zwei Knoten gefunden. Falls in Bild 47 eine Nachricht von Router A an die Gruppe gesendet werden soll, wird das Paket über den Baum geschickt. B erhält das Paket jedoch nicht über die kürzeste Verbindung zu A. Diese Verbindung gehört in diesem Fall nicht zum CBT.

Core-Router: Der Core-Router ist der wichtigste Teil der gesamten Architektur. Wenn er ausfällt, zerfällt der Baum in mehrere Teilbäume. Aus diesem Grund führt man weitere Core-Router ein, die alle in einer Core-Liste nach ihrer Priorität geordnet werden.

CBT-Protokoll Bei CBT-Routing unterscheidet man zwischen Kontroll-Paketen und Daten-Paketen. Beide Paketarten werden in IP-Datagrammen versendet. Ein Router muß Pakete, die ein MC-Daten-Paket beinhalten, erkennen. Die IP-Pakete tragen die Core-Adresse im Zieladressfeld und die Gruppen-Id im Optionsfeld ihres Headers. Dazu muß eine neue IP-Optionsnummer definiert werden. Sie werden durch den Unicast-Dienst in Richtung Core-Router gesendet. Wenn das Paket auf den MC-Baum trifft, wird die Gruppen-Id vom Optionsfeld ins Zieladressenfeld kopiert. Dann wird das Paket auf dem Baum versendet (Bild 47).

Ein Host wird auf folgende Art und Weise an den Baum einer Gruppe "angehängt": Der Host sendet eine Nachricht über das IGMP-Protokoll an einen Router in seinem Subnetz. Die Nachricht enthält die Gruppen-Id und Core-Liste. Der Router verschickt dann eine JOIN-REQUEST Nachricht in Richtung Core-Router. Die Join-Request Nachricht wird von Router zu Router versendet, bis sie den Core erreicht oder bis sie einen Router erreicht, der Teil des entsprechenden Baumes ist. Der empfangende Router schickt ein Join-ACK zurück. Durch dieses Join-ACK wird nun der Baum aufgebaut, indem alle Router, die das Join-ACK weitersenden, zu non-core Routern werden (sie ändern ihren Zustand). In der Zeit zwischen dem Request und ACK sind die Router in einem schwebenden Zustand.

Jeder Router im Baum speichert seine Eltern- und Kinder-Anschlüsse bezüglich dieses Baumes ab. Über den Elternanschluß wurde das ACK empfangen; über den Kinderanschluß wurde das ACK versendet. Zu beachten ist, daß jeder non-core Router mehrere Kinderanschlüsse besitzen kann.

Wenn ein non-core Router eine Gruppe verlassen will, sendet er ein Quit-Request an seinen Eltern-Router. Dazu dürfen jedoch keine Mitglieder dieser Gruppe direkt anliegenden Subnetzen des Routers angehören und auf allen seinen Kinderanschlüssen muß er ein Quit-Request empfangen haben. Darauf antwortet der Eltern-Router mit einem Quit-ACK und löscht den Kinderanschluß.

Falls die Verbindung eines Routers zu seinem Eltern-Router ausgefallen ist, gibt es zwei Möglichkeiten dieses Problem zu lösen:

1. Zum einen kann der Router versuchen, sich dem Baum wieder anzuschließen, indem er ein Join-Request an den Core sendet. Die baumabwärtsliegenden Router merken in diesem Fall nichts vom Ausfall des Routers.

2. Als zweite Möglichkeit kann er eine Flush-Tree Nachricht an alle Router schicken, die unter ihm im Baum liegen. Danach muß jeder einzelne Router selbst versuchen, sich wieder an den Baum dranzuhängen.

Der Hauptvorteil der CBT-Architektur liegt in ihrer Skalierbarkeit und ihrer Unabhängigkeit vom darunterliegenden UC-Protokoll. CBT ist ein recht einfaches und effizientes Protokoll.

3.4 PIM

Das Protocol Independent Multicasting (PIM) gibt es in zwei Versionen: PIM dense und PIM sparse. Dabei beziehen sich dense und sparse auf die Dichte der Gruppenmitglieder im Internet. Eine Gruppe nennt man "dense", wenn die Wahrscheinlichkeit groß ist, daß ein beliebiges Gebiet des Internet ein Gruppenmitglied enthält, auch wenn das Gebiet relativ klein gewählt wird. Eine Gruppe heißt "sparse", wenn das Gegenteil gilt.

PIM dense Der Vorteil dichter Gruppen besteht darin, daß man einen recht einfachen Routing-Algorithmus einsetzen kann: RPF und Kürzen (flood and prune). Dieser Algorithmus wurde im vorigen Kapitel schon vorgestellt. PIM ist - wie der Name schon sagt - ein Protokoll, welches unabhängig vom Unicast-Routing ist. Es geht nur davon aus, daß ein solches Protokoll existiert.

"RPF und Kürzen" ist wie folgt implementiert:

1. Wenn ein Router ein MC-Paket von der Quelle S an Gruppe G empfängt, untersucht er, ob er selbst UC Pakete an Q auf dem gleichen Weg versenden würde. Falls nicht, schickt er eine prune(Q,G)-Nachricht an den benachbarten Router, von dem er das Paket erhalten hat.
2. Sonst schickt er eine Kopie der Nachricht in alle Richtungen, aus denen er keine prune(Q,G)-Nachricht erhalten hat, mit Ausnahme der Richtung, aus der das Paket kam. Falls es keine solche Richtung gibt, sendet er ein prune(Q,G) zurück.

PIM sparse Dieses Verfahren hat große Ähnlichkeit mit dem CBT-Algorithmus. Dem Core Router entspricht der sogenannte "Rendezvous Point" (RP), wobei jeder Gruppe einer oder mehrere dieser Rendezvous Points zugewiesen wird. In dieser Arbeit wird der Einfachheit wegen angenommen, daß jede Gruppe nur einen RP besitzt. Der größte Unterschied zwischen PIM und CBT ist die Möglichkeit von PIM zwei Arten von Bäumen zu benutzen: Die "shared trees", für ein Multicast mit niedrigen Kosten, und die "source-based trees", für die Übertragung mit niedrigster Verzögerung [Ram95].

Beim Aufbau eines "shared trees" unterscheidet man zwischen:

Kommunikation zwischen einem Empfänger und dem RP: Wenn ein Host einen "shared tree" aufbauen möchte, stößt sein Designated Router (DR) einen sogenannten "flow setup" an, der vom RP der Gruppe zum DR verläuft. Der DR kennt die

RP-Id und Gruppen-Id und sendet Join-Req Paket in Richtung RP. Der erste Knoten, der zum Baum (mit dieser RP-Id und Gruppen-Id) gehört bzw. der RP selbst, schicken ein Join-Rep zurück (Bild 48).

Kommunikation zwischen einem Sender und dem RP: Eine Quelle Q, die an eine Gruppe G senden will, initiiert ihrerseits einen “flow setup“ zwischen Q und dem RP der Gruppe. Der DR des sendewilligen Hosts kennt die RP-Id und Gruppen-Id und sendet eine RP-Register Nachricht an den RP der Gruppe. Darauf sendet der RP dem Sender eine Nachricht, daß der Sender nun zum Baum gehört (TREE-Join).

Der Empfänger hat die Möglichkeit, nachdem der shared-tree aufgebaut ist, zu einem source-rooted shortest-path tree zu wechseln.

Beim Wechseln vom shared tree zum source-rooted shortest path tree kann man zwischen zwei Phasen unterscheiden:

Empfänger-Quellen Kommunikation: Ein Host, der Pakete über den shared tree von der Quelle Q empfängt, kann die Pakete auch über den source-rooted Baum versenden lassen. Sein DR Initiiert dazu ein Tree-Join zur Quelle Q. Daraufhin werden die Pakete alle über den kürzesten Pfad zum Empfänger gesendet.

Empfänger-RP Kommunikation: Nachdem der Empfänger dem Source-rooted shortest-path tree angeschlossen ist, kann er den shared tree verlassen, indem er Tree-Leave versendet. Wenn der Empfänger daraufhin ein Antwort-Paket erhält, ist er vom shared tree abgehängt und empfängt alle Pakete dieser Gruppe über den RPF-Baum von der Quelle.

Alle weiteren Pakete von dieser Quelle werden jetzt über den kürzesten Pfad versendet. Auf diesem kürzesten Pfad muß jedoch nicht der RP liegen. Wenn ein Host eine Gruppe verlassen will, sendet er eine “prune“-Nachricht an die Quelle des source-based trees bzw. an den RP des shared trees.

Eine Besonderheit des PIM sparse Protokolls ist die Bereitstellung zweier verschiedener MC-Bäume. Beim shared tree bestehen viele Ähnlichkeiten zum CBT-Verfahren. Das Besondere bei PIM sparse Protokoll ist die Möglichkeit, zwischen diesen zwei MC-Bäumen zu wählen.

4 Vergleich der Multicast-Protokolle

Die beiden Protokolle DVMRP und MOSPF werden zur Zeit zum MC-Routing im Internet eingesetzt. CBT und PIM sind bisher nur Protokollspezifikationen, die noch nicht im breiten Einsatz sind. Ein großer Unterschied zwischen den im Augenblick genutzten Protokollen (DVMRP und MOSPF) und den Protokollen CBT und PIM besteht in ihrer Abhängigkeit von einem Unicast-Dienst. DVMRP baut auf RIP auf, MOSPF ist eine Erweiterung von OSPF. Beide Protokolle funktionieren nur, wenn sie gemeinsam mit den speziellen Unicast-Protokollen genutzt werden. CBT und PIM sind dagegen unabhängig

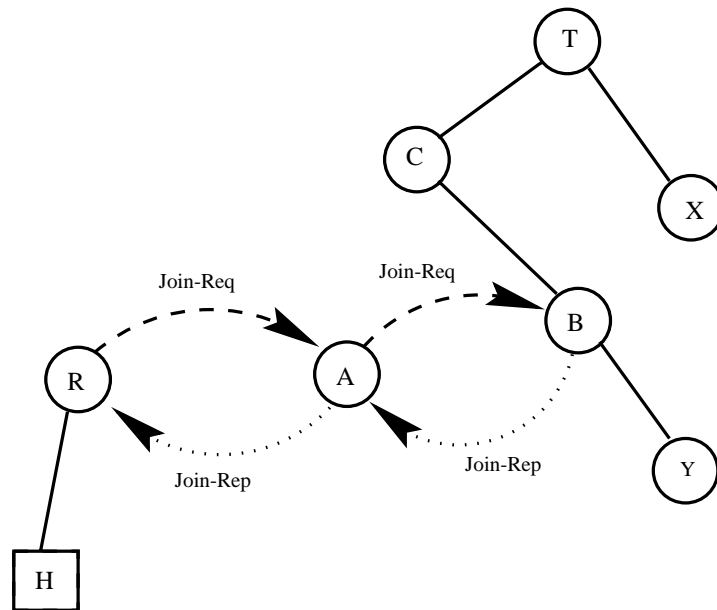


Abbildung 48. Beitreten des Host H zum MC-Baum

vom Unicast-Dienst. Es spielt keine Rolle, welche Unicast-Protokolle zur Verfügung stehen. CBT und PIM setzen nur die Existenz eines beliebigen Unicast-Protokolls voraus und sind damit auch für das Routing über Domänengrenzen hinweg geeignet.

Die vorgestellten Protokolle unterscheiden sich wie schon erwähnt in ihrer Skalierbarkeit. Das beste Ergebnis erzielt hier das CBT, welches nur einen MC-Baum pro Gruppe aufbaut. Die anderen Verfahren brauchen einen MC-Baum pro Gruppe und Quelle. Dies ist der größte Nachteil dieser Verfahren. Wenn die Anzahl der Gruppen im Internet ansteigt, steigt auch die Anzahl der möglichen Quellen an, und somit die Anzahl der zu berechnenden MC-Bäume. Der Speicherbedarf von DVMRP, MOSPF und PIM wird dann relativ groß.

Die Protokolle unterscheiden sich auch in der Art des Baumaufbaus. Bei CBT initiiert der Empfänger die Speicherung der Zustandsinformation in den Routern, bei PIM der Empfänger oder der Sender und bei den beiden anderen Verfahren sind es die einzelnen Router selbst, ohne daß Sender bzw. Empfänger einwirken.

Bei den Protokollen DVMRP und MOSPF sind alle Router im Netz in den Aufbau des Baumes einer Gruppe involviert. Auch wenn ein Router nicht auf dem Baum einer Gruppe liegt, muß er "prune"- und "graft"-Nachrichten dieser Gruppe entgegennehmen und auch speichern. Diesen Nachteil hat das CBT-Protokoll nicht. CBT braucht nur die Router, die auf dem Pfad zwischen dem Core und den Gruppenmitgliedern liegen, um Zustände zu speichern.

Der größte Vorteil von PIM gegenüber den anderen Protokollen besteht in seiner Möglichkeit, zwei verschiedene Modi zu benützen. Dadurch steigt die Flexibilität von PIM. Es ist möglich, das Protokoll an die Gegebenheiten des Netzwerks anzupassen.

Mechanismen zur Gruppenkommunikation in globalen Netzen

Jens Clever

Kurzfassung

Durch die Verbreitung und immer stärkere Vernetzung von Rechnern ist die Gruppenkommunikation per Computer fast schon zu einer normalen Kommunikationsform geworden. Ältere Protokollarchitekturen, die für die Regelung von Unicastverbindungen konzipiert wurden, können die neuen Anforderungen kaum zufriedenstellend bewältigen. Im Rahmen dieser Seminararbeit werden sieben neue Protokollmechanismen vorgestellt, die unterschiedliche Ansätze zur Lösung der Problematik der Gruppenkommunikation in Weitverkehrsnetzen aufzeigen. Das Ziel ist eine Analyse dieser Protokollmechanismen. Die Vor- und Nachteile der einzelnen Mechanismen sollen aufgezeigt werden, um eine Einschätzung zu geben, für welche Aufgaben und unter welchen Voraussetzungen der Einsatz eines bestimmten Protokolls sinnvoll ist.

1 Einleitung

Der Datenaustausch über Rechnernetze ist in der heutigen Zeit schon fast zu einer normalen Kommunikationsform geworden. Es findet Gruppenkommunikation zwischen Teilnehmern mit verschiedenen Rechnern, die örtlich weit voneinander entfernt und über heterogene Netze verbunden sind, statt. Man muß davon ausgehen, daß diese Form des Informationsaustausches mit der zunehmenden Vernetzung von Computern weiter anwachsen wird. Diesem Umstand Rechnung tragend, werden derzeit schon neue Protokollarchitekturen diskutiert, welche durch neuartige Transportmechanismen und Kommunikationsstrukturen besser auf die zu erwartenden und teilweise heute schon vorhandenen Probleme der Gruppenkommunikation ausgerichtet sind.

2 Begriffe

Im folgenden werden Begriffe erläutert, die zum Verständnis der Ausarbeitung wichtig sind.

1-1-V Klassische Zweiparteienverbindung (Unicast)

1-n-V Mehrparteienverbindung mit einem Sender und n Empfängern (Multicast)

n-m-V Mehrparteienverbindung mit n Sendern und m Empfängern (Multipeer)

Abstandsmetrik Die „Entfernung“ zweier Kommunikationseinheiten kann nach verschiedenen Metriken angegeben werden. Gebräuchliche Kriterien zur Abstandsbestimmung sind z.B.:

- Die **Verzögerung** der Übertragung zwischen den Endgeräten
- Die **Bandbreite** der Verbindung zwischen den Systemen
- Der maximal mögliche oder der durchschnittliche **Datendurchsatz** auf der Verbindung zwischen den Kommunikationssystemen
- Die **Fehlerwahrscheinlichkeit** der Übertragung zwischen den Kommunikationssystemen
- Die **Zuverlässigkeit** der Verbindung zwischen den Systemen
- Die **Betriebskosten** der Verbindung zwischen den Systemen
- Die **geographische Distanz** zwischen den Systemen
- Die **Anzahl der Hops** zwischen den Systemen

Abstandsmaß Der Abstand zweier Kommunikationssysteme ist abhängig von der gewählten Metrik und der Lokation der Systeme.

3 Problematik

Bei Kommunikationsvorgängen zwischen n Parteien in Weitverkehrsnetzen ist es ungünstig und aufwendig n unterschiedliche 1-1-V aufzubauen und zu verwalten. Die Kosten für Mietleitungen, der Verwaltungsaufwand und die damit verbundene Speicherbelastung sprechen gegen diese Methode. Der Aufwand zur Verwaltung des Betriebes eines solchen Systems wächst linear mit der Anzahl der Teilnehmer. Es ist abzusehen, daß aus der beschränkten Rechenleistung des Senders eine Beschränkung der Teilnehmer an der Gruppenkommunikation folgt. Auch wenn keine Echtzeitkommunikation gefordert und ein zeitlich versetztes Senden möglich ist, bleiben die hohen Kosten durch den Betrieb von Mietverbindungen und eine unerwünschte Mehrbelastung des Senders durch den längeren Einsatz. Aus diesen Gründen werden neue Verbindungsmechanismen entwickelt um eine zuverlässige Gruppenkommunikation in globalen, heterogenen Netzen auch bei einer großen Anzahl von Teilnehmern zu ermöglichen. Das Ziel einer Optimierung ist eine Minimierung der zum Datenaustausch benötigten Bandbreite und Rechenzeit, bei einer möglichst kurzen Auslieferungsverzögerung. Um einen breiten Einsatz zu ermöglichen, sollten die Mechanismen sehr allgemein gehalten sein. Hohe Hardwareanforderungen oder Umrüstungen von Zwischensystemen würden den Einsatz der Mechanismen zusätzlich erschweren.

4 Analyse des Protokollablaufs

Um Ansatzpunkte für Verbesserungen in Protokollarchitekturen aufzeigen zu können muß zuerst geklärt sein, in welche Schritte ein Protokollablauf gegliedert ist.

4.1 Die Datentransferphase

Die Datentransferphase einer Multicast-Operation läßt sich in folgende Schritte unterteilen:

1. Zusammensetzen der Datenpakete beim Sender
2. Übertragen der Multicast-Nachrichten
3. Empfangen und Auswerten der Pakete bei den Empfängern
4. Erzeugen von Quittungen bei den Empfängern
5. Übertragen von Quittungen über das Kommunikationsnetz
6. Empfangen und Auswerten von Quittungen beim Multicast-Sender
7. Durchführen von Übertragungswiederholungen (optional)

Rechenzeit wird von den beteiligten Kommunikationsknoten beim Zusammensetzen und Auswerten von Paketen während der Schritte **1.**, **3.**, **4.** und **6.** verbraucht.

Bandbreite wird bei der Übertragung von Quittungen und Datenpaketen während der Schritte **2.**, **5.** und **7.** benötigt.

4.2 Ansatz zu Verbesserungen

Der Aufwand für die Analyse von Nachrichten wird stark vom verwendeten Paketformat beeinflußt. Durch feste Längen von Paketkopf und -anhang, Platzierung von Kontroll- und Statusinformationen im Paketkopf sowie der Ausrichtung des Formats auf eine den Prozessoren angepaßte Bytelänge, können Pakete „on the fly“ analysiert werden. Dem wird in einigen modernen Protokollen Rechnung getragen. Durch geschickten Quittierungsbetrieb läßt sich das Gesamtaufkommen der zu transportierenden Statusmeldungen reduzieren und damit Bandbreite einsparen. So läßt sich durch Zusammenfassen der Quittungen mehrerer Empfänger die Übertragung redundanter Statusmeldungen und somit eine Überhäufung des Senders mit Quittungen vermeiden. Beim Senden von Datenpaketen kann nur während Übertragungswiederholungen Bandbreite eingespart werden. Wiederholt zu übertragende Pakete sollten nicht das gesamte Kommunikationsnetz durchqueren, sondern möglichst wenige Verbindungen belasten. Dies führt neben einer Entlastung des globalen Netzes auch zu einer Verkürzung der Auslieferungsverzögerung. Eine Optimierung von Gruppenkommunikationsprotokollen kann an drei Punkten ansetzen:

1. Optimierung des Paketformates
2. Optimierung des Quittierungsbetriebes
3. Optimierung der Übertragungswiederholung

Im folgenden werden vor allem die Punkte 2 und 3 näher untersucht.

5 Wertungskriterien

Der Einsatz von neuen Protokollmechanismen ist davon abhängig, wie komfortabel sie, die an sie gestellten, Anforderungen erfüllen. Dabei ist auch entscheidend, welche Systemvoraussetzungen erfüllt sein müssen, um ein einwandfreies Funktionieren der Mechanismen zu garantieren. Daher müssen die Mechanismen nach folgenden Gesichtspunkten untersucht werden.

- Speicheranforderungen
- Verwaltungsaufwand
- Flexibilität (Erweiterungen)
- Quittierungsstrategie (Entlastung von Netz und Sender)
- Auslieferungsverzögerung (Sendewiederholungen)
- Netzbelastung
- Fehlertoleranz (Ausfall von Systemkomponenten)

6 Beschreibung der Protokollmechanismen

Bemerkung

Der Begriff „lokal“ ist im folgenden nicht unbedingt auf seine geographische Bedeutung beschränkt. Vielmehr ist darunter ein geringer Abstand der Kommunikationssysteme im Sinne der gewählten Metrik (siehe Abschnitt 2) zu verstehen. Auch werden nur 1-n-Verbindungen betrachtet. Die Überlegungen und Lösungsansätze lassen sich aber mit relativ geringen Änderungen auch auf n-m-Verbindungen übertragen.

6.1 Lokale Gruppen (LG)

Nahe beieinanderliegende Empfänger schließen sich zu Gruppen, den lokalen Gruppen, zusammen. Innerhalb dieser können Empfängerquittungen gesammelt und fehlende Datenpakete ausgetauscht werden. Dem Sender werden nur Gruppenquittungen übermittelt. Die Verantwortung für die Zuverlässigkeit der Übertragung liegt verteilt sowohl beim Sender als auch bei den Verwaltern der lokalen Gruppen. Jeder Empfänger einer Multicastverbindung, der keiner lokalen Gruppe angehört, startet eine Suchroutine, welche bei bereits existierenden lokalen Gruppen um Aufnahme bittet. Wird keine lokale Gruppe im maximal zulässigen Abstand gefunden oder die Aufnahme bei allen in Frage kommenden lokalen Gruppen abgelehnt, bildet der Empfänger eine eigenständige lokale Gruppe und ernennt sich selbst zu deren Verwalter. Dieser kann während der Übertragung wechseln. Der Verwalter fungiert als Bindeglied zwischen dem Sender und den Empfängern. Er sammelt die Einzelquittungen, bildet Gruppenquittungen und sendet diese an den Sender. Der Sender erhält nur Gruppenquittungen, womit eine

Quittungsschwemme vermindert wird. Der Gruppenverwalter bearbeitet Aufnahme- und Abmeldevorgänge innerhalb der lokalen Gruppe. Optional kann der Verwalter auch die Struktur der Gruppe analysieren und bei dynamischen Systemen gegebenenfalls einen Verwalterwechsel initiieren. Um einem Ausfall des lokalen Gruppenverwalters und den daraus resultierenden Störungen des Sendebetriebs vorzubeugen, können ein oder mehrere Empfänger die Rolle von redundanten Gruppenverwaltern übernehmen. Während des normalen Sendebetriebs übernehmen sie keine besondere Funktionen, verfügen aber über die gleichen Gruppenstatusinformationen wie der Gruppenverwalter. Für den Fall, daß der Gruppenverwalter sich abmeldet oder ausfällt, übernimmt dann ein redundanter Gruppenverwalter die Verwalterfunktionen. Welches Mitglied der lokalen Gruppe Verwalter ist, hängt von verschiedenen Kriterien, die sich aus der Anwendung ableiten ab. Wenn ein Empfänger die Gruppe verlassen will, meldet er dies dem Gruppenverwalter, der die Gruppenstatusinformation entsprechend ändert und den Abmeldevorgang an den Sender weitergibt. Die Verbindung wird erst unterbrochen, wenn der Gruppenverwalter eine Bestätigung an den Austrittswilligen gesendet hat. Tritt ein Verwalter aus, muß vorher noch ein neuer Gruppenverwalter bestimmt werden. Der Gruppenverwalter steht mit jedem anderen Gruppenmitglied in Verbindung. Das Sammeln von Einzelquittungen geht wie folgt vor sich: Der Sender verlangt über Multicast von jedem Empfänger eine Statusmeldung. Innerhalb einer lokalen Gruppe werden diese zufällig zeitversetzt (um Kollisionen zu vermeiden) an den Gruppenverwalter gesendet. Dieser erstellt eine Gruppenstatusmeldung. Nur Datenpakete, die von keinem Gruppenmitglied korrekt empfangen wurden, werden beim Sender erneut angefragt. Der Gruppenverwalter sammelt innerhalb der Gruppe die Datenpakete, die einigen Gruppenmitgliedern fehlen, von anderen aber korrekt empfangen wurden, und überträgt diese an die entsprechenden Empfänger.

6.2 Designated Status Protocol (DSP)

Die Empfängermenge ist in disjunkte Gruppen eingeteilt. Das Mitglied jeder Gruppe, das den geringsten Abstand zum Sender hat (hier die geringste Anzahl von Vermittlungsstellen) übernimmt die Rolle eines lokalen Gruppensenders (LGS). Bei diesem Mechanismus wird zwischen drei Komponenten der Multicastverbindung unterschieden: Dem Sender der Multicastverbindung, ausgezeichneten Empfängern (LGS) und normalen Empfängern. Die Verantwortung für die Zuverlässigkeit der Verbindung liegt verteilt beim Sender und den LGS.

Der Sender verbreitet die Daten an jeden Empfänger über die Multicastverbindung. Die LGS verfügen jeweils über zwei verschiedene Puffer: Den Eingangs- und den Wiederholungspuffer. Die lokalen Gruppensender speichern ankommende Daten im Eingangsspeicher und verschieben sie in den Wiederholungspuffer, wenn dort genügend Speicherplatz vorhanden ist. Hat ein normaler Empfänger Daten nicht korrekt erhalten, bittet er bei dem LGS seiner Gruppe um Sendewiederholung. Der LGS verbreitet dann über Multicast die angeforderten Daten innerhalb der Gruppe. Ein Datenaustausch zwischen den Empfängern innerhalb einer Gruppe findet nicht statt. Wenn alle Gruppenmitglieder den korrekten Empfang der betreffenden Daten bestätigt haben, werden diese aus dem Wiederholungspuffer des lokalen Gruppensenders gelöscht. Die LGS melden in regelmäßigen

Abständen nur ihren eigenen Zustand an den Sender. Diese Statusmeldungen enthalten Angaben über den verfügbaren Speicherplatz im Eingangspuffer und Quittungen. Da der Sender nur von den LGS Quittungen erhält, wird so eine Quittungsschwemme vermieden. Ist ein Datenpaket von einem LGS nicht korrekt empfangen worden, verbreitet es der Sender erneut über die Multicastverbindung an alle Empfänger. Neue Pakete werden erst dann übertragen, wenn alle LGS über ausreichend Empfangsspeicher verfügen.

6.3 Consolidated Status Protocol (CSP)

Bei diesem Mechanismus wird zwischen drei Komponenten der Multicastverbindung unterschieden: Dem Sender der Multicastverbindung, ausgezeichneten Routern und normalen Empfängern. Die Verantwortung für die Zuverlässigkeit der Verbindung liegt bei dem Sender und den Routern. Die Empfängergruppe ist in disjunkte Gruppen eingeteilt. Die Router, über welche jeweils ein Subnetz erreicht wird, übernehmen drei zusätzliche Funktionen: Sammeln von Empfängerstatusmeldungen, Erstellen von Gruppenstatusmeldungen und das Senden derselben an den Sender.

Innerhalb einer Gruppe senden die Empfänger in regelmäßigen Intervallen ihre Statusmeldungen an den für diese Gruppe zuständigen Router. Der Router erstellt nun eine Gruppenstatusmeldung. Diese identifiziert die Gruppe, gibt an, welche Blöcke von allen Gruppenmitgliedern empfangen wurden, welche Daten von mindestens einem Gruppenmitglied vermisst werden und wieviel Speicherplatz den Empfängern mindestens zur Verfügung steht. Diese Gruppenstatusmeldungen werden von den Routern in regelmäßigen Intervallen an den Sender geschickt. Die Router speichern keine Daten zwischen. Es findet kein Datenaustausch innerhalb der Gruppe, oder zwischen der Gruppe und dem zuständigen Router statt. Der Sender wiederholt Sendungen, wenn auch nur ein Empfänger nicht alle Daten korrekt erhalten hat. Diese Übertragung kann per Multicast oder über eine 1-1-V geschehen, sofern zusätzlich 1-1-V aufgebaut werden können. Neue Daten werden erst dann übertragen, wenn alle Empfänger über ausreichend freien Speicher verfügen.

6.4 Combined Protocol (CP)

Bei diesem Mechanismus wird zwischen drei Komponenten der Multicastverbindung unterschieden: Dem Sender der Multicastverbindung, ausgezeichneten Routern und normalen Empfängern. Die Verantwortung für die Zuverlässigkeit der Verbindung liegt verteilt bei dem Sender und den Routern. Die Empfängergruppe ist in disjunkte Gruppen eingeteilt. Die Router, über welche jeweils ein Subnetz erreicht wird, übernehmen vier zusätzliche Funktionen: Sammeln von Empfängerstatusmeldungen, Erstellen von Gruppenstatusmeldungen, das Senden derselben an den Sender und den Aufbau von lokalen Multicastverbindungen.

Der Quittungsbetrieb erfolgt analog zu CSP. Aber diese Gruppenstatusmeldung unterscheidet sich in folgenden Punkten von der Gruppenstatusmeldung des CSP: Es wird ermittelt, welche Datenpakete von einem beliebigen Empfänger innerhalb der Gruppe korrekt erhalten wurden. Nur die Daten, die bei keinem Empfänger korrekt ankamen

werden beim Sender als vermißt gemeldet. Durch die Empfängerstatusmeldungen ist dem Router bekannt, welche Blöcke nicht bei allen Empfängern korrekt ankamen und bei welchen Empfängern gegebenenfalls eben diese Daten vorliegen. Basierend auf diesen Informationen bauen die Router nun gruppeninterne Multicastverbindungen auf und weisen die Empfänger an, die betreffenden Daten zu senden. Die Quittungen für diese Sendungen werden weiterhin an den Router geschickt. Dieser Informationsfluß innerhalb der Gruppe besteht solange, bis alle Gruppenmitglieder auf dem selben Stand sind. Die Daten, die von keinem Gruppenmitglied empfangen wurden, werden vom Sender erneut an alle Empfänger per Multicast gesendet. Neue Daten werden erst dann übertragen, wenn alle Empfänger über ausreichend freien Speicher verfügen.

6.5 Tree-based Multicast Transport Protocol (TMTP)

Die Multicastverbindung ist hierarchisch in eine Baumkontrollstruktur gegliedert. Die Empfängermenge ist in Gruppen eingeteilt. Jeder Gruppe steht ein Verwalter vor, in Zukunft Vater genannt. Auch der Sender gilt als Vater. Die Verantwortung für die Zuverlässigkeit der Übertragung ist auf den Sender und die Väter verteilt. Vor der Bildung wird festgelegt, welche maximale Anzahl von Söhnen ein Vaterknoten haben kann. Söhne sind Verwalter anderer Gruppen oder normale Empfänger die vor Beginn der Sendung nicht in Gruppen eingeteilt wurden. Eine nachträgliche Teilnahme an dieser Verbindung erfolgt durch eine Ringsuche, die gemäß Abstandsvorgaben eine Anfrage bei existierenden Gruppen startet oder eine neue Gruppe bildet. Das Ausscheiden aus dem Verbund ist ohne Komplikationen jedoch nur für Empfänger möglich, die kein Vater sind. Möchte nun ein solcher Gruppenverwalter die Verbindung unterbrechen, muß er zuerst sicherstellen, daß alle seine Söhne in anderen Gruppen Aufnahme gefunden haben, bevor er sich abmeldet. Für die Neueinteilung von Empfängern gibt es zwei Möglichkeiten: 1. Die Empfänger starten erneut eine Ringsuche und scheiden aus der Verbindung aus, wenn sie einen neuen Vater gefunden haben. Dies kann zu Datenverlusten während der Reintegration führen. 2. Der Vater startet diese Suche für jeden seiner Söhne und sorgt dafür, daß die Daten, die während der Reintegrationsphase gesendet werden, bis zur endgültigen Eingliederung zwischengespeichert bleiben. Der Verwalter regelt die Aufnahme eines neuen Mitgliedes der Gruppe oder verweigert die Aufnahme, wenn er bereits die vorgegebene maximale Anzahl von Söhnen hat. Datenpakete werden vom Vater zwischengespeichert. Der Verwalter sendet, sobald er ein Paket korrekt erhalten hat, eine positive Quittung an seinen Vater (Verwalter). Sollten Daten verloren oder verstümmelt worden sein, wird per Zufall eine Verzögerungszeit ermittelt, die bis zum Absenden der negativen Quittung zu verstreichen hat. Sollte während dieser Zeit ein gleichgestelltes Gruppenmitglied (kein Kind) dasselbe Paket anfordern, wird die eigene Anfrage gelöscht ansonsten weitergegeben. Fehlende Daten werden nur in dem (Sub-)zweig des Baumes wiederholt per Multicast übertragen, in dem sie vermißt werden. Erst wenn alle Kinder eines Knotens eine positive Quittung signalisiert haben, werden die zwischengespeicherten Datenpakete gelöscht.

6.6 Log-Based Receiver-reliable Multicast (LBRM)

Die Verantwortung für die Zuverlässigkeit der Verbindung liegt bei den Empfängern. Jedes Datenpaket hat eine eindeutige Sequenznummer. Innerhalb einer festgelegten Zeitspanne muß der Sender mindestens ein Paket verschicken. Wenn keine neuen Daten vorliegen, wird ein „Leerpaket“ mit Sequenznummer, aber ohne Inhalt gesendet. Zusätzlich zu den Empfängern werden die Daten in einem Logging-Server gespeichert. Die Empfänger können anhand der Sequenznummern feststellen, ob sie alle Pakete erhalten haben. Wenn die Empfänger auf die Daten Wert legen, können verlorengegangene Datenpakete bei dem Logging-Server über eine Unicastverbindung angefragt und übertragen werden. Wie lange Daten im Logging-Server gespeichert bleiben und innerhalb welcher Zeitspanne eine Übertragung erfolgen muß, hängt stark vom Datenaufkommen und den Anwendungsbedingungen ab. Handelt es sich beim Logging-Server um ein externes Gerät, werden die Daten über eine 1-1-V an den Server übertragen. Für die korrekte Datenübertragung zum Logging-Server ist in diesem Fall der Sender verantwortlich.

Bei Architekturen mit mehreren Logging-Servern wird ein, durch seinen Abstand zum Sender ausgezeichneter Logging-Server (LS1) wie beschrieben behandelt. Alle weiteren Logging-Server werden vom Sender und von LS1 wie normale Empfänger behandelt. Auch ein Konzept, in dem die weiteren Logging-Server von LS1 bevorzugt gegenüber normalen Empfängern behandelt werden ist denkbar. Die weiteren Logging-Server müssen jedoch sämtliche Daten korrekt erhalten und speichern. Fehlende Daten beziehen sie, wie andere Empfänger auch, von LS1. Die Empfänger richten ihre Wiederholungsanfragen immer an den Logging-Server, zu dem sie den geringsten Abstand haben.

6.7 Scalable Reliable Multicast (SRM)

Es erfolgt keine Einteilung der Empfängermenge in Gruppen. Jedes Datenpaket wird mit einem global eindeutigen Namen versehen, anhand dessen die Empfänger Übertragungslücken feststellen können. Dies ist gefordert, da dieser Mechanismus speziell für Whiteboards entwickelt wurde. Die Zuständigkeit für die Zuverlässigkeit der Verbindung liegt bei den Empfängern. Wenn die Empfänger auf die jeweiligen Daten Wert legen, können verlorengegangene Datenpakete, da jeder Teilnehmer mit jedem anderen per Multicast in Verbindung steht, bei allen Beteiligten nachgefragt werden.

Wenn ein Teilnehmer „A“ um eine Sendewiederholung bittet, oder auf eine solche antwortet, wartet er jeweils eine kurze, zufällig gewählte Zeitspanne. Wenn ein anderer Teilnehmer innerhalb dieser „Wartezeit“ den Vorgang durchführt, verzichtet „A“ auf seine Aktion. Dies vermindert das Senden von redundanten Anfragen und Antworten. Bei der Beantwortung einer Sendewiederholungsanfrage wird nicht zwischen Sender und Empfängern unterschieden, d.h. es existieren im System nur Teilnehmer die auf die Anfrage antworten können oder nicht. Dieses System kann derart modifiziert werden, daß bei Sendewiederholungsanfragen nur Teilnehmer bis zu einem bestimmten Maximalabstand angesprochen werden. Das wird ohne eine Gruppeneinteilung dadurch erreicht, daß die Lebensdauer von Anfragen und Antworten, durch setzen der TTL, geschickt gewählt werden, so daß weit entfernte Teilnehmer nicht in die Situation kommen, auf Anfragen antworten zu müssen.

7 Bewertung der Protokolle

7.1 Lokale Gruppen (LG)

Speicheranforderungen Alle Teilnehmer müssen über genügend Pufferspeicher verfügen, da Daten innerhalb der lokalen Gruppe ausgetauscht werden sollen.

Verwaltungsaufwand Die Gruppenverwalter sammeln die Einzelquittungen der Empfänger und erstellen die Gruppenstatusmeldungen. Allerdings müssen die Gruppenverwalter zu jedem Mitglied der Gruppe eine 1-1-V unterhalten und verwalten. Dazu kommt die Bearbeitung von An- und Abmeldevorgängen, die Analyse der Gruppenstruktur und die Initiierung von Verwalterwechseln. Um diesen Aufgabenstellungen gerecht zu werden, müssen die Gruppenverwalter über eine ausreichende Prozessorleistung verfügen.

Flexibilität Durch die regelmäßige Analyse der Gruppenstruktur und die Möglichkeit des Verwalterwechsels ohne eine Unterbrechung der Verbindung, ist diese Methode auch bei sehr dynamischen Empfängermengen geeignet.

Quittierungsstrategie Durch die Einteilung der Empfängermenge in Gruppen ist das Bilden von Sammelquittungen möglich. Durch eine Beschränkung der Gruppenanzahl bzw. der Gruppengröße ist das Quittungsaufkommen für den Sender bzw. die Gruppenverwalter beherrschbar.

Auslieferungsverzögerung Da in den meisten Fällen von nicht korrekt empfangenen Datenpaketen ein gruppeninterner Datenaustausch das Problem löst und die Verbindungen innerhalb der lokalen Gruppen im Durchschnitt besser oder kürzer sind als die Verbindung Sender-Empfänger, ist mit geringeren Auslieferungsverzögerungen zu rechnen.

Netzbelastung Durch die Strategie des Informationsaustausches innerhalb der lokalen Gruppe, werden Weitverkehrsverbindungen nur dann zur Übertragung von Wiederholungssendungen benutzt, wenn eine Korrektur innerhalb der lokalen Gruppe nicht möglich ist.

Fehlertoleranz Fällt ein Gruppenverwalter aus, tritt der redundante Gruppenverwalter an seine Stelle. Dadurch kann in den meisten Fällen ein Datenverlust verhindert und sogar gruppenintern behoben werden.

Speicheranforderungen	hoch
Verwaltungsaufwand	erhöht
Flexibilität	sehr hoch
Quittierungsstrategie	gut
Auslieferungsverzögerung	sehr gering
Netzbelastung	global sehr vermindert/lokal erhöht
Fehlertoleranz	sehr gut

7.2 Designated Status Protocol (DSP)

Speicheranforderungen Sowohl der Sender, als auch die Gruppensender müssen über genügenden Pufferspeicher verfügen um Datenpakete zwischenspeichern, bis alle Empfänger im jeweiligen Zuständigkeitsbereich den korrekten Empfang bestätigt haben.

Verwaltungsaufwand Die lokalen Gruppensender sammeln die Einzelquittungen der Empfänger und erstellen die Gruppenstatusmeldungen.

Flexibilität Ein Mechanismus zur Aufnahme neuer Empfänger während der Sendung ist nicht vorgesehen.

Quittierungsstrategie Durch eine Beschränkung der Gruppenanzahl bzw. der Gruppengröße ist das Quittungsaufkommen für den Sender bzw. die Gruppensender beherrschbar, da der Sender nur von den Gruppensendern bzw. der Gruppensender von den Empfängern Quittungen erhält.

Auslieferungsverzögerung Da der lokale Gruppensender Datenpakete zwischenspeichert, müssen verlorene Daten, die der LGS korrekt empfangen hat, nicht noch einmal vom Sender angefordert werden.

Netzbelastung Auch wenn viele verlorene Datenpakete durch eine lokale Übertragungswiederholung die Empfänger erreichen, wiederholt der Sender die Übertragung sobald auch nur ein lokaler Sender ein Datenpaket nicht korrekt empfangen hat.

Fehlertoleranz Es ist kein Mechanismus vorgesehen, der den Ausfall eines lokalen Senders während der Datenübertragung kompensiert. Alle Empfänger der betreffenden Gruppe können ab dem Zeitpunkt des Ausfalls des lokalen Gruppensenders keine Übertragungsfehler mehr korrigieren.

Speicheranforderungen	erhöht
Verwaltungsaufwand	sehr gering
Flexibilität	keine
Quittierungsstrategie	gut
Auslieferungsverzögerung	verkürzt
Netzbelastung	leicht vermindert
Fehlertoleranz	keine

7.3 Consolidated Status Protocol (CSP)

Speicheranforderungen Nur der Sender muß über genügend Speicherkapazität verfügen, um die Daten so lange zu speichern, bis alle Empfänger den korrekten Empfang aller Daten bestätigt haben.

Verwaltungsaufwand Die Router empfangen die Einzelquittungen der Empfänger und erstellen die Gruppenstatusmeldungen.

Flexibilität Ein Mechanismus zur Aufnahme neuer Empfänger während der Sendung ist nicht vorgesehen.

Quittierungsstrategie Durch die Einteilung der Empfängermenge in Gruppen ist das Bilden von Sammelquittungen möglich. Durch eine Beschränkung der Gruppenanzahl bzw. der Gruppengröße ist das Quittungsaufkommen für den Sender bzw. die Gruppensender beherrschbar.

Netzbelastung Der Mechanismus nimmt nur insofern Einfluß auf Wiederholungsübertragungen, als daß durch die Angabe der Speicherkapazitäten der Empfänger das verfrühte Senden von Daten verhindert und somit eine Fehlerquelle beseitigt wird. Dies führt auch zu einer minimalen Entlastung des Gesamtnetzes.

Fehlertoleranz Es ist kein Mechanismus vorgesehen, der den Ausfall eines Rooters während der Datenübertragung kompensiert. Alle Empfänger der betreffenden Gruppe können ab dem Zeitpunkt des Ausfalls des Rooters keine Übertragungsfehler mehr korrigieren.

Speicheranforderungen	normal
Verwaltungsaufwand	sehr gering
Flexibilität	keine
Quittierungsstrategie	gut
Auslieferungsverzögerung	durchschnittlich
Netzbelastung	minimal vermindert
Fehlertoleranz	keine

7.4 Combined Protocol (CP)

Speicheranforderungen Sowohl der Sender als auch die Empfänger müssen über eine ausreichende Speicherkapazität verfügen um Daten zwischenspeichern zu können.

Verwaltungsaufwand Neben dem Sammeln von Einzelquittungen und dem Erstellen von Gruppenstatusmeldungen kommt speziellen Routern die Aufgabe zu, lokale Multicastverbindungen aufzubauen und zu verwalten.

Flexibilität Ein Mechanismus zur Aufnahme neuer Empfänger während der Sendung ist nicht vorgesehen.

Quittierungsstrategie Durch die Einteilung der Empfängermenge in Gruppen ist das Bilden von Sammelquittungen möglich. Durch eine Beschränkung der Gruppenanzahl bzw. der Gruppengröße ist das Quittungsaufkommen für den Sender bzw. die Gruppensender beherrschbar.

Auslieferungsverzögerung Da in den meisten Fällen von nicht korrekt empfangenen Datenpaketen ein gruppeninterner Datenaustausch das Problem löst und die Verbindungen innerhalb der Gruppe im Durchschnitt besser oder kürzer sind als die Verbindung Sender-Empfänger, ist mit sehr geringen Auslieferungsverzögerungen zu rechnen.

Netzbelastung Durch die Gruppenstruktur und die Strategie des Informationsaustausches innerhalb der Gruppe, werden Weitverkehrsverbindungen nur dann zur Übertragung von Wiederholungssendungen benutzt, wenn eine Korrektur innerhalb der lokalen Gruppe nicht möglich ist. Durch den gruppeninternen Datenaustausch ist jedoch mit einer erhöhten Netzbelastung auf „Kurzstrecken“ zu rechnen. Auch wird wieder das verfrühte Senden von Daten, durch Angaben über die zur Verfügung stehenden Speicherkapazitäten, verhindert.

Fehlertoleranz Es ist kein Mechanismus vorgesehen, der den Ausfall eines Rooters während der Datenübertragung kompensiert. Alle Empfänger der betreffenden Gruppe können ab dem Zeitpunkt des Ausfalls des Rooters keine Übertragungsfehler mehr korrigieren.

Speicheranforderungen	erhöht
Verwaltungsaufwand	erhöht
Flexibilität	keine
Quittierungsstrategie	gut
Auslieferungsverzögerung	sehr gering
Netzbelastung	global sehr vermindert/lokal erhöht
Fehlertoleranz	keine

7.5 Tree-Based Multicast Transport Protocol (TMTP)

Speicheranforderungen Sowohl der Sender, als auch die Väter müssen über genügend Pufferspeicher verfügen um Datenpakete zwischenspeichern, bis alle Söhne im jeweiligen Zuständigkeitsbereich den korrekten Empfangen bestätigt haben. Durch die Gliederung der Empfänger Menge in eine Baumstruktur, können sehr lange Wartezeiten entstehen.

Verwaltungsaufwand Die Väter entlasten den Sender durch das zweiginterne Sammeln der Einzelquittungen der Söhne und dem Erstellen von Zweigstatusmeldungen. Dazu kommt die Bearbeitung von Anmelde- bzw. Abmeldevorgängen. Um diesen Aufgabenstellungen gerecht zu werden, müssen die Väter über eine ausreichende Prozessorleistung verfügen. Zu geringe Prozessorleistung führt dazu, daß weniger Söhne verwaltet werden können.

Flexibilität Durch die fest vorgegebene maximale Anzahl von Söhnen kann es geschehen, daß eine Anfrage um Aufnahme, von einem nahe bei einem Vater X liegenden Endgerät, abgelehnt wird und dies zur Bildung eines neuen Zweiges oder einer unnötigen Umrootung zu einem anderen Vater Y führt. Das Ausscheiden eines Vaters aus der Verbindung ist nicht ohne Weiteres möglich, da zuerst alle Söhne auf andere Väter verteilt werden müssen.

Quittierungsstrategie Durch die Baumstruktur und die fest vorgegebene maximale Anzahl von Söhnen ist eine Quittungsflut ausgeschlossen.

Auslieferungsverzögerung Die Baumstruktur sorgt bei Übertragungswiederholungen für einen kürzeren Weg und eine bessere Verbindung als die Verbindung Sender-Empfänger, da die Daten bei jedem Vater zwischen Sender und Empfänger zwischengespeichert werden.

Fehlertoleranz Wenn ein Vater ausfällt, starten die betroffenen Söhne erneut ein Ring-suche um bei anderen Vätern aufgenommen zu werden. Je nach Dauer dieser Reintegrationsphase kommt es zu Auslieferungsverzögerungen oder gar zu Datenverlusten. Das ist abhängig davon, ob die fehlenden Daten noch bei einem Mitglied des neuen Zweiges vorliegen.

Speicheranforderungen	hoch
Verwaltungsaufwand	erhöht
Flexibilität	gering
Quittierungsstrategie	gut
Auslieferungsverzögerung	sehr gering
Netzbelastung	global sehr vermindert/lokal erhöht
Fehlertoleranz	gering

7.6 Log-Based Receiver-reliable Multicast (LBRM)

Speicheranforderungen Da man bei diesem Mechanismus einen oder mehrere spezielle Logging-Server benutzt, sind in dieser Hinsicht keine besonderen Anforderungen an die anderen Teilnehmer der Multicastverbindung gestellt.

Verwaltungsaufwand Die Logging-Server sollten so konstruiert sein, daß mehrere Unicastverbindungen geschaltet und verwaltet werden können. Die genaue Beschaffenheit solcher Server hängt von der zu erbringenden Leistung ab.

Flexibilität Das System verhält sich gegenüber allen Erweiterungen und Verkleinerungen der Empfängermenge sehr flexibel.

Quittierungsstrategie Da nur der Haupt-Logging-Server (LS1) Quittungen an den Sender verschickt und die Verantwortlichkeit für die Zuverlässigkeit der Verbindung bei den Empfängern liegt, können keinerlei Probleme durch eine Quittungsflut entstehen.

Auslieferungsverzögerung Je nach Konstruktion und Anzahl der Logging-Server kann auf Anfragen sehr schnell oder eher langsam reagiert werden. Bei einem durchschnittlich konstruierten (statischen) System ist mit einer leicht verkürzten Auslieferungszeit zu rechnen.

Netzbelastung Durch den Verzicht auf Quittungen wird das Netz entlastet. Bei einem System mit mehreren Logging-Servern ist auf jeden Fall von einer Entlastung der Weitverkehrsverbindungen auszugehen, wenn gute Verbindungen zwischen dem Sender und den Logging-Servern bestehen.

Fehlertoleranz Wie bei jedem zentralisierten System führt auch hier ein Ausfall der Hauptkomponente, des Haupt-Logging-Servers (LS1), zu einer starken Störung des Kommunikationssystems. Schon eine unzuverlässige Verbindung zwischen dem Sender und LS1 zieht längere Auslieferungsverzögerungen nach sich. Es wird davon ausgegangen, daß die Sender-Haupt-Logging-Server-Verbindung gut und schnell ist und die Logging-Server für einen stabilen Dauerbetrieb konstruiert sind.

Speicheranforderungen	keine/extrem hoch bei Logging-Servern
Verwaltungsaufwand	minimal
Flexibilität	sehr hoch
Quittierungsstrategie	sehr gut
Auslieferungsverzögerung	leicht verringert
Netzbelastung	global vermindert
Fehlertoleranz	gut

7.7 Scalable Reliable Multicast (SRM)

Speicheranforderungen Jede Komponente sollte über einen ausreichend großen Puffer verfügen um Wiederholungsanfragen beantworten zu können. Theoretisch funktioniert der Mechanismus auch dann zufriedenstellend, wenn nur ein Bruchteil (z.B. die Hälfte) der Teilnehmer über genügend Speicherplatz verfügt. Dies geht allerdings zu Lasten der Auslieferungsverzögerung und belastet eben diese Teilnehmer besonders stark.

Verwaltungsaufwand Da keine Verwaltungsstruktur besteht, ist nur die Vollständigkeit der Daten zu überprüfen, Sendewiederholungsanfragen zu senden und zu beantworten.

Flexibilität Das System verhält sich gegenüber allen Erweiterungen und Verkleinerungen der Teilnehmermenge sehr flexibel.

Quittierungsstrategie Da die Verantwortung für die Zuverlässigkeit der Verbindung bei den Empfängern liegt, können keine Probleme durch eine Quittungsflut entstehen.

Auslieferungsverzögerung Im günstigsten Fall kann ein sehr nahe gelegener Teilnehmer die Anfrage beantworten, was eine sehr kurze Auslieferungsverzögerung zur Folge hätte. Im ungünstigsten Fall kann der Sender die Anfrage beantworten, womit wegen der Wartelgorithmen eine längere Auslieferungsverzögerung als bei einer 1-1-Verbindungen einträte.

Netzbelastung Da auch Anfragen und Wiederholungssendungen per Multicast abgewickelt werden, ist mit einer stark erhöhten Netzbelastung zu rechnen.

Fehlertoleranz Bei diesem Mechanismus sind alle Teilnehmer, was die Beantwortung von Wiederholungsanfragen betrifft, gleichberechtigt und, abhängig von der Größe des Wiederholungspuffers, gleichgestellt. Dadurch kann eine Multicastverbindung, die sich dieses Mechanismus bedient, nur durch den Zusammenbruch des Netzes oder den Ausfall des Senders gestört werden.

Speicheranforderungen	erhöht
Verwaltungsaufwand	gering
Flexibilität	sehr hoch
Quittierungsstrategie	gut
Auslieferungsverzögerung	verringert
Netzbelastung	sehr hoch
Fehlertoleranz	sehr gut

8 Zusammenfassung

Durch die unterschiedlichen Anforderungen der verschiedenen Multicast-Applikationen wird es auch in Zukunft nur schwer möglich sein, ein allgemeingültiges Verfahren als Standard durchzusetzen. Es kann nicht einmal allgemeingültig entschieden werden, ob man den oder die Sender oder die Empfänger zur Sicherstellung des Sendebetriebs verpflichten soll. Zusammenfassend lassen sich nur folgende Schlüsse und auch diese nur eingeschränkt ziehen: Das Konzept der lokalen Gruppen stellt dann einen praktikablen und lohnenden Ansatz dar, wenn sich ein Datenaustausch innerhalb der lokalen Gruppe sicher und kostengünstig vollziehen läßt. Baumstrukturen erweisen sich für dynamische Systeme als zu starr. Man betrachte den Fall, daß ein neuer Teilnehmer aus einem lokalen Netz einem weit entfernten Vaterknoten zugeordnet wird, weil ein Vaterknoten aus dem selben Netz schon die maximale Anzahl von Söhnen erreicht hat. Dies führt zu dem Schluß, daß sich dynamische Systeme regelmäßig optimieren sollten, was aber durch zunehmende Größe erschwert wird. Kontrollmechanismen, die diese Aspekte abwägen könnten, wären sehr komplex und damit langsam. Eine optimale Einteilung läßt sich also besten Falls bei statischen Gruppen erreichen. Allerdings müßten dann auch die Verbindungen statisch sein. Der Ansatz, die Verantwortlichkeit den Empfängern zu übertragen, ist bei zuverlässiger Kommunikation nur praktikabel, wenn ein Großteil der Sendungen korrekt empfangen wird. Ansonsten der oder die Sender mit Rückfragen überhäuft werden. Dieses Problem läßt sich mit guten „Wartealgorithmen“ (siehe 6.7) zwar mindern, aber nur auf Kosten einer erhöhten Netzbelastung, da zumindest die Übertragungswiederholungen per Multicast übertragen werden müssen.

IP und ATM: Konkurrenz oder Symbiose?

Virginie J. L. Njeck Nounga

Kurzfassung

Im Rahmen dieser Arbeit werden zunächst die Grundprinzipien des *Asynchronous Transfer Mode (ATM)* vorgestellt. Diese umfassen die Adressierung innerhalb von ATM-Netzen, die zellenbasierte Übertragungstechnologie und das Prinzip der virtuellen Pfade und virtuellen Kanäle. Darauf aufbauend werden die Vor- und Nachteile der ATM-Technologie diskutiert. Im Vergleich zu ATM wird anschließend eine Einführung in die Protokollarchitektur *TCP/IP* des Internets gegeben. Aufbauend auf diesen Grundlagen werden die Probleme bei der Integration von ATM in bestehende Netzinfrastrukturen erläutert und Lösungsmöglichkeiten vorgestellt. Dies sind etwa *IP über ATM* und *LAN-Emulation*. Schließlich werden Prognosen für die künftige Entwicklung im Netzwerkbereich diskutiert.

1 Einleitung

Diese Ausarbeitung beschäftigt sich mit dem Thema "IP und ATM: Konkurrenz oder Symbiose?" In Kapitel 2 werden die grundlegenden Prinzipien der ATM-Technologie vorgestellt. In Kapitel 3 werden die Vor- und Nachteile dieser Technologie dargestellt. Kapitel 4 und 5 beschäftigen sich mit den grundlegenden Prinzipien und den Vor- und Nachteilen der Protokollfamilie des Internets. In Kapitel 6 werden die Probleme, die beim Einsatz von ATM auftreten, dargestellt. Kapitel 7 beschäftigt sich mit *IP über ATM* und *LAN-Emulation*. Schließlich werden mögliche Entwicklungen im Netzwerkbereich dargestellt.

2 Grundlegende Prinzipien der ATM-Technologie

In ATM werden zwei Schnittstellen definiert:

- UNI (User Network Interface) als Netzzugangsschnittstelle.
- NNI (Network Network Interface) als Schnittstelle zwischen 2 Vermittlungsknoten.

ATM arbeitet mit Zellen fester Länge. Jede der 53-Byte langen Zelle enthält 48 Datenbytes und fünf Bytes für die Adressierung und Fehlererkennung. Jeder Zellkopf enthält mehrere Felder mit Protokollkontrollinformation [AB94]:

- *GFC* — *Generic Flow Control*: Die Generic Flow Control gewährleistet den geregelten Zugang verschiedener Endsysteme zum ATM-Netz und dient somit der Verkehrskontrolle an der UNI-Schnittstelle. Sie ermöglicht es, ein Endsystem in seiner Aktivität zu bremsen, um einer Überlastsituation im Netz vorzubeugen.

- *VCI/VPI — Virtual Circuit/Path Identifier*: Die Zellen werden entsprechend der Angaben im Header einem virtuellen Kanal und einem virtuellen Pfad (Kanal-Bündel) zugeordnet. Der Parameter VCI stellt die Nummer des virtuellen (logischen) Kanals dar. Der virtuelle Pfad wird durch den Parameter VPI gekennzeichnet.
- *PT — Payload Type*: Der Nutzlasttyp PT dient dazu, innerhalb einer Verbindung zwischen Zellen mit Benutzerinformation und Zellen mit Managementinformation zu unterscheiden.
- *CLP — Cell Loss Priority*: Die Angabe der Zellenverlustpriorität CLP ermöglicht es, verlustempfindliche Zellen innerhalb einer Verbindung zu markieren, damit sie in Überlastsituation nach Möglichkeit nicht verworfen werden. Die Zellen mit niedrigerer Priorität werden vom Netz in Überlastsituation zuerst verworfen.
- *HEC — Header Error Control*: Die Fehlerkontrolle im Zellenkopf ist erforderlich, um Übertragungsfehler zu erkennen und ggf. zu korrigieren. Falls bei der Auswertung nicht korrigierbare Übertragungsfehler im Zellenkopf entdeckt werden, wird die betroffene Zelle verworfen.

ATM realisiert eine asynchrone Übertragung von Zellen fester Länge. Dies bedeutet, daß die Zellen im Gegensatz zur synchronen Übertragung nicht in fest vorgegebenen Zeitschlitzten gesendet werden. Vielmehr können ATM-Zellen "je nach Bedarf" in unregelmäßigen Abständen gesendet werden. Bei ATM werden auf jedem Übertragungsabschnitt ununterbrochen Zellen übertragen. Ist gerade keine Nutzinformation zu senden, so werden speziell markierte Leerzellen gesendet.

ATM ist ein Multiplexverfahren, so daß ein Übertragungsweg von mehreren Verbindungen zeitgleich genutzt werden kann. Dazu werden die zu übertragenden Bitströme in ATM-Zellen segmentiert, die auf den gemeinsamen digitalen Übertragungsweg geschickt werden. Die verfügbare Übertragungskapazität (Bandbreite) eines physikalischen Kanals wird zunächst in virtuelle Pfade aufgeteilt. Ein virtueller Pfad repräsentiert eine Gruppe von virtuellen Kanälen. Beim Aufbau einer virtuellen Verbindung wird zunächst ein virtueller Pfad ausgewählt. Anschließend wird ein Teil der Bandbreite als virtueller Kanal dieses Pfades belegt. Ein physikalischer Übertragungsabschnitt kann — bildhaft gesprochen — als eine breite Straße mit einer großen Anzahl von Fahrspuren angesehen werden. Eine Fahrspur stellt einen virtuellen Kanal dar, und wird mit einer Kennung (VCI — Virtual Channel Identifier) identifiziert. Mehrere Fahrspuren werden zu einem Pfad zusammengefaßt. Für die Identifizierung der Pfade werden VPIs (Virtual Path Identifier) verwendet. Physikalisch gesehen, findet im Breitbandkanal eine serielle Übertragung von ATM-Zellen, in denen die VPI und VCI Angaben enthalten sind, statt.

Ein ATM-Netz kann sowohl privat als auch öffentlich sein. Der ATM-Standard definiert auch die Kopplung von privaten ATM-Netzen über ein öffentliches ATM-Netz. Die öffentliche UNI definiert hierbei die Anbindung eines privaten ATM-Systems an das öffentliche ATM-Netz.

In ATM-Netzen werden zwei Vermittlungsarten unterscheiden:

- Vermittlung von virtuellen Pfaden (VP-Vermittlung): Bei dieser Vermittlung werden nur die VPI(Virtual Path Identifier) Werte umgesetzt.
- Vermittlung von virtuellen Kanälen und virtuellen Pfaden(VC/VP-Vermittlung): Bei dieser Vermittlung werden sowohl die VCI als auch die VPI-Werte umgesetzt.

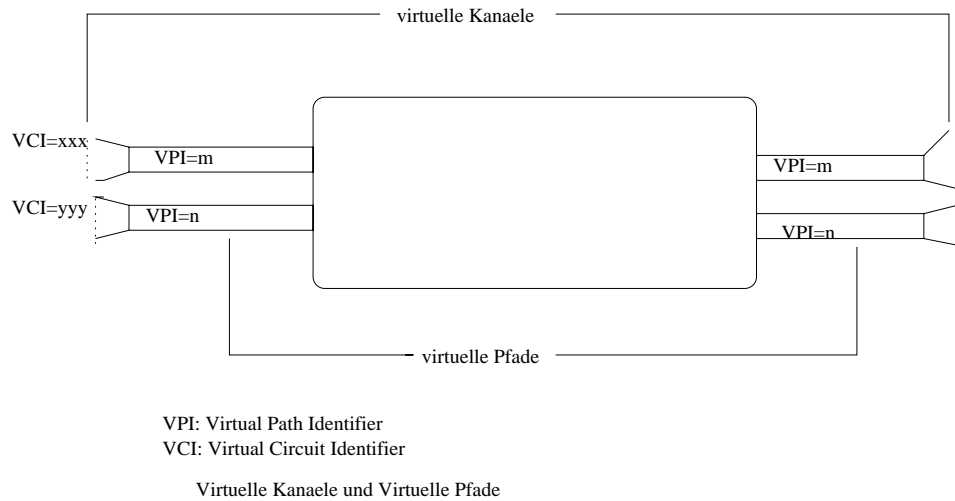


Abbildung 49. Virtuelle Pfade und Kanäle

ATM eignet sich auf Grund seiner Flexibilität für die Nutzung durch zahlreiche unterschiedliche Anwendungen, wie etwa:

- Datenübermittlungsdienste
- Videodienste (z.B. Video-Verteildienste, Videokonferenzen, Video Datenbanken)
- Dienste für die Übermittlung von Multimedia-Information (Z.B. Sprache, Daten und Video komprimiert).
- Sprachübermittlungsdienste.

Diese Anwendungen haben eine Vielzahl unterschiedlicher Anforderungen an die Qualität eines Übertragungsdienstes. Durch die Flexibilität, die Möglichkeit zur Reservierung von Netzressourcen und dem asynchronen Charakter eignet sich ATM als Übertragungstechnologie für ein breites Anwendungsspektrum.

3 Vor- und Nachteile von ATM

In diesem Abschnitt werden Vor- und Nachteile der ATM-Technologie diskutiert [Sch95], [Cid92], [Ste94].

3.1 Vorteile

Ein ATM-Netz eignet sich für die Übertragung von Bitströmen mit unterschiedlichen Bitraten, da ATM eine asynchrone Übertragung unterstützt. Die Kommunikationskomponenten innerhalb eines ATM-Netzes wandeln sowohl zeitkontinuierliche als auch zeitvariable Bitfolgen in eine Folge von Zellen mit fester Länge um. Zwischen den beteiligten Endsystemen können mehrere virtuelle Verbindungen aufgebaut werden. Ein ATM-Netzwerk besteht aus einzelnen Punkt-zu-Punkt Verbindungen. Dies erhöht im Gegensatz zu Netzwerken mit geteilten Medien die Flexibilität, die Wartbarkeit und die Fehlertoleranz.

ATM ist ein verbindungsorientiertes Protokoll, das eine zuvor ausgehandelte Dienstgüte (QoS — Quality of Service) garantiert. Beim Verbindungsaufbau werden wichtige Verkehrsparameter, wie beispielsweise die durchschnittliche Übertragungsrate (MCR — Mean Cell Rate) und die maximale Übertragungsrate (PCR — Peak Cell Rate), zwischen allen beteiligten Kommunikationssystemen ausgehandelt.

3.2 Nachteile

Da jeder ATM-Zelle zusätzliche Protokollkontrollinformationen im Zellkopf vorangestellt werden, resultiert aus der kurzen Zellenlänge (nur 48 Datenbytes für Daten) ein relativ großer Protokolloverhead. Zudem steht, bedingt durch die kurze Zellenlänge und die hohe Übertragungsrate, nur eine äußerst kurze Bearbeitungszeit pro Zelle zur Verfügung. Dadurch ist eine Realisierung von ATM-Komponenten mittels Software kaum zu realisieren. Stattdessen muß auf kostspielige und wenig flexible Hardwarekomponenten zurückgegriffen werden.

Es gibt keine abschnittsweise Fehlerkontrolle: Auf dem Übertragungsabschnitt sendet die Empfangsseite keine Rückmeldung (Bestätigungen, Quittungen) bezüglich der empfangenen Zellen. Wurden in einer Zelle keine Übertragungsfehler festgestellt, so wird sie direkt weitergeleitet. Die Verfälschung eines Bits im Zellkopf wird mit der HEC-Hilfe korrigiert. Die Zellen mit nicht behebbaren Fehlern werden verworfen. Für die Dienste, die eine fehlerfreie Übermittlung von Zellen erfordern, muß eine Fehlerkontrolle zwischen den beteiligten Endsystemen realisiert werden. Ebenso gibt es keine abschnittsweise Flußkontrolle. Weil die Empfangsseite keine Rückmeldungen an die Sendeseite sendet, ist eine Flußkontrolle auf dem Übertragungsabschnitt nicht möglich.

Um den Protokolloverhead nicht noch mehr zu erhöhen, enthalten ATM-Zellen keine Sequenznummern. Da von einem ATM-basierten Übertragungsdienst jedoch eine reihenfolgetreue Auslieferung der Zellen gefordert wird, muß die ursprüngliche Reihenfolge der Zellen bei der Übertragung erhalten bleiben. Dies schließt beispielsweise die Nutzung alternativer Netzwerkpfade aus und schränkt die Struktur von Switching- oder Routingkomponenten stark ein [Ste94].

4 Grundlegende Prinzipien von IP

4.1 Geschichte des Internet

Das Internet kommt ursprünglich vom ARPANET, das im Jahre 1969 im Rahmen eines Forschungsprojekts ins Leben gerufen wurde. Nach ein paar Jahren war das ARPANET viel erfolgreicher als ursprünglich geplant, und zusätzlich begannen viele Organisationen das Netz nicht nur für experimentelle Zwecke zu nutzen. Mehrere Firmen zeigten Interesse an dieser Technologie. Dies führte zu der Änderung des Status vom Forschungsnetz hin zu einem Dienstleistungsnetzwerk. In den frühen achtziger Jahren stellte man das Protokoll zwischen den zentralen Netzknoten auf TCP/IP um.

4.2 Das Internet Protocol IP

Das Internetprotokoll transportiert Daten in Form von Datenpaketen (Datagrammen) über verschiedene Netze hinweg. Im TCP/IP-Protokoll sind die Dateneinheiten der höheren Schichten gewissermaßen in das Datenpaket eingekapselt.

IP ist ein verbindungsloses Protokoll. IP kennt keinen Verbindungszustand. Es ist nicht notwendig, eine Verbindung aufzubauen, bevor man Daten zu einem Rechner senden kann. Vielmehr genügt es, das IP-Paket einfach abzusenden. Zu IP gehört ein einheitliches Adressierungsschema. Jedes IP-Paket enthält zwei Adressen in Form von 32-Bitworten: die Absender- und die Empfängeradresse. Eine Internet-Adresse wird meist in Form von vier, durch Punkte getrennte Bytes notiert. Man spricht in diesem Fall von der Dotted-Quad-Schreibweise. Um das Routen von IP-Paketen zu vereinfachen, unterteilt man die Adresse in zwei Teile: den Netzwerkteil und den Rechnerteil. Ein Router muß, um ein Datenpaket weiterleiten zu können, nur den Netzwerkteil einer Adresse analysieren. Den Rechnerteil einer Adresse wertet erst der Router im Zielnetzwerk aus. Um den verschiedenen Anforderungen gerecht zu werden, was die Größe von Netzwerken angeht, unterscheidet man verschiedene Aufteilungen der 32 Adreßbits:

- Adressen der Klasse A: Hier ist das erste Bit der IP-Adresse “0”. Die folgenden 7 Bits bilden den Netzwerkteil der Adresse und die restlichen 24 Bits bilden den Rechnerteil. Es kann also weniger als 128 verschiedene Klasse-A-Adressen geben, aber jedes dieser Netze kann aus Millionen von Rechnern bestehen. Ein typischer Vertreter ist zum Beispiel das amerikanische MILNET, das sich über das ganze Land erstreckt.
- Adressen der Klasse B: Hier fängt die Adresse mit der Bitfolge “10” an. Die folgenden 14 Bits stellen den Netzwerkteil der Adresse dar, die restlichen 16 Bits den Rechnerteil. Es kann also tausende von Klasse-B-Netzen mit jeweils tausenden von Rechnern geben. Ein typischer Vertreter ist das Campusnetzwerk einer großen Universität.
- Adressen der Klasse C: Hier fängt die Adresse mit der Bitfolge “110” an. Die folgenden 21 Bits stellen den Netzwerkteil der Adresse dar, die restlichen 8 Bits den Rechnerteil. Es kann also Millionen von Netzwerken der Klasse C geben, von denen jedes weniger als 256 Rechner beherbergt. Standardmässig erhält man ein solches Klasse-C-Netz, wenn man ein Netzwerk beim NIC (Network Information Center) beantragt.

- Adressen der Klasse D: Hier fängt eine Adresse mit den Bits “111” an. Sie hat eine spezielle Bedeutung. Sie ist für Gruppenkommunikation (Multicast-Adresse) reserviert.

Klasse A fuer Netze mit mehr als 65.536 Knoten

0	Netz-ID	Knoten-ID
---	---------	-----------

Klasse B fuer Netze zwischen 256 und 65536 Knoten

1	0	Netz-ID	Knoten-ID
---	---	---------	-----------

Klasse C fuer Netze mit weniger als 256 Knoten

1	1	0	Netz-ID	Knoten-ID
---	---	---	---------	-----------

Klasse D fuer Gruppenkommunikation (Multicast)

1	1	1	0	Multicast-Adresse
---	---	---	---	-------------------

Verschiedene Adressklassen von IP

Abbildung 50. IP-Adreßklasse

Normalerweise existiert bei allen Netzwerken eine maximale Größe für ein Datenpaket. Im IP-Kontext nennt man diese Grenze die *Maximum Transmission Unit (MTU)*. Die Netzwerkschicht teilt IP-Pakete, die größer als die MTU des verwendeten Netzwerkes sind, in kleinere Stücke, sogenannte Fragmente auf. Der Zielrechner setzt diese Fragmente dann wieder zu vollständigen Paketen zusammen, bevor er sie an die darüberliegende Protokollschicht weiterreicht[Kön93].

5 Vor- und Nachteile von IP

Vorteile Das Internet-Protokoll bietet dem Anwendungsprogrammierer eine Kommunikationsschnittstelle an, die von den unterliegenden Übertragungstechnologien unabhängig ist und diese vor dem Anwender verbirgt. Zudem erlaubt es die netzwerkweite, einheitliche Adressierung von Kommunikationsteilnehmern. Auf den höheren Protokollschichten bietet die Internet-Protokollfamilie mehrere standardisierte Protokolle für die verschiedenen Netzwerkdienste an. Es ist eine offene Protokollfamilie, die unabhängig von Herstellern entwickelt wurde.

Nachteile In IP ist die Dienstqualität nicht gewährleistet. IP garantiert nicht, daß ein Datenpaket überhaupt beim Empfänger ankommt. Wenn ein Datenpaket abgeschickt

wird, können mehrere Kopien davon beim Empfänger ankommen. Schließlich ist IP kein reihenfolgeerhaltendes Protokoll. Denn IP garantiert nicht, daß Datenpakete in Reihenfolge beim Empfänger ankommen, in der sie beim Sender in das Netz eingespeist wurden [Kön93].

6 Probleme bei der Integration von ATM

Die Struktur von ATM-basierten Netzen und diejenige der heute weit verbreiteten Netze unterscheiden sich stark. ATM-Netze verlangen nach einer vollständig neuen Adreßstruktur.

ATM definiert im Prinzip drei verschiedene Adreßtypen von jeweils 20 Byte Länge. Jede ATM-Adreßstruktur hat ein Routing-Feld, eine Endstationskennung, und ein Selector-Byte (SEL). Die Adreßstrukturen unterscheiden sich durch ihr Format. Die verschiedenen Formate der ATM-Adreßstrukturen sind:

- DCC: Data-Country-Code-Format
- ICD: International-Code-Designator-Format
- E164: Private-Adress-Format

Dabei ist der erste Teil der Adresse der Identifikationsteil für das Netzwerk, der zweite ein Stationsidentifikator. Bei der Initialisierung eines Rechners tauschen der ATM-Switch und die angeschlossene Endstation Informationen aus, die zu einer Registrierung der Station und ihrer Adresse führen. Diese Prozedur ist durch das Signaling and *Interim Link Management Interface (ILMI)* des ATM-Forums beschrieben. Leider sind die ATM-Adressen mit den oben beschriebenen und weit verbreiteten IP-Adressen nicht identisch. Es stellte sich nun die Frage, auf welcher Ebene eine Abbildung beziehungsweise Zuordnung der ATM-Adresse zu einer korrespondierenden IP-Adresse vonstatten gehen soll. Grundvoraussetzung bei der Definition der Abbildungsvorschrift sollte sein, daß die existierenden Adressen transparent für ein ATM-Netzwerk sind, und daß heutige Anwendungen aufgrund ihrer Entwicklungskosten auch beim Einsatz von ATM unverändert bleiben können.

Herkömmliche Netzwerktreiber bieten in der Regel keinen verbindungsorientierten Dienst an. Die von ihnen bereitgestellten Dienste sind jedoch wohldefiniert und können nicht verändert werden. Würde man dies tun, so müßten sämtliche Anwendungen umgeschrieben werden. Ein nicht realisierbarer Aufwand. Statt dessen müssen Netzwerktreiber entwickelt werden, so daß ein ATM-Interface zur Verfügung steht, welches wie ein verbindungsloser Treiber auf der Sicherungsschicht funktioniert.

Mit der Einführung von ATM ergibt sich eine vermaschte Netzstruktur, die aus vielen einzelnen, voneinander unabhängigen Datenkanälen besteht und im gesamten leicht einen Durchsatz von mehreren Millionen Zellen pro Sekunde erreichen kann. Echte Hochleistungsbackbones verlangen deshalb nach einer Lösung, die effektiv und kostengünstig eine Abbildung zwischen verbindungslosem und verbindungsorientiertem Datenverkehr

herstellt, wobei innerhalb des aufzubauenden Netzes so viele Komponenten wie möglich (außer Routern auch Server und Endknoten) direkt an das ATM-Netz angeschlossen werden sollten [Dei95].

ATM befindet sich hinsichtlich des täglichen Einsatzes in Dienstleistungsnetzen noch in den Anfängen, und es wird noch längere Zeit dauern, bis ATM flächendeckend verfügbar sein wird. ATM verspricht jedoch die Fähigkeit, den gesamten Datenverkehr mit verschiedenen Übertragungsqualitäten (Quality of Service) handhaben zu können. Man kann zum Beispiel eine Telefonkonferenz in Echtzeit mit einem Datentransfer im Hintergrund zeitgleich durchführen. ATM wird nur dann ein Erfolg werden, wenn alle Implementierungen folgende Punkte realisieren:

- Überlastkontrolle
- Verkehrsmanagement
- Sicherheitsmechanismen
- Erfolgreiche Integration in existierende und weltweit anerkannte Protokollarchitekturen und Netzstrukturen, wie zum Beispiel dem Internet

Für die Integration von ATM in existierende Netzstrukturen sind mehrere Möglichkeiten denkbar. Zwei von ihnen werden in dem nächsten Abschnitt näher betrachtet [Lau94].

7 Integration von ATM in existierende Netze

Bisher wurden bereits hohe Summen in bestehende Netzstrukturen investiert. Daher muß aus finanziellen Gründen eine sanfte Migration hin zu ATM vorgesehen werden. Insbesondere stellt das Internet mit seiner Protokollarchitektur einen weltweit anerkannten und weit verbreiteten Standard dar, weswegen eine Integration von ATM in das Internet notwendig erscheint.

7.1 Probleme bei der Integration von IP und ATM

Problem Bei der Integration von IP und ATM stellt sich das Problem der Adreßauflösung, d.h. der Abbildung logischer IP-Adressen auf physikalische Adressen auf der Schicht 2 (ATM, Ethernet oder Token Ring). Das bisher für die Adreßauflösung eingesetzte ARP (Address Resolution Protocol) legt ein Rundruf-Medium zugrunde (geteiltes Medium). Auf der anderen Seite unterstützt ATM nur Punkt-zu-Punkt-Verbindungen (dediziertes Medium). Bei herkömmlichen Netzwerken schickt ARP einen Rundruf an alle Teilnehmer, um die Adresse einer Endstation zu erfahren. Dies ist bei der ATM zugrundeliegenden Sternstruktur nicht mehr möglich.

Lösungsansätze Um das Problem zu lösen, gibt es zwei verschiedene Möglichkeiten zur Integration von ATM in bestehende Netzstrukturen:

- ATM als MAC-Protokoll (LAN-Emulation)
- ATM als LLC für IP (IP über ATM)

7.2 LAN-Emulation

LAN-Emulation als Ansatz der SWG-Drafting-Gruppe innerhalb des ATM-Forums beschreibt eine Möglichkeit zur Bereitstellung einer traditionellen Schnittstelle auf der Sicherungsschicht für ATM-Netzwerke. Dazu wird auf dem ATM-Adapter einer Endstation ein LAN-Emulation-Client installiert, der dem Endanwender das Gefühl vermittelt, sich in einem klassischen LAN zu bewegen. Dieser soll sowohl Token-Ring (IEEE.802.5) als auch Ethernet (IEEE.802.3) unterstützen.

Client-Server-Architektur Auf Seite des Endanwenders gibt es pro Rechner und pro emuliertem LAN, an das der Rechner angeschlossen ist, einen LAN-Emulation-Client (LE-Client). Die Server-Komponente (LAN-Emulation-Service — LE-Service) zerfällt in drei Teile, die zentral oder verteilt realisiert sein können. Die sind:

- Der LEC-Server (LAN-Emulation Configuration-Server): Er verwaltet eine Datenbank mit den Konfigurationsdaten der emulierten LANs (z.B. Adressen ihrer LE-Server). Zusätzlich gibt es eine bidirektionale Verbindung (Configuration Direct) zwischen LE-Client und LEC-Server zum Austausch von Konfigurationsdaten.
- Der LE-Server (LAN-Emulation-Server): Jedes emulierte LAN enthält genau einen LE-Server. Der LE-Server ist zuständig für die Auflösung von ATM-Adressen. Es gibt eine bidirektionale Punkt-zu-Punkt-Verbindung (Control Direct) zwischen LE-Client und LE-Server zur Übertragung von Adreßauflösungs-Anfragen und -Antworten. Optional ist eine unidirektionale Punkt-zu-Mehrpunkt-Verbindung (Control Distribute) zwischen LE-Server und allen LE-Clients.
- der BUS (Broadcast-and-Unknown-Server): Jedes emulierte LAN enthält genau einen BUS. Der BUS ist zuständig für die Übermittlung von Multicast und Broadcast-Nachrichten ("Multicast-Server"). Es gibt eine bidirektionale Punkt-zu-Punkt-Verbindung (Multicast Send) zwischen LE-Client und BUS zur Übermittlung von Multicast-Nachrichten und eine unidirektionale Punkt-zu-Mehrpunkt-Verbindung (Multicast Forward) zwischen BUS und allen LE-Clients zur Verteilung von Multicast-Nachrichten.

Funktionsweise Der Protokollablauf bei der LAN-Emulation untergliedert sich in verschiedene Teile:

- Konfiguration

- Registrierung beim LE-Server
- Registrierung beim BUS
- Datentransfer

Vor der ersten Datenübertragung durchläuft ein LE-Client verschiedene Initialisierungsphasen. Während dieser Phasen werden die unterschiedlichen ATM-Verbindungen zwischen den LAN-Emulation-Komponenten aufgebaut. Als Beispiel sei der Aufbau von bidirektionalen Kontrollverbindungen und Punkt-zu-Mehrpunkt-Verbindungen zwischen dem LE-Client und den verschiedenen Elementen des LE-Service genannt. Ebenso werden die notwendigen Konfigurationsparameter (z.B. ATM und MAC-Adressen, LAN-Typ) ausgetauscht. Der Abbau der ATM-Verbindungen erfolgt zeitbergesteuert.

Datentransfer: Beim Datentransfer unterscheidet man zwischen Unicast-Datentransfer und den Multicast- und Broadcast-Datentransfer. Beim Unicast-Datentransfer findet die Abbildung der MAC-Adresse auf die korrespondierende ATM-Adresse mit Hilfe des LE-Server statt. Während der Adreßauflösung kann der Sender bereits Daten über den Broadcast-and-Unknown Server aussenden. Sobald die ATM-Zieladresse bekannt ist, baut der Sender eine direkte ATM-Verbindung zum Empfänger auf oder benutzt eine bereits existierende ATM-Verbindung. Beim Multicast- und Broadcast-Datentransfer übermittelt der sendende LE-Client die Nachrichten über die Multicast-Send-Verbindung zum Broadcast-and-Unknown-Server. Der Broadcast-and-Unknown-Server verteilt dann die Nachrichten über die Multicast-Forward (Punkt-zu-Mehrpunkt-Verbindung zwischen BUS und LE-Clients) -Verbindung an alle Zielsysteme.

7.3 IP über ATM

Zur Unterstützung von IP über ATM werden logische IP-Subnetze (LIS) spezifiziert. Darunter versteht man ein lokales ATM-Netz mit einer eindeutigen IP-Netzadresse. Die logischen Subnetze sind über IP-Router miteinander verbunden. Die Stationen müssen IP-Adressen mit Hilfe des ATMARP-Protokolls auf ATM-Adresse abbilden.

Adreßauflösung Der ATMARP-Server verwaltet innerhalb des LIS eine globale Adreßauflösungstabelle. Jede Station kennt die ATM-Adresse des ATMARP-Servers. Es gibt bidirektionale Verbindungen zwischen den Stationen und dem ATMARP-Server. Falls die lokale Adreßtabelle einer Station keinen Eintrag für eine bestimmte IP-Adresse enthält, wird eine Anfrage an den ATMARP-Server gestellt (kein Broadcast).

ATMARP-Protokoll-Ablauf Der ATMARP-Protokoll-Ablauf gliedert sich in zwei Phasen:

- Die Registrierphase: Der Client baut eine bidirektionale ATM-Verbindung zum ATMARP-Server auf. Der Server sendet einen InATMARP-Request, um die IP-Adresse

des Clients zu erfahren. Der Client sendet seine IP-Adresse in einen InATMARP-Reply an den ATMARP-Server. Der ATMARP-Server trägt daraufhin die Adresse in seine Tabelle ein.

- Adreßauflösung: Jeder Client verwaltet eine lokale Adreßtabelle. Ist eine Adreßauflösung über die lokale Tabelle nicht möglich, dann stellt er eine Anfrage (ATMARP-Request) an den ATMARP-Server. Der ATMARP-Server antwortet mit einem ATMARP-Reply, das die gesuchte ATM-Adresse des Systems enthält oder einer ARP-NAK falls die Adreßauflösung nicht durchgeführt werden konnte.

IP über ATM ist vergleichbar mit IP über Ethernet/802.3. Es gibt aber ein paar Unterschiede.

Gemeinsamkeiten

- Beide Subnetzarchitekturen stellen den IP-Protokollimplementierungen die gleiche Schnittstelle zur Verfügung.
- Die IP-Netzadresse wird dem ATM-Netz wie eine IP-Adresse eines Ethernet-Netzes zugewiesen. Verbunden mit der Subnetz-Maske, definiert diese IP-Adresse das logische IP-Unternetz (LIS).

Unterschiede

- Ethernet/802.3 ist ein Netz, dem ein geteiltes Medium zugrundeliegt. Alle Stationen teilen sich das gleiche Kabel, weshalb jede Station die gesendeten Daten anderer Stationen im selben Subnetz mithören und Konflikte entdecken kann. ATM liegt ein dediziertes Medium zugrunde. Jede Station hat ihre eigene Anbindung an das Netz.
- Ethernet/802.3 wird durch mehrere technische Richtlinien (z.B. Länge eines Subnetzes, minimale Rahmenlänge) eingeschränkt.
- Im Ethernet/802.3-Netzwerk warten die Stationen mit dem Senden, bis das gemeinsam benutzte Medium nicht mehr belegt ist. Erst danach können sie ihre Daten übertragen. In ATM-Netzen wird zuerst ein virtueller Kanal zwischen der übertragenden und den empfangenden Stationen aufgebaut. ATM benutzt PVCs (Permanent Virtual Channels) und SVC (Switched Virtual Channels). PVCs werden "von Hand" durch die Netzwerkverwalter initialisiert. SVCs werden dagegen mittels der ATM-Signalisierung auf- und abgebaut.
- Im Ethernet/802.3-Netzwerk sind an der Adreßauflösung mittels ARP (Address Resolution Protocol) alle Stationen unmittelbar beteiligt. Zum Beispiel schickt eine Station eine ARP-Anforderung an alle Stationen im Ethernetsegment und die verantwortliche Station antwortet. In ATM wird hierzu eine Server-basierte Adreßauflösung eingesetzt [Lau94].

8 Prognose für künftige Entwicklungen

ATM ist ein revolutionärer Netzstandard, der in naher Zukunft die Kommunikationstechnologie nachhaltig beeinflussen wird. Im Gegensatz zu High-Speed-Ethernet oder FDDI2 ist der Asynchrone Transfermodus eine vollkommen neue Entwicklung ohne die prinzipiellen Leistungs- und Distanzbegrenzung der herkömmlichen Systeme. Doch anders als es oft zu lesen ist, wird ATM nicht(nur) als Konkurrent heutiger Netze antreten; ATM versucht heutige Standards zu integrieren, um Investitionen zu schützen und einen allmählichen Übergang möglich zu machen.

8.1 Bewertung der Ansätze

LAN-Emulation ermöglicht den Aufbau von virtuellen LANs, also die Emulation mehrerer lokaler Netze innerhalb eines ATM-Netzes.

IP über ATM gemäß RFC 1577 hält das Router-Modell für den Backbone-Bereich aufrecht und wird deshalb von den Herstellern solcher Systeme befürwortet. Durch dieses Prinzip greift man auf die Filter- und Sicherheitsmöglichkeit traditioneller Router zurück. Dieser Ansatz ist einfach und bereits auf einigen ATM-Komponenten verfügbar, jedoch auf IP-Netzwerke beschränkt.

LAN-Emulation besitzt eindeutig den Vorteil des Aufbaus von virtuellen LANs, ist jedoch dem klassischen Bridging ähnlich. IP über ATM hat durch seine Verfügbarkeit am Markt einen gewissen Vorsprung, ist jedoch auf ein Protokoll beschränkt. RFC 1577 wird deshalb von vielen auch als reine Interimslösung bis zur Realisierung echter ATM-APIs betrachtet.

Virtuelle LANs in Verbindung mit LAN-Emulation erlauben erstmalig den Aufbau logischer und nicht mehr physikalischer zusammenhängender Arbeitsgruppen. Die virtuellen LANs bieten hierbei die Sicherheits- und Filterfunktionen, wie man sie bis dato bei zentral im Netz platzierten Routern schätzte. Die Router wandern an die Außenränder des Gesamtnetzes und verbinden die virtuellen LANs lokal miteinander [Dei95].

Protokollmechanismen zur Lokalisierung von Diensten im Internet

Manfred Rohrmüller

Kurzfassung

Mit steigender Teilnehmerzahl im Internet und der damit verbundenen zunehmenden Netzbelastung wird eine Regulierung des weltweiten Datenaustauschs immer notwendiger. Diese Regulierung kann durch die Replizierung von Diensten und Dokumenten erfolgen, so daß Netzbewutzer vorzugsweise auf Dienste in ihrer Umgebung zurückgreifen können. In diesem Text werden zuerst verschiedene Metriken zur Erfassung der Entfernung zweier Rechner in einem Netzwerk und daran anschließend Protokollmechanismen zur Auffindung „optimaler“ Anbieter von replizierten Diensten besprochen.

1 Einführung in die Problematik

Mit zunehmender Popularität verteilter Informationssysteme im Internet (wie z.B. dem World Wide Web) werden die Probleme der steigenden Netzwerkbelastung durch eine drastisch verlängerte Antwortzeit offensichtlich. Zu diesen Problemen gehören die steigende Belastung der Server von vielbenutzten Dokumenten, die unnötige Belegung von Netzwerkkapazitäten durch den Austausch redundanter Informationen und die langen Wartezeiten bei der Übertragung von Dokumenten über leistungsschwache Leitungen.

Eine Möglichkeit, die Netzwerkbelastung zu verringern, besteht in der Replizierung von Diensten oder Dokumenten, d.h. eine Dienstleistung wird von mehreren Servern an verschiedenen Standorten in gleicher Form angeboten. Das Ziel einer solchen Replizierung ist die Beschränkung des Datenaustauschs auf lokale Bereiche und wenig belastete Netzwerkverbindungen.

In diesem Text werden zunächst verschiedene Metriken zur Bestimmung der Distanz zwischen zwei Rechnern vorgestellt und deren Eignung für einen Einsatz im Internet besprochen. Dies sind im wesentlichen die Anzahl der Hops und die Paketumlaufzeit. Solche Distanzmetriken werden zur Definition des „lokalen Bereichs“ eines Rechners benötigt. Darauf aufbauend werden die drei Methoden Anycast, Suche mit Verbindungsgraphen und Triangulation zur Auswahl „optimaler“ Dienstanbieter vorgestellt.

Vorteile einer dynamischen Serverauswahl: Bei der Benutzung von replizierten Diensten wird eine Strategie zur Auswahl eines geeigneten Anbieters benötigt. Eine statische Vorgabe des zu nutzenden Servers ist nur in den Fällen sinnvoll, in denen der optimale Dienstanbieter bekannt, immer verfügbar und nicht vom gegenwärtigen Netzwerkzustand abhängig ist. Im Normalfall ist eine dynamische Auswahl erforderlich, welche den aktuellen Netzwerkzustand (wie z.B. Netzwerktopologie, Verfügbarkeit von

Servern, Netzwerkauslastung) berücksichtigt. Eine dynamische Strategie ist auch im Bereich der Mobilkommunikation notwendig, wobei die Auswahl vom jeweiligen Standort des Clients abhängt.

Bild 51 zeigt den Nutzen einer dynamischen Serverauswahl anhand einer Messung von Zugriffszeiten auf WWW-Dokumente mit unterschiedlichen Größen [CC95]. In diesem Diagramm ist die Abhängigkeit der Zugriffszeit von der Auswahl des Servers und der Dateigröße dargestellt.

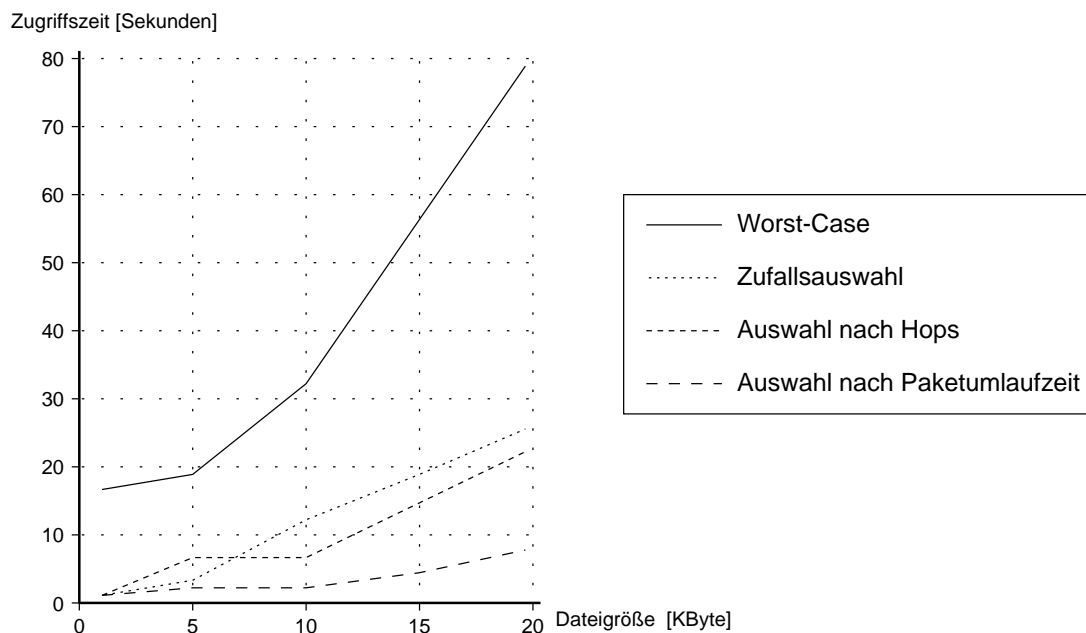


Abbildung 51. Zugriffszeiten bei dynamischen Auswahlstrategien

Zur Ermittlung der Daten in Bild 51 wurden zunächst 10 WWW-Server mit dem gleichen Dienstangebot fest vorgegeben, und dann je ein Server für jedes der drei folgenden Kriterien ausgewählt (diese „Distanzmetriken“ werden in Abschnitt 2 genauer beschrieben):

- Auswahl nach dem Kriterium geringste Anzahl von Hops,
- Auswahl nach dem Kriterium beste durchschnittliche Paketumlaufzeit,
- zufällige Auswahl eines Servers¹⁶.

Die eigentliche Messung bestand darin, daß für jeden der so ausgewählten Server die Zugriffszeit für ein Dokument der Größe 5, 10, 15 und 20 KByte gemessen wurde. Zudem wurde für jede Dokumentgröße die jeweils längste Zugriffszeit unter allen 10 Servern ermittelt und im Diagramm eingetragen. Man sieht, daß sich durch die dynamische Auswahl eines Servers ein zeitlicher Vorteil gegenüber dem Worst-Case ergibt, welcher mit zunehmender Dokumentgröße steigt. Ebenso zeigt sich die Abhängigkeit der Auswahl

¹⁶ Die Zufallsauswahl stellt keine Distanzmetrik im eigentlichen Sinn dar, und wird hier auch nicht weiter besprochen.

eines Servers von der verwendeten Metrik, was zu unterschiedlichen Zugriffszeiten führt. Bemerkenswert ist auch, daß selbst eine Zufallsauswahl zeitlich bessere Ergebnisse als der Worst-Case liefert.

2 Distanzmetriken

2.1 Anzahl der Hops

Als einfachste und naheliegendste Metrik soll zunächst die Anzahl der Hops zwischen zwei Rechnern betrachtet werden. Diese Metrik entspricht der Anzahl der Vermittlungsstellen, die eine Nachricht vom Sender zum Empfänger mindestens zurücklegen muß, und wird oftmals auch einfach als die „Entfernung“ zwischen zwei Rechnern bezeichnet.

Diese Entfernung zweier Rechner kann dadurch ermittelt werden, daß mehrere Nachrichten vom Sender an den Empfänger mit variierenden TTL-Werten¹⁷ gesendet werden. Beginnend bei einem initialen TTL-Wert wird dieser analog zu einem binären Suchalgorithmus erhöht oder erniedrigt, bis die Entfernung zum Empfänger eindeutig feststeht. Zu beachten ist hierbei, daß nicht wie bei einem normalen binären Suchen der TTL-Wert auf Gleichheit geprüft werden kann, da vom Empfänger keine Rückmeldung über den erhaltenen TTL-Wert erfolgt. Die Entfernung muß also durch eine obere und eine untere Schranke eingegrenzt werden. Der initiale TTL-Wert ist vom jeweiligen Netzwerk abhängig, und sollte etwa der durchschnittlichen Entfernung zweier Rechner entsprechen.

In einer Untersuchung wurden die Entfernungen von einem festen Rechner zu 5000 zufällig ausgewählten WWW-Servern gemessen [CC95]. Das Ergebnis ist in Bild 52 dargestellt. Die durchschnittliche Entfernung zweier Rechner im Internet beträgt etwa 17 Hops. Wie dieses Diagramm zeigt, liegen die am häufigsten auftretenden Entfernungen in einem engen Bereich beieinander. Da die geographischen Entfernungen zwischen zwei Rechnern aber nicht so dicht beieinander liegen, läßt sich aus der Anzahl der Hops also nicht auf die räumliche Distanz der beiden Rechner schließen. Insbesondere garantiert eine geringe Entfernung gemäß der Metrik „Anzahl der Hops“ also nicht, daß sich beide Rechner in derselben Region des Netzwerkes befinden.

Zusätzlich zu dem oben beschriebenen Algorithmus zur Ermittlung der Entfernung zwischen zwei Rechnern A und B wird bei der Serverauswahl durch Triangulation (im Abschnitt 3.4) ein Verfahren benötigt, mit dem diese Entfernung von einem dritten Rechner C aus bestimmt werden kann. Ein solches Verfahren liefert die nachfolgend beschriebene *AVG*-Funktion.

Die *AVG*-Funktion: Mit dieser Funktion kann aus den gegebenen Hop-Entfernungen eines Clients C und eines Servers S zu einem oder mehreren ausgewählten Rechnern (man könnte diese Rechner in diesem Zusammenhang als „Vermessungspunkte“ bezeichnen) ein Wert für die Entfernung zwischen C und S berechnet werden.

¹⁷ TTL=„Time-To-Live“, wobei dieser Wert im Internet die Anzahl der Teilstrecken angibt, über die eine Nachricht maximal weitergeleitet wird.

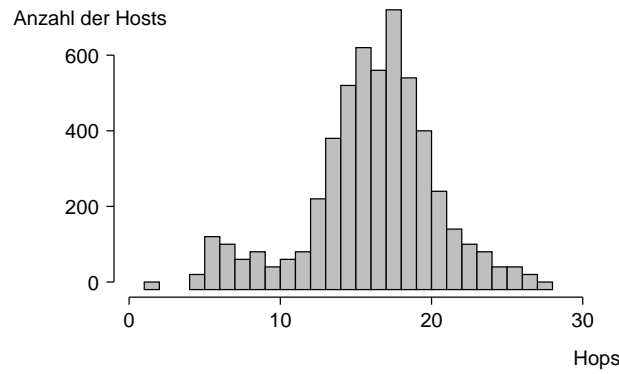


Abbildung 52. Entfernungen zu WWW-Servern im Internet

Dazu wird für jeden Server ein N -Tupel $\langle s_1, s_2, \dots, s_N \rangle$ als die Distanz zu jedem der N -Vermessungspunkte definiert. Dann kann die Entfernung zwischen einem Server an den Koordinaten $S = \langle s_1, s_2, \dots, s_N \rangle$ und einem Client bei $C = \langle c_1, c_2, \dots, c_N \rangle$ mit der *AVG*-Funktion bestimmt werden:

$$AVG(S, C) = \frac{MAX(S, C) + MIN(S, C)}{2}$$

wobei MIN und MAX definiert sind als:

$$MIN(S, C) = \max(|s_1 - c_1|, |s_2 - c_2|, \dots, |s_N - c_N|)$$

$$MAX(S, C) = \min(s_1 + c_1, s_2 + c_2, \dots, s_N + c_N)$$

Eine beispielhafte Erläuterung dieser Funktion gibt Bild 53. Wird das Netzwerk vom Punkt B aus vermessen, so kann ein Server S mit der Distanz $S = \langle 2 \rangle$ an den Stellen s' oder s'' liegen, ein Client C mit $C = \langle 3 \rangle$ könnte sich in diesem Fall bei c' oder bei c'' befinden. Die Entfernung zwischen C und S wird also durch $|C - S|$ und $C + S$ begrenzt. Verallgemeinert man diese Überlegung auf N solche Vermessungspunkte, so erhält man die obigen MIN und MAX Funktionen.

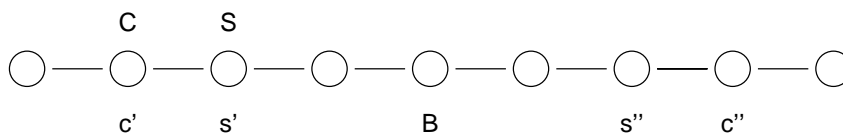


Abbildung 53. Erklärung der *AVG*-Funktion

2.2 Paketumlaufzeit

Die Paketumlaufzeit gibt die Zeitdauer an, die vom Absenden einer Nachricht bis zum Empfang der Antwort vergeht. Sie stellt also die Antwortzeit dar, in der sowohl die Übertragungszeiten von Anfrage und Antwort, als auch die Bearbeitungsdauer des Auftrags beim Server enthalten sind.

Auch bei dieser Metrik kann nicht entschieden werden, ob sich zwei Rechner in derselben Netzwerkregion befinden, da eine Übertragung über große Entfernungen durchaus schneller sein kann, als eine Übermittlung in einer Netzwerkregion (z.B. bei schnellen Backbone-Verbindungen). Ein weiteres Problem der Paketumlaufzeiten sind mögliche Schwankungen bei den Übertragungszeiten innerhalb des Netzwerkes. Diese Schwankungen müssen bei der Messung und der weiteren Verwendung der Umlaufzeiten mitberücksichtigt werden. Daher ist es unbedingt notwendig die Paketumlaufzeit zu einem Rechner mehrmals, über einen möglichst großen Zeitraum verteilt zu messen.

Eine Untersuchung der Paketumlaufzeiten zu den 5000 WWW-Servern, für die auch die Anzahl der Hops gemessen wurden, ergab die in Bild 54 dargestellten Werte [CC95].

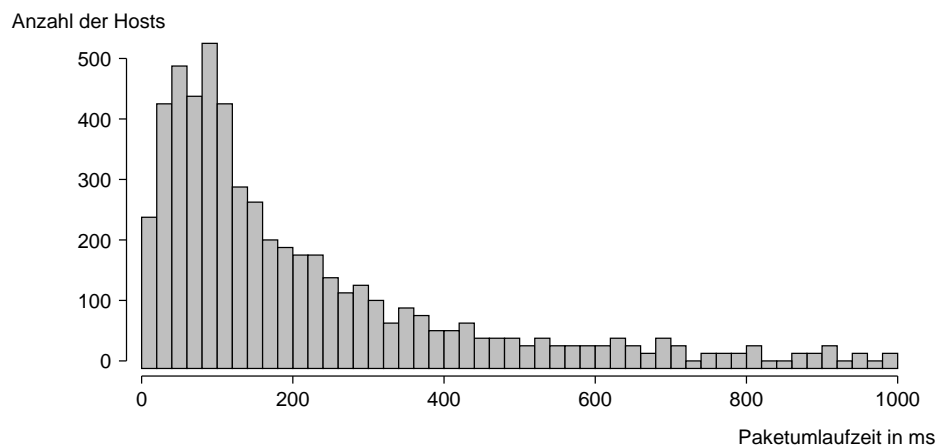


Abbildung 54. Paketumlaufzeiten zu WWW-Servern im Internet

2.3 Vergleich von Paketumlaufzeit und Anzahl der Hops

Vergleicht man die gemessenen Entfernungen und Paketumlaufzeiten im Internet aus Bild 52 und Bild 54, so erkennt man keinen Zusammenhang zwischen der Entfernung und der zugehörigen Paketumlaufzeit. Eine kurze Antwortzeit wird offensichtlich nicht durch eine geringe Anzahl von Hops impliziert.

Die fehlende Korrelation zwischen den beiden Metriken legt die Vermutung nahe, daß entweder Schwankungen in der Übertragungszeit oder Verzögerungen aufgrund starker Belastung des Netzwerkes oder des Servers die Paketumlaufzeit nachhaltig beeinflussen.

Die Auswirkung der Netzwerkbelastung auf die Paketumlaufzeit zeigt folgende Untersuchung [CC95]: Gemessen wurde die Paketumlaufzeit zu einem einzelnen Host periodisch alle 30 Sekunden über eine Dauer von 2 Tagen. Die in den Abschnitten 2.1 und 2.2 beschriebenen Untersuchungen ergaben bei diesem Rechner eine Entfernung von 19 Hops zum Meßrechner und eine Paketumlaufzeit von 289 ms. Die periodischen Messungen der Umlaufzeiten ergaben dagegen Werte zwischen 80 ms und 740 ms, wobei der Großteil der Werte zwischen 120 ms und 380 ms lag. Dabei zeigten sich bereits in Zeitintervallen von wenigen Minuten große Schwankungen bei den Antwortzeiten.

Dieses Experiment zeigt auch, daß Verfahren, welche als Metrik nur die Anzahl der Hops

benutzen, oftmals keine zeitoptimalen Server finden können, andererseits ist die Paketumlaufzeit aufgrund der gezeigten Schwankungen in vielen Protokollen nicht effizient einsetzbar, da meist nur mit Mittelwerten gerechnet werden kann.

Die starken Unterschiede zwischen der Anzahl von Hops und der Paketumlaufzeit haben auch direkte Auswirkungen auf die Platzierung von replizierten Servern. Werden in einem Netzwerk Server nach dem Kriterium Anzahl der Hops ausgewählt, so sind Replikate anders zu platzieren, als wenn nach dem Kriterium Paketumlaufzeit ausgewählt wird.

2.4 Alternative Metriken

Als Alternative zu den oben vorgestellten Metriken wäre z.B. die Einteilung von Rechnerentfernungen anhand der entstehenden Übertragungskosten denkbar. Dies könnte z.B. dadurch erfolgen, daß die einzelnen Übertragungsstrecken jeweils mit ihren Gebühren gewichtet werden. Als weitere Kosten könnten eventuell anfallende Gebühren bei der Auftragsbearbeitung durch den Server mit berücksichtigt werden. Ein Problem hierbei wäre aber sicherlich die Überlastung von günstigen Strecken (oder billigen Servern), welche zudem oftmals eine geringere Kapazität als kostenintensivere Verbindungen besitzen.

Eine weitere Möglichkeit bestünde darin, den Auslastungsgrad eines Servers oder einer Netzwerkverbindung als Entscheidungskriterium zu verwenden. Bei einem solchen Verfahren müßte die momentane Belastung der einzelnen Verbindungen und Server ständig erfaßt werden. Es wäre also eine Art Netzwerkmonitor erforderlich, der jederzeit darüber Auskunft gibt, in welchem Teil des Netzes noch Kapazitäten vorhanden sind.

Abschließend bleibt zu diesen Metriken zu sagen, daß diese in der Praxis sicher nur schwer zu handhaben sind, und oftmals einen nicht unerheblichen Verwaltungsaufwand erfordern. In der Regel werden daher wohl nur die Metriken Entfernung und Paketumlaufzeit benutzt.

3 Verfahren und Protokollmechanismen zur Auswahl „optimaler“ Dienste

3.1 Einteilung der verschiedenen Techniken

Eine grundsätzliche Entscheidung bei der Realisierung von Diensten zur Auffindung von speziellen Dienstleistungen liegt darin, ob die Daten über die verfügbaren Dienstleister erst bei der Dienstanforderung (Reaktiv) oder vorab durch ein separates Verfahren (Proaktiv) erfaßt werden.

Wird ein reaktives Verfahren benutzt, so ist keine genaue Analyse der Netzwerktopologie oder des Netzwerkzustands möglich, da diese aufgrund ihres hohen Aufwandes die eigentliche Dienstanforderung zu lange verzögern würden. Daher ist bei diesen Verfahren meist keine genaue Bestimmung des bestmöglichen Servers durchführbar. Bei proaktiven Techniken ist dagegen die Bestimmung eines optimalen Dienstansbieters in akzeptabler Zeit erreichbar. Jedoch erfordern diese Verfahren einen höheren Aufwand, da Datentabellen

über die vorhandenen Server erstellt und bei verändertem Netzwerkzustand aktualisiert werden müssen.

Die verschiedenen Mechanismen lassen sich anhand der folgenden Merkmale genauer klassifizieren:

- Platzierung der Protokollunterstützung in der Vermittlungsschicht oder in der Anwendungsschicht
- mit Unterstützung durch vorhandene Routingtabellen oder ohne eine solche Unterstützung
- analysieren der Netzwerktopologie oder Messung der Entfernung zu verschiedenen Rechnern

Die möglichen Verfahren sind in Bild 55 aufgeführt und nach den obigen Kriterien eingeteilt [GS95].

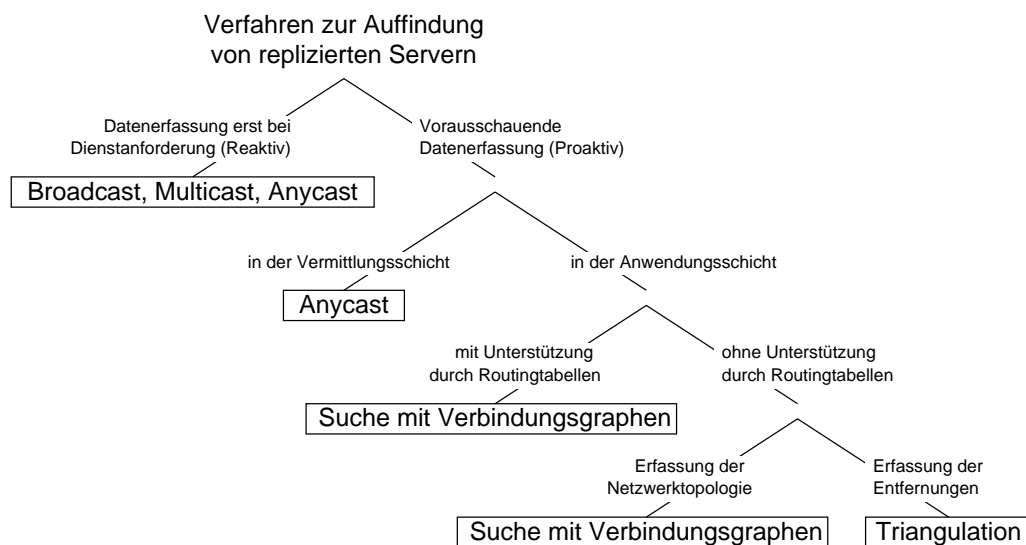


Abbildung 55. Einteilung der Techniken zur Auffindung von Servern

Auf die Verfahren des Broadcast und Multicast wird in diesem Text nicht weiter eingegangen. Aufgrund ihrer unkomplizierten Handhabung stellen sie zwar eine sehr einfache Lösung des Auswahlproblems dar, jedoch erscheint die Vorgehensweise dieser Verfahren nicht sehr effizient, da das Netzwerk mit vielen unnötigen Anforderungsnachrichten überschwemmt wird. Trotz dieser offensichtlichen Nachteile sind Broadcast und Multicast aufgrund ihrer Einfachheit die momentan am meisten eingesetzten Techniken.

3.2 Anycast

Allgemeines Prinzip des Anycastdienstes: Anycast bezeichnet eine zustandslose Datagrammübertragung, welche von Rechnern hauptsächlich dazu benutzt wird, um eine bestimmte Dienstleistung in einem Netzwerk ausfindig zu machen. Das Ziel von Anycast

besteht darin, die Anforderung eines Benutzers an eine Resource „Anycastadresse“ an einen einzelnen, vorzugsweise naheliegenden Server zu übertragen. Im Gegensatz zum Broad- und Multicast wird beim Anycast nur eine Anforderung an genau einen Server gestellt. Ein Anwendungsprogramm, welches einen Anycastdienst benutzt, braucht sich also nicht mehr um die Auswahl eines Dienstanbieters zu kümmern, sondern gibt nur noch die Adresse eines gewünschten Typs von Dienstbringer an. Die Vermittlungsschicht stellt daraufhin die Weiterleitung dieser Anforderung an den unter den aktuellen Netzwerkeigenschaften und der vorgegebenen Distanzmetrik bestmöglichen Server sicher.

Um eine Vermittlung an einen geeigneten Server sicherstellen zu können, muß das zuständige Anycastprotokoll natürlich die Anbieter eines Dienstes sowie deren Distanz aufgrund einer vorgegebenen Metrik kennen. Eine einfache Lösung dieses Problems stellt die Benutzung vorhandener Routingtabellen dar. Der Anycastdienst braucht bei dieser reaktiven Vorgehensweise keine eigenen Datentabellen aufzubauen. Da jedoch in den meisten Fällen nicht alle in Frage kommende Server durch Routingtabellen erfasst sind, müssen diese um die Informationen für den Anycastdienst ergänzt werden. Diese Erweiterungen können nach dem folgenden Prinzip durchgeführt werden: Ein Server, der einen replizierten Dienst neu anbietet, teilt seinen benachbarten Routern die Existenz dieser Dienstleistung über die normalen Routingnachrichten mit. Die so benachrichtigten Router bestimmen daraufhin die Entfernung zu dem neuen Anycastserver mit der vorgegebenen Distanzmetrik und aktualisieren ihre Routingtabellen entsprechend. Im nächsten Schritt geben diese Router die neuen Informationen an die ihnen benachbarten Router weiter, die ihrerseits wieder ihre Routingtabellen erneuern. Mit dieser Technik würde ein neuer Anycastserver im gesamten Netzwerk bekanntgegeben.

Der Nachteil dieser Technik ist natürlich der immense Speicherbedarf um alle Server eines Netzwerkes in jedem Router zu erfassen. Um andererseits eine gewisse Flexibilität zu gewährleisten, müssen einem Router möglichst viele Anbieter bekannt sein. Es muß also eine sinnvolle Entscheidung über die Anzahl der zu speichernden Server getroffen werden. Daher ist nicht unbedingt sichergestellt, daß auch wirklich der naheliegende Server gefunden wird.

Ein besserer (proaktiver) Ansatz zur Auffindung des bestmöglichen Servers im Anycastdienst besteht in der Verwendung von sogenannten Core-Based-Trees (CBT). Diese Technik wird weiter unten genauer beschrieben.

Probleme bei Anycasting im Internet: Bei der Einbeziehung eines Anycastdienstes in das bestehende Internet sind einige Entscheidungen und Probleme zu lösen die sich durch die vorhandenen Implementierungen der TCP/IP-Protokolle ergeben [PMM93].

Grundsätzlich können IP-Nachrichten dupliziert oder falsch vermittelt werden. Daher ist es nicht ausgeschlossen, daß eine Anycastnachricht an mehrere Server gesendet wird, oder auch verlorenght. Eine Definition des Anycastdienstes im Internet sollte also lauten, daß ein IP-Anycast eine bestmögliche Vermittlung an einen, aber möglicherweise mehrere Server bietet.

Ein weiteres wesentliches Problem besteht in der Verbindungslosigkeit des IP-Dienstes, aufgrund der es dem Routingprotokoll des Clients nicht möglich ist, sich zu merken,

an welchen Server die letzte Anycastnachricht gesendet wurde. Werden von einem Anwendungsprogramm nun zwei aufeinanderfolgende Nachrichten an eine Anycastadresse verschickt, so kann also nicht vorausgesetzt werden, daß diese Nachrichten von IP auch an ein und denselben Server gesendet werden. Die Lösung dieses Problems besteht darin, daß sich Anwendungen, die mehrere Nachrichten an den gleichen Server senden möchten, die tatsächliche Adresse des Empfängers nach der ersten Anforderung speichern. Alle weiteren Nachrichten können dann als Unicast an die konkrete Adresse der Gegenstelle gesendet werden, und müssen nicht mehr den Anycastdienst von IP benutzen.

Anycastadressen im Internet

Möglichkeiten zur Einbindung von Anycastadressen in IPv4: Eine weitere Entscheidung bei der Einbindung eines Anycastdienstes in die TCP/IP-Protokolle ist die Frage nach der zu verwendenden Klasse von IP-Adressen. Hier stehen zwei Möglichkeiten zur Auswahl: Einbeziehung der Anycastadressen in eine der bestehenden Adressklassen oder Definition einer neuen Klasse.

Der große Vorteil bei der Benutzung eines vorhandenen Adressraumes ist die Einfachheit der Router, da diese nicht zwischen Unicast- und Anycastnachrichten unterscheiden müssten. Dem stehen zwei wesentliche Nachteile dieses Ansatzes gegenüber: Es gibt praktisch keine Möglichkeit um eine Anycastadresse als solche zu identifizieren, und es ist sehr schwierig Anycastadressen bereitzustellen, die im gesamten Internet öffentlich bekannt sind¹⁸. Diese beiden Nachteile wären in einer neuen Klasse von Internetadressen nicht vorhanden.

Die Möglichkeit, einen Teil des vorhandenen Adressraumes als Anycastadressen zu deklarieren, erscheint wenig sinnvoll, da mit dieser Technik der Adressraum sehr ineffizient genutzt würde und somit auch eine starke Einschränkung hinsichtlich der Anzahl von Anycastadressen bestünde.

Anycastadressen in IPv6: Die Spezifikation der IPv6-Adressarchitektur [HD95], welche eine Erweiterung der bestehenden IPv4-Architektur darstellt, definiert eine Adresse als einen 128-Bit langen Bezeichner eines einzelnen Interfaces oder einer Menge von Interfaces. Alle Unicastadresse werden in IPv6 anhand der ersten drei Bits der Adresse von allen anderen Adressierungsarten unterschieden. Es stehen somit noch 125 Bits zur Kennzeichnung eines Interfaces mit einer Unicastadresse zur Verfügung.

In IPv6 werden Anycastadressen in den Unicastadressenraum eingelagert, d.h. Anycastadressen sind syntaktisch nicht von Unicastadressen zu unterscheiden. Eine Anycastadresse ist in IPv6 als ein Bezeichner definiert, welcher mehreren Interfaces zugeordnet ist.

Zur Unterscheidung zwischen Anycast und Unicast werden spezielle Einträge in den Routingtabellen benutzt, die nach dem folgenden Prinzip erstellt werden: Für jede zugewiesene Anycastadresse gibt es einen längsten Adresspräfix P , der die topologische Region bezeichnet, in der sich alle Interfaces dieser Anycastadresse befinden. Innerhalb

¹⁸ Solche Adressen wären z.B. zur Auffindung des nächstgelegenen DNS-Servers wünschenswert.

der von P bezeichneten Region muß jedes Mitglied des Anycasts durch einen separaten Eintrag im Routingsystem gekennzeichnet sein. Außerhalb dieser Region wird die Anycastadresse im Routingprotokoll wie jede andere Adresse aus P behandelt. Mit dieser Vorgehensweise wird das Problem der „großen“ Routingtabellen umgangen.

Im schlechtesten Fall kann das Präfix P einer Anycastadresse das Nullpräfix sein, z.B. falls die Mitglieder dieser Gruppe keiner zusammenhängenden Region angehören. In diesem Fall muß die Adresse allen Routern im gesamten Internet explizit bekannt sein. Es ist daher anzunehmen, daß solche globalen Anycastadressen in IPv6 nicht oder nur in sehr begrenztem Umfang unterstützt werden.

Beispiel für Anycastadressen im bestehenden IP: Eine beispielhafte Implementierung von Anycast auf IP besteht darin, daß eine neue Familie von Socketadressen definiert wird, die ein Duplikat der IP-Familie darstellt [Mil95]. Diese „Virtual-Internet-Family“ erlaubt die Erzeugung eines virtuellen Netzinterfaces, welches an eine Anycastadresse gebunden ist. Da diese virtuellen Sockets wie Standardinterfaces erscheinen, können sie auch als solche von den üblichen Routingprotokollen behandelt werden. Die mehrfachen Wege zu einer Anycastadressen können vom IP-Routing so gehandhabt werden, als ob es sich hierbei um mehrfache Verbindungen innerhalb des Netzwerkes handelt.

Eine solche Realisierung des Anycasting ist mit vorhandenen Routingprotokollen kompatibel, und bedarf nur einer geringfügigen Anpassung der Netzwerksoftware der Rechner dahingehend, daß die Anycastadressen von diesen erkannt werden können.

Routing von Anycastnachrichten mittels CBTs

Einführung in die CBT-Architektur: Eine Möglichkeit zur Vermittlung von Anycastadressen besteht in der Verwendung von sog. Core-Based-Trees (CBT) [Bal95]. Ein CBT ist eine Datenstruktur, welche die Topologie mehrerer Router des Internets in der Form eines Baumes repräsentiert, in dem die Wege zur Vermittlung einer Nachricht durch die Äste des Baumes vorgegeben sind. In diesem Abschnitt wird zuerst der allgemeine Aufbau eines solchen CBTs besprochen, bevor im nächsten Abschnitt der Einsatz dieser Bäume in der Anycasttechnik vorgestellt wird.

Ein CBT besteht aus einem Netzwerkrouter, dem „Core-Router“¹⁹, welcher die Wurzel des Baumes darstellt, und von dem ein oder mehrere Verzweigungen ausgehen. Diese Zweige werden aus weiteren Routern, den „Non-Core-Routern“, gebildet, welche einen kürzesten Weg zwischen den Hosts, die direkt an einem der Router hängen, und dem Core-Router bilden. Ein Beispiel für einen CBT ist in Bild 56 gezeigt, in dem die Kreise die Router und die Sechsecke die Hosts darstellen.

Für jede Gruppe von Hosts (z.B. alle Server einer speziellen Anycastadresse, alle Empfänger einer Multicastnachricht, etc.) wird ein eigener CBT aufgebaut, der all diese Hosts miteinander verbindet. In diesem Baum können sich dynamisch Hosts anmelden oder abmelden. Mit dieser Datenstruktur wird eine Übertragung von Nachrichten nur über solche Verbindungen erreicht, die direkt zu den entsprechenden Empfängern führen. Es

¹⁹ „Core“ = Kern

ergibt sich damit der Vorteil, daß Router, die nicht im CBT liegen (die sich also auch nicht auf direktem Weg zwischen Sender und Empfänger befinden), keine unnötigen Pakete vermitteln müssen.

Die CBT Architektur wurde hauptsächlich zur Vermittlung von Multicastnachrichten entwickelt. Ist für jede Multicastgruppe ein CBT vorhanden, so kann eine Nachricht an diese Gruppe nach der folgenden Vorgehensweise übertragen werden: Die Mitteilung wird zunächst als Unicastnachricht an den Core des zugehörigen CBTs gesendet. Auf dem Weg dorthin wird diese Nachricht von irgendeinem Router des Baumes, entweder vom Core selbst, oder von einem Non-Core-Router, empfangen. Von diesem Router ausgehend wird die Nachricht nun rekursiv an alle benachbarten Router weitergeleitet, bis die Rechner an den Blättern, also die Empfänger der Multicastnachricht, erreicht sind. Auf diese Weise erreicht eine Multicastnachricht alle Hosts der entsprechenden Gruppe.

Einsatz von CBTs im Anycast: Die CBT Architektur bietet neben der Vermittlung von Multicastnachrichten nun auch eine effektive Lösung des Anycastproblems, ohne daß eine neue Klasse von IP-Adressen definiert werden muß.

Dazu wird für jede Anycastadresse ein eigenständiger CBT aufgebaut, in den alle Server einer speziellen Dienstleistung aufgenommen werden. Die Übertragung einer Dienstanforderung an eine Anycastadresse erfolgt nun analog zur Übertragung einer Multicastnachricht in zwei Schritten: Ein Client, der natürlich nicht in den CBT dieser Anycastadresse integriert sein muß, sendet eine Unicastnachricht an den entsprechenden Core-Router. Diese Nachricht trifft nun entweder am adressierten Core oder an einem der Non-Core-Router, welcher auf dem Weg zwischen dem Client und dem Core liegt, auf den CBT. Im zweiten Schritt der Übertragung wird die Nachricht nun von dem Router nicht wie beim Multicast an alle benachbarten Router übertragen, sondern nur direkt an den nächstgelegenen Server der Anycastadresse weitergeleitet. Diese Technik wird auch als das „2-Phasen Routing“ des CBTs bezeichnet. Ein Beispiel für eine solche Übertragung ist in Bild 56 dargestellt.

Dieses Verfahren hat trotz seiner offensichtlichen Effektivität zwei große Probleme, die hauptsächlich den Core-Router betreffen: Es besteht erstens das Problem, einen geeigneten Router als Core auszuwählen, und zweites ist die Adresse dieses ausgesuchten Routers den Clients bekanntzugeben, damit diese eine Anycastanforderung absetzen können. Da es von der Auswahl des Core-Routers abhängt, ob eine Anforderung wirklich den nächstgelegenen Server findet, und da die Netzwerktopologie im allgemeinen nicht bekannt sein dürfte, ist es auch einsichtig, daß diese Methode nicht dafür garantieren kann, den wirklich nächsten Server anzusprechen.

3.3 Serverauswahl durch Verbindungsgraphen

Allgemeine Vorgehensweise: Bei der Auswahl eines Servers durch einen Verbindungsgraphen wird ein möglichst großes Teilgebiet des gesamten Netzwerkes auf einen Graph abgebildet, in dem dann auf Anfrage eines Clients ein nahegelegener Dienstanbieter gesucht wird. Die Konstruktion und ein späteres Updating dieses Netzwerkgraphen wird

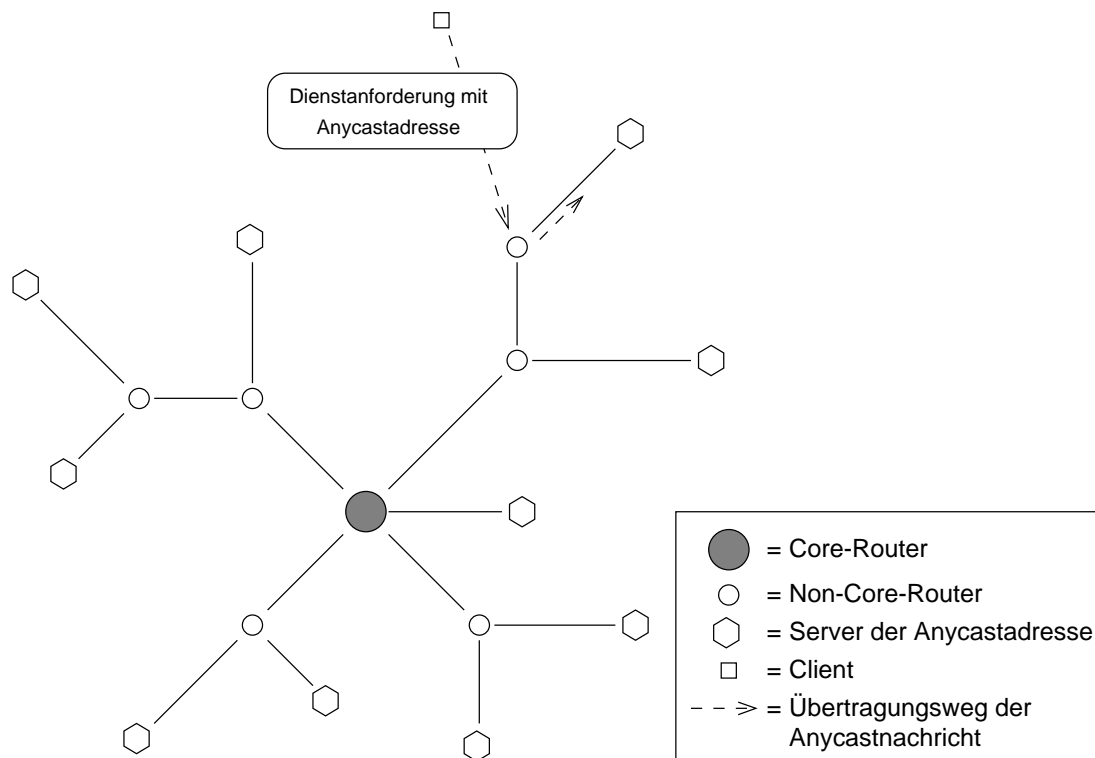


Abbildung 56. CBT Architektur in der Anycasttechnik

von einem oder mehreren ausgewählten Rechnern innerhalb des Netzwerkes, den sog. „Measurement-Servern“, durchgeführt.

Die allgemeine Vorgehensweise dieser Technik besteht darin, daß zuerst ein Graph erstellt wird, der die Verbindungen der Measurement-Server zu den replizierten Dienstern widerspiegelt. Dabei stellen die einzelnen Rechner des Netzwerkes die Knoten des Verbindungsgraphen dar, die Verbindungen der Rechner untereinander bilden die Kanten des Graphen. Der Aufbau des Verbindungsgraphen ist davon abhängig, ob vorhandene Routingtabellen der Vermittlungsschicht benutzt werden können, oder ob die Netzwerktopologie selbst untersucht werden muß. Konkrete Verfahren hierzu werden im nächsten Abschnitt besprochen.

Fragt nun ein Client einen der Measurement-Server nach einem für ihn naheliegenden Server eines replizierten Dienstes, so erstellt der Measurement-Server einen weiteren Graph, der den Weg zwischen ihm und dem Client darstellt, und fügt diesen Graph an den bestehenden Verbindungsgraph an. Die Suche des nächstgelegenen Servers kann danach mit einer einfachen Breitensuche, beginnend beim Knoten des anfragenden Clients, erfolgen.

Eine entscheidende Frage bezüglich der Effektivität dieser Technik liegt darin, wieviele Measurement-Server benutzt werden um den initialen Verbindungsgraphen aufzubauen, und wieviele davon vom Client angefragt werden müssen um einen möglichst nahen Dienstbringer zu finden. Andererseits dürfen nicht zu viele Measurement-Server eingesetzt werden, da der Aufwand zur Konstruktion des Graphen sonst zu aufwendig und das gesamte Verfahren dadurch ineffizient wird.

Aufbau des Verbindungsgraphen

Aufbau des Verbindungsgraphen mit Routingtabellen: Sind zur Konstruktion des Verbindungsgraphen vorhanden Routingtabellen verfügbar, so ist diese recht einfach durchführbar. Ausgehend von einem Measurement-Server kann der Weg zu einem der Server festgestellt werden, indem zuerst die eigene Routingtabelle betrachtet wird und die erste Teilstrecke des Weges ermittelt wird. Bei dem so erreichten Router wird nun dessen Routingtabelle betrachtet um die nächste Teilstrecke herauszufinden. Dieses Verfahren wird solange wiederholt, bis der gewünschte Server erreicht ist. Dieser Algorithmus kann dahingehend erweitert werden, daß die Wege zu allen Servern von einem der Measurement-Server berechnet wird. Bei diesem Verfahren ist darauf zu achten, daß jede Routingtabelle nur einmal betrachtet wird, da ansonsten manche Wege doppelt aufgenommen würden.

Aufbau des Verbindungsgraphen ohne Routingtabellen: Sind keine Routingtabellen verfügbar, z.B. weil ein Zugriff auf diese nicht möglich ist, so ist eine aufwendige Untersuchung der Netzwerktopologie nötig. Dieses als „Expanded Ring Search“ bezeichnete Verfahren arbeitet ähnlich wie der Algorithmus zur Ermittlung der Hop-Entfernung aus Abschnitt 2.1 mit variierenden TTL-Werten. Zur Bestimmung eines Weges zwischen einem der Measurement-Server und einem Server wird im ersten Schritt eine Nachricht an diesen Server mit dem TTL-Wert 1 gesendet und derjenige Router vermerkt, welcher mit einer ICMP-„Time-Exceeded“ Nachricht antwortet²⁰. Danach wird der TTL-Wert um eins erhöht und die Nachricht wieder abgesendet. Dies wird solange fortgesetzt, bis der Empfänger erreicht ist. Die so entstandene Reihenfolge von antwortenden Rechnern bildet dann den Weg zu dem gewünschten Server. Auf diese Weise können die Wege zu allen Servern ermittelt werden.

Der große Nachteil dieser Technik besteht in dem immensen Nachrichtenaustausch, welcher benötigt wird um die Wege zu allen möglichen Servern zu erfassen. Hieraus ergibt sich natürlich auch ein nicht unerheblicher Zeitaufwand zum Aufbau des gesamten Verbindungsgraphen. Im Gegensatz zur Datenerfassung mit Routingtabellen ist es daher kaum möglich alle Dienstanbieter des Netzwerkes in den Graphen aufzunehmen.

3.4 Serverauswahl durch Triangulation

Eine weitere Methode zur Auffindung nahegelegener Server besteht in der Vermessung der Positionen replizierter Server durch einer Menge von ausgewählten Rechnern, welche kurz als Vermessungsrechner bezeichnet werden. Dieses Verfahren ist deshalb interessant, weil zur Vermessung der Positionen nur einfache Berechnungen der Entfernungen notwendig sind, welche mit der in Abschnitt 2.1 beschriebenen Technik schnell durchzuführen sind. Der zeitaufwendige Aufbau eines Verbindungsgraphen entfällt komplett. Da diese Methode analog zur Triangulation bei Positionierungssystemen funktioniert wird sie auch als Triangulationsmethode bezeichnet.

Die allgemeine Vorgehensweise bei der Triangulationsmethode sieht folgendermaßen aus: Beginnend mit einer Menge von replizierten Servern und N Rechnern zur Vermessung

²⁰ ICMP = „Internet Control Message Protocol“

der Netzwerktopologie, welche über das gesamte Netzwerk verteilt sein sollten, berechnet zuerst jeder dieser Vermessungsrechner seine Distanz zu jedem der Server. All diese Messungen werden dann zu einem Triangulationsserver gesendet, welcher die einzelnen Entfernungen zu den replizierten Servern als Koordinaten eines N -dimensionalen Vektorraumes, dem Triangulationsraum, wie in Abschnitt 2.1 beschrieben zusammenfaßt. In der Praxis wäre es durchaus möglich, daß die Koordinaten von den Vermessungsrechnern selbst ermittelt werden, indem sie ihre Meßwerte untereinander austauschen. Ein Beispiel für die Berechnung solcher Koordinaten zeigt Bild 57 zusammen mit Tabelle 5.

Fragt nun ein Client nach einem ihm nahegelegenen Server, so ermittelt jeder der Vermessungsrechner seine Distanz zu dem Client und der Triangulationsserver berechnet hieraus die Koordinaten des Clients. Aufgrund der Koordinaten des Clients und denen der replizierten Server erstellt der Triangulationsserver eine Liste derjenigen Server, die sich in geringer Entfernung zum Client befinden. Bei dieser Berechnung der Abstände wird die *AVG* Metrik aus Abschnitt 2.1 benutzt.

Ein großer Nachteil dieser Technik ist es, daß der Koordinatenraum der Triangulation von den Standorten der Vermessungsrechnern abhängig ist. Ein Beispiel zur Erläuterung dieses Problems ist in Bild 57 dargestellt, die zugehörigen Abstände und Koordinaten zeigt Tabelle 5. Die Koordinaten vermitteln den Eindruck als ob S_1 und S_2 den gleichen Standort hätten, tatsächlich liegen sie aber 8 Hops von einander entfernt. Ebenso sind C_1 und C_2 6 Hops entfernt, besitzen aber auch die selben Koordinaten. Da kein Vermessungsrechner auf der „horizontalen Achse“ liegt, ist es nicht möglich zwischen den Standorten von S_1 und S_2 zu unterscheiden.

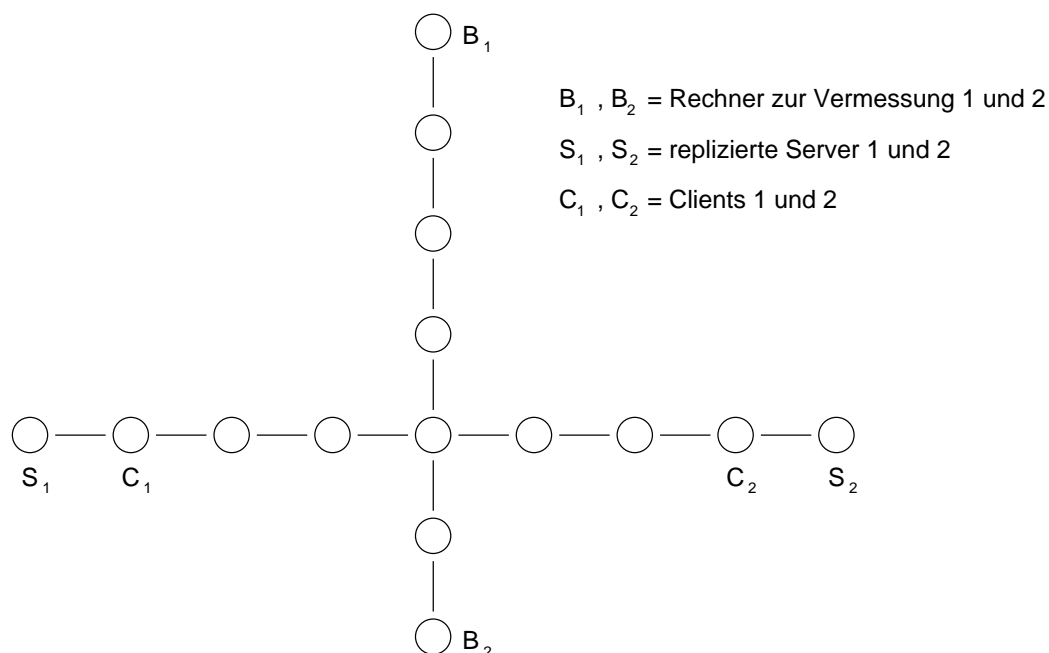


Abbildung 57. Beispiel zur Triangulation

Dieses Beispiel zeigt, daß mit der Triangulationstechnik keine eindeutige Entscheidung über den Standort eines Rechners getroffen werden kann, es ist nur möglich obere und

Rechner	Entfernung zu B_1	Entfernung zu B_2	Koordinaten
S_1	8	6	$\langle 8, 6 \rangle$
S_2	8	6	$\langle 8, 6 \rangle$
C_1	7	5	$\langle 7, 5 \rangle$
C_2	7	5	$\langle 7, 5 \rangle$

Tabelle 5. Entfernungen und Koordinaten zu Bild 57

untere Grenzen für den Abstand zweier Punkte (Rechner) innerhalb des Triangulationsraumes anzugeben. Diese Grenzen können als erster Schritt zur Vorauswahl mehrerer Server verwendet werden, unter denen ein Client dann durch eine genaue Analyse der Abstände einen optimalen Server finden kann.

3.5 Vergleich der verschiedenen Techniken

Vergleicht man die vorgestellten Techniken hinsichtlich ihres Nachrichtenaustausches, welcher bei der Datenerfassung und der Auswahl eines Dienstanbieters anfällt, so stellt der Anycastdienst die beste Lösung dar. Dieser Vorteil entsteht dadurch, daß sämtliche Daten über das bestehende Netzwerk in der Vermittlungsschicht vorhanden sind, und nicht erst mit aufwendigen Untersuchungen erfaßt werden müssen. Auf der Vermittlungsschicht ist es auch möglich eine Datenbank über vorhandene Server inkrementell durch Überwachen der Routingnachrichten zu verwalten. Genau dies stellt den Hauptnachteil aller Verfahren auf der Anwendungsschicht dar: alle Daten über das vorhandene Netzwerk müssen periodisch neu erstellt werden, was eine komplette Untersuchung des Netzwerkes zur Folge hat.

Auf der anderen Seite kann ein Anycastdienst je nach verwendeter Technik nicht immer garantieren, daß wirklich der nächstgelegene Server gefunden wird, was mit einem Verbindungsgraphen der das gesamte Netzwerk abdeckt ohne weiteres möglich ist. Auch können bei Diensten in der Vermittlungsschicht nur Eigenschaften berücksichtigt werden, die sich aus der Netzwerktopologie ergeben, die Qualität eines Servers kann also z.B. nicht in die Auswahlkriterien miteinbezogen werden.

Bei genauerer Untersuchung der Kosten für die Verfahren basierend auf Verbindungsgraphen stellt sich heraus, daß beide Konstruktionstechniken etwa denselben Aufwand zum Aufbau und zur Verwaltung des Graphen verursachen. Da jedoch bei der Datenerfassung ohne Routingtabellen nur ein Teil des gesamten Netzwerkes betrachtet werden kann, ist es nicht sichergestellt, daß der bestmögliche Server gefunden wird.

Das Triangulationsverfahren erscheint aufgrund der Tatsache, daß nur einfachen Abstandsberechnungen zur Vermessung des Netzwerkes notwendig sind zunächst als eine sehr kostengünstige Technik. Da sich die Auffindung von naheliegenden Dienst Anbietern jedoch als sehr ineffizient herausstellt, müssen für ein brauchbares Ergebnis viele Vermessungspunkte vorhanden sein. Dies ergibt zusammen einen sehr hohen Verwaltungsaufwand, bei einer Effektivität die noch unter der der Graphensuche liegt.

Verfahren	Nachrichtenaufkommen	Effektivität
Anycast	gering	schlecht vorhersagbar weil vom Verfahren abhängig
Graphensuche m. Routingtabellen	hoch	sehr gut
Graphensuche o. Routingtabellen	hoch	ausreichend
Triangulation	sehr hoch	schlecht

Tabelle 6. Zusammenfassender Vergleich der vorgestellten Techniken

Tabelle 6 zeigt für die vorgestellten Methoden einen Überblick über das bei der Datenerfassung jeweils entstehende Nachrichtenaufkommen. Zur Bewertung der einzelnen Techniken ist zusätzlich deren Effizienz bei der Auffindung des bestmöglichen Servers in Bezug auf den von ihnen verursachten Nachrichtenaustausch eingetragen.

In Anbetracht des sich ständig verändernden und größer werdenden Internets werden sich zukünftige Forschungen sicher zunehmend mit Techniken befassen, die ihre Datenbestände inkrementell erneuern können (wie z.B. beim Anycastdienst), und somit eine große Kostenersparnis gegenüber Verfahren haben, die ihre Daten auf der Anwendungsschicht verwalten. Ein weiteres Forschungsgebiet wäre sicherlich die Erweiterung der besprochenen Distanzmetriken dahingehend, daß komplexere Netzwerkeigenschaften, wie etwa die Übertragungskosten, mitberücksichtigt werden können.

Das ATM-Management-Modell

Michael Ebner

Kurzfassung

Der asynchrone Transfermodus, kurz ATM, bildet die Grundlage für das kommende Breitband-ISDN. Dabei bietet ATM eine vielseitige, multifunktionale Plattform an, welche verschiedene Verkehrstypen (Telefon, herkömmliche Daten oder Multi-Media-Daten) im gleichen Netzwerk übertragen kann und schwer mit traditionellen Lösungen vereinbar ist. Dies alles macht ATM sehr populär, stellt aber auch neue Anforderungen an das Netzwerkmanagement.

Das ATM-Forum hat dazu ein Fünf-Schnittstellen-ATM-Managementmodell definiert, welches hier genauer vorgestellt wird. Dazu werden erstmal die Anforderungen an das Management nach ISO/OSI vorgestellt. Dann gibt es einen kleinen Überblick über das ATM-Managementmodell. Im Hauptteil werden dann die einzelnen Schnittstellen vorgestellt.

1 Einführung

Der asynchrone Transfermodus, kurz ATM, bildet die Grundlage für das kommende Breitband-ISDN. Dabei verbindet ATM das Beste aus TDM (Time Division Multiplex) und Paketübertragung und bietet eine vielseitige, multifunktionale Plattform an, welche verschiedene Verkehrstypen (Telefon, herkömmliche Daten oder Multi-Media-Daten) im gleichen Netzwerk übertragen kann. Zusätzlich bietet ATM eine Skalierbarkeit an, welche schwer mit traditionellen Lösungen vereinbar ist. Dies alles macht ATM sehr populär.

Zum Beispiel wollen die privaten Nutzer die Vorteile des öffentlichen Netzes nutzen, aber die Kontrolle, zur Garantie der Dienste, behalten. Der Betreiber von öffentlichen Netzwerken will einen Einblick in das private Netzwerk erhalten, mit Einschränkungen natürlich, um dem Benutzer Managementfunktionen anbieten zu können. Wenn im weiteren Verlauf von Betreibern und Benutzern die Rede ist, dann sind mit Betreibern immer Personen von öffentlichen Netzwerken und mit Benutzern immer Personen von privaten Netzwerken gemeint.

Dies alles erfordert einen neuen Rahmen für ein ATM-Management, welcher mit neuen Anforderungen an die Managementfunktionen fertig werden muß. Ohne diesen neuen Standard können Netzmanager auch nicht mit der Entwicklung von Managementfunktionen für ATM beginnen.

Netzmanager benötigen eine Ende-zu-Ende Sicht für eine Anzahl von Betriebsbedingungen. Sie müssen auch in der Lage sein, die Infrastruktur des physikalischen Netzes überwachen und kontrollieren zu können. SNMP-basierende Systeme sind nicht in der Lage, dieses *umfassende Ende-zu-Ende Management* von ATM-Netzwerken zu handhaben, weil die dazu notwendigen Protokollfähigkeiten nicht vorhanden sind. Da SNMP (Simple Network Management Protocol) aber bei den Endverbrauchern weit verbreitet ist, wird es als Ausgangspunkt für das ATM-Managementsystem benutzt, d.h. es wird auf SNMP aufgesetzt.

Das ATM-Forum entwickelt deshalb ein ATM-Management-Modell. Dieses Forum ist ein weltweiter, nicht kommerzieller Zusammenschluß von Firmen, Institutionen, Benutzern usw. (mehr als 700 Mitglieder), welche an der Verbreitung, Weiterentwicklung und Standardisierung von ATM (Asynchronous Transfer Mode) interessiert sind. Dabei wird mit anderen Organisationen, wie ANSI (American National Standard Institute) und ITU-T (International Telecommunications Union), zusammengearbeitet.

2 ISO/OSI-Netzwerkmanagementmodell

In diesem Kapitel wird kurz das Netzwerkmanagementmodell nach ISO/OSI beschrieben [HA93].

Netzwerkmanagement läßt sich nach OSI umschreiben als “Aktivitäten zur **Kontrolle, Koordination und Überwachung** aller Ressourcen, welche eine Kommunikation in der OSI-Umgebung ermöglichen”.

Das Modell legt grundlegende Konzepte und Begriffe fest, wobei von einem kommunikationsorientierten Modell ausgegangen wird. Es wird keine Implementierung vorgegeben. Die Architektur ist in vier Teilmodelle

- Informationsmodell
- Organisationsmodell
- Kommunikationsmodell
- Funktionsmodell

aufgeteilt.

Im folgenden werden die vier Teilmodelle vorgestellt.

2.1 OSI-Informationsmodell

Da sich alle Managementaktivitäten mit der Erfassung, Verwaltung, Steuerung und Überwachung von unterschiedlichen *Kommunikations-Ressourcen* befassen, ist eine standardisierte Beschreibung von Managementinformationen notwendig. Das Informationsmodell hat genau diese Aufgabe.

Im Informationsmodell werden alle managementrelevanten Ressourcen, mit der Hilfe eines objektorientierten Ansatzes, abstrahiert. Somit werden Datenabstraktionen und Vererbungen, sowie Methodenaufrufe verwendet. Dazu werden **Managementobjekte (MOs)** benutzt, welche nach ISO/OSI definiert sind als

die OSI-Managementsicht einer Netzressource in einer OSI-Umgebung, die unter Zuhilfenahme eines OSI-Managementprotokolls verwaltet werden kann [Sei94] .

Ein Managementobjekt (MO) (Bild 58) wird dabei durch folgendes festgelegt :

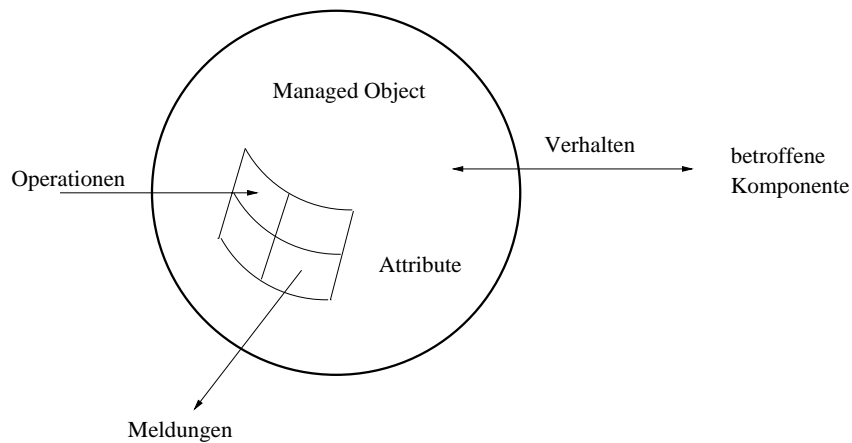


Abbildung 58. Modell eines Managementobjekts

- *Attribute*, welche an seiner Schnittstelle sichtbar sind
- *Management-Operationen*, die auf das Objekt anwendbar sind
- *Verhalten* des Objekts in Reaktion auf eine Management-Operation
- *Meldungen*, die vom Objekt zu erwarten sind
- eingeschlossene *Bedingungen*, welche nach außen unsichtbar sind
- Position in der *Vererbungshierarchie*
- Spezifikation ähnlicher *Objekt-Klassen*

Die *Attribute* beschreiben die allgemeinen Eigenschaften eines MOs. Dazu zählt z.B. die Beschreibung des aktuellen Zustands eines bestimmten Objekts, Identifikation oder Statistiken über das vorherige Verhalten des Objekts.

Management-Operationen können in zwei Typen aufgespalten werden, welche die sich auf die Attribute eines Objektes beziehen, z.B. Setzen von Attributen, und welche die sich auf das ganze Managementobjekt beziehen, wie z.B. Erzeugen eines Objekts.

Bei einem MO können asynchrone Ereignisse auftreten, d.h. Ereignisse, welche durch das zu managende Objekt (z.B. kein Papier mehr bei einem Drucker) auftreten und nicht durch Anforderung des Managers. Um diese Ereignisse an den Manager weiterleiten zu können, werden *Meldungen* benutzt, welche ein MO von sich aus an einen Manager schicken kann, ohne vorher dazu aufgefordert worden zu sein.

Das *Verhalten* eines MOs wird durch eine informelle Beschreibung der Semantik angegeben.

Die MOs eines einzelnen OSI-Systems bilden die *Management Information Base (MIB)*, welche baumähnlich aufgebaut für die Definition und den Gebrauch von Informationen zuständig ist. Für die MIB wurden Operationen und eine abstrakte Syntax/Semantik definiert, aber keine Implementierungsvorgaben.

Das Informationsmodell sagt nicht, was in eine MIB aufgenommen werden muß, sondern nur wie.

2.2 OSI-Organisationsmodell

Grundsätzlich wird von einem verteilten kooperativen Management ausgegangen. Das Organisationsmodell legt Rollen und gegebenenfalls Domänenkonzepte fest. Es werden zwei Rollen unterschieden, nämlich die Manager-Rolle und die Agenten-Rolle. Manager erbringen eine bestimmte Managementaufgabe durch Manipulation von MOs und Agenten verwalten auf Anforderung eines Managers ihnen zugeordnete MOs und leiten deren Meldungen weiter.

Folgende Punkte können zwischen Manager und Agent ausgetauscht werden :

- Operationen auf MOs
- Ergebnisse von Operationen
- Meldungen von MOs
- Fehlermeldungen

MOs können ohne vorherige Aufforderung Meldungen abgeben.

2.3 OSI-Kommunikationsmodell

Grundlage für das Management ist der Austausch von Managementinformationen zwischen den kooperierenden offenen Systemen. Das Kommunikationsmodell regelt diesen Austausch und setzt auf das ISO/OSI-Referenzmodell auf.

Hierzu werden drei verschiedene Managementkategorien

- Schichtübergreifendes Management (Systems Management, SM)
- Schichtenmanagement (Layer Management, LM)
- Protokollmanagement (Layer Operation)

eingeführt.

Die Zusammenarbeit zwischen diesen Kategorien ist nicht festgelegt, auch nicht die Zusammenarbeit dieser Kategorien mit dem lokalen System (z.B. Betriebssystem). Dies ist für eine spätere Implementierung unbefriedigend.

Schichtübergreifendes Management Das schichtübergreifende Management betrifft das Gesamtmanagementverhalten kooperierender Systeme. Dazu werden Managementanwendungen (Systems Management Applications, SMA) benutzt, welche aus der Kooperation von entsprechenden Anwendungsprozessen (System Management Application Processes, SMAP) bestehen. Ein SMAP kann je nach Anwendung die Managerrolle oder die Agentenrolle einnehmen, da Management eine *unsymmetrische* Beziehung ist.

Der kommunikationsrelevante Teil dieser SMAPs ist die Systemmanagement-Instanz (Systems Management Application Entity, SMAE), welche über ein Anwendungsprotokoll

(Systems Management Protocols) Managementinformationen mit anderen SMAEs austauscht. Dafür wurde von OSI eigens ein Dienst entwickelt, nämlich der Common Management Information Service (CMIS), und ein entsprechendes zugehöriges Management-Protokoll, das Common Management Information Protocol (CMIP).

CMIS dient dem Zugriff und der Manipulation von (entfernten) MOs und erlaubt den Zugriff auf die gesamte MIB. Dabei unterscheidet CMIS folgende Dienstgruppen :

- Assoziationsverwaltung
- Ausführen von Operationen
- Mitteilen von Ereignissen

Somit kann mit CMIS/CMIP die MO-Auswahl und Parameterübergabe durchgeführt werden.

CMIS/CMIP setzen auf ACSE (Association Control Service Elements) und ROSE (Remote Operation Service Elements) auf.

Schichtenmanagement Das Schichtenmanagement betrifft Funktionen, Dienste und Protokolle, welche spezifisch für eine Schicht sind und höhere OSI-Schichten nicht benötigen. Bis jetzt wurden nur Protokolle für Routinginformationen und Kataloge für Schicht-3-Objekte und Schicht-4-Objekte festgelegt.

Die Kommunikationsinstanz wird (N)-Layer Management Entity (LME) genannt und das entsprechende Protokoll heißt (N)-Layer Management Protocol.

Protokollmanagement Als mögliche Objekte lassen sich Fenstergrößen, Timer und Testframes in HDLC, Fehlerinformationen und Protokollparameter bei Verbindungsabbau und -aufbau nennen.

2.4 OSI-Funktionsmodell

Managementaktivitäten können in verschiedene Funktionsbereiche (System Management Functional Areas, SMFAs) unterteilt werden.

Im OSI-Funktionsmodell wurden folgende fünf Funktionsbereiche festgelegt :

Konfigurationsmanagement Definieren, Benennen, Initialisieren und Löschen von MOs, Einstellen und Ändern von MO-Attributen, Sichern des Normalbetriebs; usw.

Fehlermanagement Erkennen, Lokalisieren und Beheben von Störungen

Leistungsmanagement Sammeln von statistischen Daten und Aufzeichnungen der Netzhistorie zur Verbesserung des Leistungsverhaltens von Betriebsmitteln.

Abrechnungsmanagement Benutzerverwaltung, Festlegung und Abrechnung von MO-Nutzung und Betriebsmittelverbrauch, Autorisierung

Sicherheitsmanagement Authentisierung, Zugangsüberwachung, Schlüsselverwaltung

Für diese Bereiche wird folgendes festgelegt :

- die erwartete Funktionalität des Bereiches
- Prozeduren, die notwendig sind, um die Bereichsfunktionalität zu erbringen
- Managementobjektklassen, die in diesem Bereich wichtig sind

Die Einteilung in verschiedene Funktionsmodelle ermöglichte es, generische Netzwerkmanagement-Hilfsfunktionen (Systems Management Functions, SMFs) zu entwickeln.

Als Beispiel läßt sich der Event Forwarding Discriminator (EFD) nennen. Mit dem EFD können Meldungen von MOs gefiltert werden, und dann kann über die Weiterleitung entschieden werden.

3 ATM-Managementmodell

Das ATM-Forum, welches für die Standardisierung von ATM verantwortlich ist, entwickelt ein *Fünf-Schnittstellen-ATM-Managementmodell* und ein *OAM (Operationen, Verwaltung und Wartung) Transportwesen*. Das Modell definiert Schnittstellen zum Managen von hybriden Netzwerkumgebungen, welche aus öffentlichen und privaten Netzwerken der Größe LAN bis MAN bestehen können. Außerdem definiert es OAM-Zellen zum automatischen Verteilen von Managementinformationen durch das ATM-Netzwerk.

Da es bis zu einer lauffähigen Vollversion noch einige Jahre dauern wird, wurde ein vorübergehendes Modell geschaffen, die *ILMI (Interim Local Management Interface)*. Die "ILMI" ist ein Bestandteil der "UNI (User-Network Interface)" und arbeitet mit SNMP-Nachrichten, wobei nur die Schnittstelle zwischen zwei Netzen geregelt werden kann, und es können keine Informationen durch das Netz von ihr geschickt werden.

Ziel des Managementmodells ist es ein *umfassendes Ende-zu-Ende ATM-Management* zu ermöglichen. Eine umfassende Architektur erzeugt Probleme und neue Herausforderungen in allen Teilen des Netzwerkmanagements. Die Dienstreservierung, z.B. das Garantieren von Bitraten, ist dafür als wichtigster Punkt zu nennen. Durch das Garantieren können die Verbindungen in jedem Switch optimiert werden.

Weiter läßt sich das Fehler- und Leistungsmanagement aufführen. Die hohe Geschwindigkeit und Unmenge von verschiedenen Datentypen erfordern eine neue Dimension der Überwachung. Desweiteren werden mehrere Sichten auf das Netz gefordert, um das Fehlermanagement zu erleichtern. Dies wird durch eine physikalische und logische Sichtweise auf das Netz ermöglicht.

Auch die Natur der Kostenzuweisung und Abrechnung ändert sich durch eine Ende-zu-Ende Architektur. Die Betreiber eines Netzwerkes müssen für jede Verbindung die verwendete Dienstart erkennen können und dann die Benutzung messen können. Dazu wird ein benutzerorientiertes Zahlungssystem vorgesehen.

Ressourcenverwaltung ist auch ein wichtiger Aspekt, da damit Probleme bzgl. der Kapazität bei bestehenden Netzwerken erkannt werden können und die Planung für neue Dienste und Netzwerke erleichtert wird.

Hierzu entwickelt das ATM-Forum nun ein ATM-Managementmodell, welches private und öffentliche Netzwerkdienste und deren Kommunikation miteinander festlegt. Außerdem werden auch Übergänge zwischen SNMP (bei Kundenseite) und CMIP (bei Betreiberseite), sowie zwischen standardisierten und geschützten (z.B. firmeneigenen) Systemen definiert. Diese Übergänge werden definiert, da zwischen privaten und öffentlichen Netzwerken eine getrennte Sicht mit verschiedenen Stufen der Funktionalität benötigt wird. ITU (International Telecommunication Union), ATM-Forum und "Network Management Forum" (zusammenschluß von Firmen, welche weltweit einen gemeinsamen Netzwerkmanagementstandard verbreiten wollen) arbeiten hierbei zusammen.

Im ATM-Managementmodell wurden fünf Schnittstellen definiert, welche mit M1 bis M5 (M für Management) bezeichnet wurden. Alle sind notwendig, um eine Ende-zu-Ende Überwachung und Kontrolle ermöglichen zu können [AC95, ATM94a, KM95].

Im Bild 59 ist eine Gesamtübersicht abgebildet.

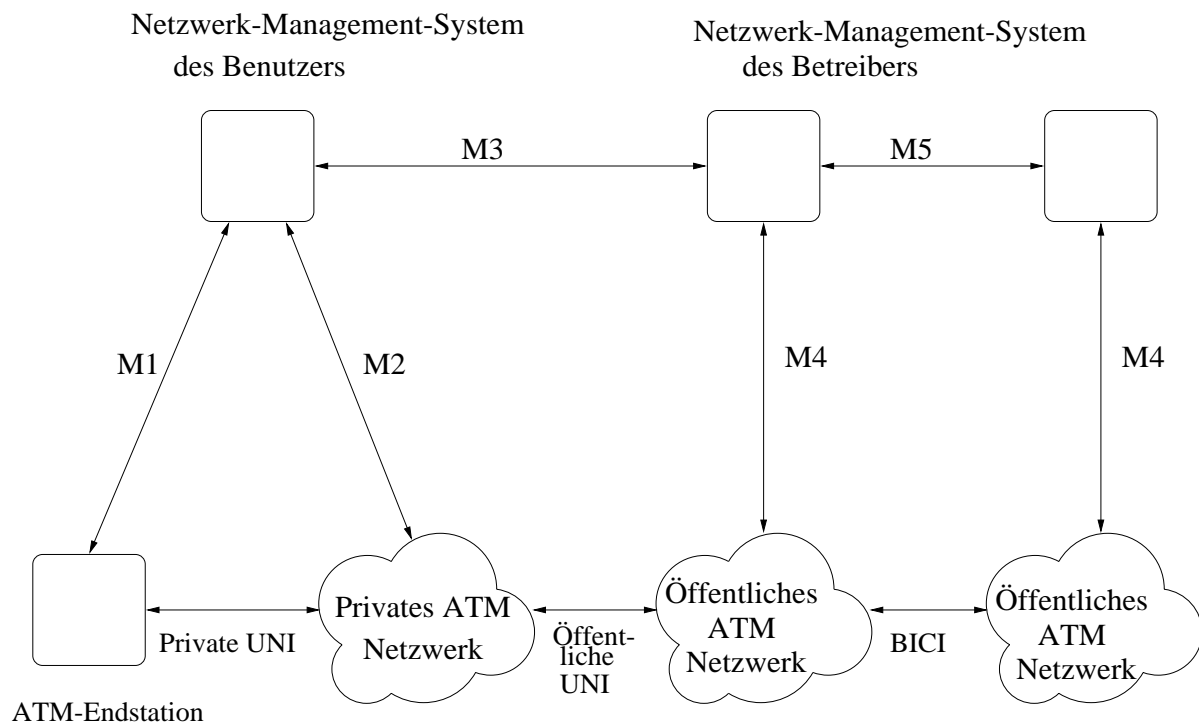


Abbildung 59. Das ATM-Management-Modell

BICI (Broadband Inter Carrier Interface)

Diese fünf Schnittstellen werden nun im nächsten Abschnitt genauer vorgestellt.

4 Die Schnittstellen des ATM-Managementmodells

4.1 M1 und M2

M1 definiert die Schnittstelle zwischen dem Netzwerkmanagementsystem des Benutzers und einer ATM-Endstation. M2 ist die Schnittstelle zwischen dem Netzwerkmanagementsystem des Benutzers und seinem ATM-Netzwerk oder ATM-Switch. Bis jetzt wurde aber keine Spezifikation für diese Schnittstellen herausgegeben.

Deswegen wurde die ILMI als vorübergehendes Modell geschaffen, welche zwischen Teilnetzen (auch Endstationen) ein Minimum an Managementmöglichkeiten zur Verfügung stellt. Die ILMI dient also nicht speziell als Vorläufer für M1 und M2, sondern nur als Managementhilfe, bis die vorgesehenen Schnittstellen (M1-M5) einsatzbereit sind. Es ist zu beachten, daß die ILMI nur das "lokale" Management übernehmen kann, da sie keine Managementinformationen über das Netz verschicken und somit keine Aktionen auslösen kann [AC95].

Da bei den Endverbrauchern SNMP weit verbreitet ist, wurde eine auf SNMP-basierende Spezifikation gewählt. Dies schließt die MIB II und die relevanten MIBs für DS-1, DS-3 (amerikanische Multiplexstandards) und SONET (Synchronous Optical NETwork) ein, sowie die AToM-MIB [AT94a]. Die AToM-MIB versucht, einen Zusammenschluß der obengenannten MIBs zu erzeugen.

Jeder ATM-Switch oder Endsystem, sowie jede UNI, müssen eine "UNI Management Entity (UME)" besitzen, welche eine ILMI-MIB unterstützt. Die UME dient als Ansprechpartner für alle managementrelevanten Anfragen. Die Anfragen werden, vergl. SNMP, in TCP/IP-Pakete gepackt. Diese Pakete werden vor der Übertragung in ATM-Pakete übersetzt und dann über das Netz verschickt.

Die ILMI-MIB sammelt Informationen über folgende Gruppen :

- **physical-layer**
 - Porttyp
 - Mediumtyp
 - Status des Ports
- **ATM-layer**
 - max. Anzahl der VPCs und VCs (Virtual Channel)
 - Anzahl der konfigurierten VPCs und VCs
 - detaillierte Informationen über die Konfiguration von VPCs und VCs
- **ATM-layer statistics**
 - Gesamtzahl der übertragenen, empfangenen und abgelehnten ATM-Zellen
- **Virtual Path Connection (VPC)**
 - Operationsstatus von jedem VPC

- Daten über den Verkehrsdeskriptor und Serviceparameter angegeben vom sendenden und empfangenden Ende des VP
- **Virtual Connection (VC)**
 - siehe VPC
- **network prefix**
 - Sorgt für die eindeutige Adresse eines Netzwerkes, indem ein Präfix der Länge 8-13 Bytes festgelegt wird.
- **address**
 - Sorgt für die eindeutige Adresse einer Endstation, indem eine 8-Bytes große Adresse festgelegt wird

Die angegebenen Informationen bzw. Objekte sind nur ein beispielhafter Auszug.

In der AToM-MIB [AT94a] werden auszugsweise folgende MOs, für die ATM-Konfigurationsinformationen bzgl. einer ATM-Schnittstelle, dazu definiert :

```
AtmInterfaceConfEntry ::= SEQUENCE {
    atmInterfaceMaxVpcs      Integer,
    atmInterfaceMaxVccs      Integer,
    atmInterfaceConfVpcs     Integer,
    atmInterfaceConfVccs     Integer,
    atmInterfaceMaxActiveVpiBits Integer,
    atmInterfaceMaxActiveVciBits Integer,
    atmInterfaceIlmiVpi      Integer,
    atmInterfaceIlmiVcc      Integer,
    atmInterfaceAddressType   Integer,
    atmInterfaceAdminAddress  OCTET STRING,
    atmInterfaceMyNeighbourIpAddress IpAddress,
    atmInterfaceMyNeighbourIfName DisplayString
}
```

```
atmInterfaceMaxVpcs OBJECT-TYPE
    SYNTAX      INTEGER (0 .. 4096)
    MAX-ACCESS  read-write
    STATUS      current
    DESCRIPTION
        "The maximum number of VPCs (PVCs and SVCs)
        supported at this ATM interface. At the ATM UNI,
        the maximum of VPCs (PVCs and SVCs)
        ranges from 0 to 256 only."
    ::= { atmInterfaceConfEntry 1}
```

atmInterfaceIlmiVci OBJECT-TYPE
 SYNTAX INTEGER (0 .. 65535)
 MAX-ACCESS read-write
 STATUS current
 DESCRIPTION
 "The VCI value of the VCC supporting
 the ILMI at this ATM interface. If the values of
 atmInterfaceIlmiVpi and atmInterfaceIlmiVci are
 both equal to zero then the ILMI is not
 supported at this ATM interface."
 DEFVAL { 16 }
 ::= { atmInterfaceConfEntry 8}

atmInterfaceAddressType OBJECT-TYPE
 SYNTAX INTEGER {
 private(1),
 nsapE1 64(2),
 nativeE1 64(3),
 other(4),
 }
 MAX-ACCESS read-only
 STATUS current
 DESCRIPTION
 "The type of primary ATM address configured
 for use at this ATM interface."
 ::= { atmInterfaceConfEntry 9}

4.2 M3

Die Schnittstelle M3 bildet die Verbindung zwischen den öffentlichen und privaten Netzwerken. M3, die Customer Network Management (CNM) Schnittstelle, ist die Schnittstelle zwischen dem Managementsystem des Kunden und dem Managementsystem des Betreibers, welche dem Kunden eine Einsicht in das Netz des Betreibers liefert. Dazu wird hauptsächlich mit dem Konfigurations-, Fehler- und Leistungsmanagement des Betreibers gearbeitet [ATM94b].

Da bei den Benutzern meist SNMP und bei den Betreibern meist CMIP verwendet wird, ist eine Schnittstelle zwischen diesen beiden Systemen vorgesehen.

Einige generelle Anforderungen an die Schnittstelle M3 sind :

- Für die Funktionen der Class I+II (siehe unten) wird SNMP verwendet.
- Die MIB soll hauptsächlich auf dem SNMP-Standard basieren

- Jeder Kunde darf vom öffentlichen Netz nur seine eigene Sicht sehen. Der Betreiber muß dafür sorgen, daß kein Kunde einen Teil eines anderen Kunden sieht.
- M3 muß das Managen von Punkt-zu-Punkt PVCs (Permanent Virtual Connection) und ATM-UNI-Schnittstellen unterstützen.

Es sind zwei Klassen von Diensten für den Kunden vorgesehen, nämlich

- Class I (verbindlich)
 - Grundsätzlich nur Leseberechtigung
 - Anzeige von Konfiguration und Status des Netzwerkes
 - Anzeige von Informationen des Fehler- und Leistungsmanagements
- Class II (optional)
 - Erlaubt Kunden, logische/virtuelle Verbindungen auf-, abzubauen und zu verändern
 - Lese- und Schreibberechtigung

In Planung ist eine Erweiterung der Möglichkeiten, so daß die Kunden eine Realzeitkontrolle über die benutzten Dienste haben.

4.3 M4 und M5

M4 ist die Schnittstelle zum öffentlichen Netzwerk. In M4 sind die Anforderungen zum Managen von ATM-Netzwerkelementen (NEs) festgelegt. In der dazugehörigen MIB wird der Informationsaustausch zwischen den ATM-Geräten und dem Managementsystem festgelegt [ATM94c]. M5 ist die Schnittstelle zwischen den Managementsystemen der Betreiber, also den öffentlichen Netzen. Für M5 gibt es bis zum jetzigen Zeitpunkt keine Spezifikation und es existiert auch kein Fertigstellungszeitpunkt dafür, da diese Schnittstelle sehr kompliziert ist. Im weiteren Verlauf wird deshalb nur die Schnittstelle M4 zu Wort kommen.

M4 ermöglicht die Ansichten NML (Network Management Level) und EML (Element Management Level) zu dem Managementsystem des Betreibers und dem öffentlichen ATM-Netzwerk. Bis jetzt wurden einige Anforderungen für das Konfigurations-, Fehler-, Leistungs- und Sicherheitsmanagement festgelegt. Dabei werden nur "Permanent Virtual Connections (PVCs)" unterstützt, während die Anwendbarkeit auf "Switched Virtual Connections (SVCs)" auf spätere Studien verschoben wurde.

Konfigurationsmanagement Das Konfigurationsmanagement, aus der Sicht des ATM-Netzwerkmanagements für die Schnittstelle M4, enthält folgendes :

1. ATM-NE (ATM Network Element) Konfigurationsidentifikation und Bericht über Änderungen

2. Konfiguration von UNIs (User Network Interface), BICIs (Broadband Inter Carrier Interface) und BISSIs (Broadband Inter Switching System Interface)
3. Konfiguration von "ATM VP/VC Link Termination Points and Cross Connections"
4. Konfiguration von "ATM VPC and VCC Termination Points"
5. Konfiguration von VPC und VCC OAM-Segmentendpunkten
6. Ereignisflußkontrolle (Kontrolle über automatisch generierte ATM-NE Berichte)

Für Punkt 1 ist es wichtig, daß die Sicht auf die verschiedenen ATM-NEs auch wirklich dem momentanen Zustand entsprechen, ansonsten ist kein effektives Management möglich. Folglich muß jedes ATM-NE dem Managementsystem eine abstrakte Sicht mit dem momentanen Zustand anbieten und alle späteren Änderungen, welche nicht ausdrücklich durch die M4-Schnittstelle geschehen, dem Managementsystem melden. Solche Änderungen können z.B. durch Netzwerkfehler oder durch menschliche Eingriffe (z.B. Einbau neuer Module) entstehen.

Um die gewünschte Funtionalität gewährleisten zu können, hat das ATM-Forum dafür einige Funktionen definiert :

- ATM-NE-Anzeigen, welche anzeigen, ob ein ATM-NE installiert/initialisiert ist und es verfügbar für das Managementsystem ist.
- Anfragen des Managementsystems über die momentane Konfiguration; Informationen über alle externen regelbaren physikalischen und logischen Komponenten eines ATM-NE und deren Beziehung zueinander müssen festgehalten und verfügbar gemacht werden.
- ATM-NE-Meldungen über Änderungen der Konfiguration
- Anzeigen von Änderungen des Operationsstatus von MOs innerhalb eines ATM-NE

Zur **Konfiguration von UNIs, BICIs und BISSIs** wurden folgende Funktionen festgelegt :

- Konfiguration der "physical path terminations"
- Anfragen zur Konfiguration
- Rekonfiguration

Zur Konfiguration von **ATM VP/VC Link Termination Points and Cross Connections** gibt es folgende Schlüsselpunkte :

- ein physikalischer Pfad kann in mehrer VPLs aufgesplittet sein
- ein VPC kann in mehrere VCLs aufgesplittet sein

- VPCs (VP Connection) sind eine Zusammensetzung von VPLs (VP Link)
- VCCs (VC Connection) sind eine Zusammensetzung von VCLs (VC Link)

Dazu wurden nun einige Anforderungen formuliert :

- Aufbau von “VPL-to-VPL cross-connections” und “VCL-to-VCL cross-connections” soll unterstützt werden
- Abbruch von VPL- und VCL-Verbindungen mit der Freigabe der belegten Ressourcen
- Sperren und Freigeben des Durchflusses von ATM-Zellen von einzelnen VP/VC-Verbindungen, wobei die Beziehungen und Ressourcen der Verbindung erhalten bleiben
- Wiederherstellen von alten VPL/VCL-Konfigurationsdaten
- Die obengenannten Punkte gelten auch für “multipoint VPL and VCL cross-connections”
- Auf- und Abbau von “VPL/VCL termination points” zu einer existierenden Mehrpunktverbindung

Fehlermanagement Das Fehlermanagement, aus der Sicht des ATM-Netzwerkmanagements für die Schnittstelle M4, enthält folgendes :

- Benachrichtigung des Managementsystems bei entdeckten Fehlern
- Aufzeichnen der Fehlerberichte
- Isolieren von Fehlern

Fehlerbenachrichtigungen sollten folgende Informationen enthalten :

1. Die fehlerhafte Komponente oder eine Liste der möglichen fehlerhaften Komponenten. Eine Komponente sollte dabei die kleinstmögliche auswechselbare oder reparierbare Einheit (Hard-, Software) darstellen.
2. Beschreibung des Fehlers anhand einer vorgegebenen Tabelle von Fehlerbeschreibungen.
3. Spezielle Probleme (für zukünftige Erweiterungen).
4. Einteilung des Fehlers in mehrere Härtegrade :

Critical Zeigt an, daß der angebotene Dienst des Netzwerkelementes davon betroffen ist und unmittelbare korrigierende Eingriffe notwendig sind (z.B. Totalausfall eines NEs).

Major Zeigt an, daß der angebotene Dienst des Netzwerkelementes davon betroffen ist und dringend korrigierende Eingriffe notwendig sind (z.B. Teilausfall eines NEs).

Minor Zeigt an, daß zwar der angebotene Dienst nicht beeinflusst wird, aber eine Korrektur sollte trotzdem durchgeführt werden, um größere Fehler zu vermeiden.

Warning Zeigt an, daß ein Fehler aufgetreten ist, welcher den angebotenen Dienst in Gefahr bringt.

Indeterminate Härtegrad konnte nicht bestimmt werden

Cleared Zeigt die Beseitigung von vorherigen Alarmen an.

5. Wurde die meldende NE ersetzt ?
6. Welche Einheit stellt einen Back-up-Service für die fehlerhafte Einheit bereit.
7. Platz für zusätzlichen Text (optional).
8. Reparaturvorschläge (optional)
9. Zeit und Datum der Fehlererkennung.

Leistungsmanagement Das Leistungsmanagement, aus der Sicht des ATM-Netzwerkmanagements für die Schnittstelle M4, enthält folgendes :

- Leistungsüberwachung
- Verkehrsmanagement
- UPC (User Parameter Control)/ NPC (Network Parameter Control) Überwachen auf Unstimmigkeiten
- Leistungsmanagementkontrolle
- Netzwerkdaten sammeln

Für die **Leistungsüberwachung** sollte der ITU-T Standard unterstützt werden, wie z.B. die Empfehlung G.774-01 und ANSI T1.231 für das Überwachen des SDH (SONET) Transportwesens. Außerdem sollte der ANSI T1 Standard, wie z.B. ANSI T1.pmnew und ANSI T1.231 für das DS-3 Transportwesen, unterstützt werden.

Für das **Verkehrsmanagement** sollte das M4-Interface im Stande sein, von allen ATM-Interfaces auf dem Netzwerkelement die aktuelle (15-Minuten) Anzahl von den folgenden Daten bereitzustellen :

1. wegen "Header Error Check (HEC)" abgelehnte Zellen
2. wegen Protokollfehler abgelehnte Zellen
3. OAM-Zellen, welche über ein Interface in das Netzwerkelement kamen.

Für jede UNI, BICI und BISSI, welche am NE endet, sollen folgende Informationen über die letzte, wegen einem Protokollfehler, abgelehnte Zelle bereitgestellt werden :

- Typ des Protokollfehlers (kein zugewiesener VPI/VCI-Wert, VPI/VCI-Wert ist außerhalb des Wertebereichs)
- VPI/VCI-Wert der abgelehnten Zelle
- Zeit und Datum

Dazu sieht die protokollunabhängige MIB für M4 z.B. folgendes MO vor :

- **Name** ATM Cell Protocol Monitoring Current Data

- **Attribute**

Managed Entity ID eindeutiger Name für das MO; (read-only)

Administrative State dient zum Ein- und Ausschalten (unlock/lock) der Funktion zum Sammeln von Daten, welche von diesem MO angeboten wird. (read/write)

Suspect Flag zeigt die momentane Zuverlässigkeit der angebotenen Dienste dieses MOs an. Es gibt zwei Zustände : reliable und unreliable (suspect). (read-only)

Elapsed Time zeigt die vergangene Zeit an, seit der angefangen wurde zu zählen. (read-only)

Threshold Data ID ist ein Zeiger auf das “Threshold Data”-MO, wo der Schwellwert für dieses MO steht. (read/write)

Number of Suppressed Intervals Anzahl der Intervalle, welche die Daten nicht an das “ATM Cell Protocol Monitoring History Data”-MO übergeben konnten. (read-only)

Discarded Cells due to protocol errors Anzahl der, wegen einem oder einer Menge von nicht erkennbaren Feldern, abgelehnten ATM-Zellen. Sollte der Schwellwert überschritten werden, dann wird dies Angezeigt. (read-only)

Received OAM Cells Anzahl der, über die verbundenen Schnittstellen, empfangenen OAM-Zellen. (read-only)

- **Meldungen :**

Threshold Crossing Alert : Benachrichtigung des Managementsystems, wenn der Schwellwert der abgelehnten ATM-Zellen überschritten wird. Folgende Werte sollten dabei mit übergeben werden :

- Die ID des meldenden MOs.
- Der Typ des Leistungsparameters, welcher den Schwellwert überschritten hat. Bei diesem MO kann dies nur *Discarded Cells due to protocol errors* sein.

Dieses MO zählt also die, aufgrund eines Protokollfehlers, abgelehnten Zellen.

UPC/NPC-Überwachung ist notwendig, um alle einkommenden Zellen auf deren Zulässigkeit bzgl. der vereinbarten Limits für diese Verbindung überprüfen zu können. Sollten diese Limits überschritten werden (z.B. Bitrate), dann werden diese Zellen abgelehnt. Da nun aber Zellen, welche von einer UPC/NPC-Funktion abgelehnt wurden, und Zellen, welche durch Übertragungsfehler oder Fehlfunktionen abgelehnt wurden, aus der Sicht der Ende-zu-Ende-Verbindung nicht zu unterscheiden sind, müssen dafür gewisse Werkzeuge geschaffen werden.

Sicherheitsmanagement Das Sicherheitsmanagement, aus der Sicht des ATM-Netzwerkmanagements für die Schnittstelle M4, enthält folgendes :

- Identifizierung
- Authentifizierung
- Kontrolle des Ressourcenzugangs
- Integrität
- Sicherheitsverwaltung

5 OAM-Zellen

In Zukunft werden noch mehr Managementfähigkeiten als bisher angeboten, allerdings mit den vom ATM-Forum definierten OAM-Zellen (Operationen, Verwaltung und Wartung) und nicht mit SNMP oder CMIP.

Es soll drei Typen von 53-Bytes OAM-Zellen geben, nämlich

- Fehlermanagement
- Leistungsmanagement
- aktivieren/deaktivieren (für das Starten und Beenden für Fehler- und Leistungsmanagementfunktionen).

OAM-Zellen ermöglichen es ATM-Netzwerkgeräten, Informationen über Ende-zu-Ende Verbindungen zu sammeln, reduzieren die Notwendigkeit, MIBs durch das Netz zu verteilen, und reduzieren den Verkehr an managementrelevanten Daten.

Beispielsweise wurde die AIS-Zelle (Alarm Indication Signal) definiert. Diese Zelle trägt Fehlerinformationen durch das Netz weiter. Sollte z.B. ein ATM-Switch ausfallen und dadurch eine VC (virtuell connection) unterbrochen werden, dann generiert jeder benachbarte ATM-Switch des Netzwerkes eine AIS-Zelle und schickt diese an alle vorherigen ATM-Switch. Die AIS-Zelle teilt diesen ATM-Switchen den Ausfall mit und gibt ihnen somit die Möglichkeit, eine alternative Route auszuwählen, falls die normale Route über den ausgefallenen ATM-Switch führen würde.

Mit einer anderen OAM-Zelle des Fehlermanagements kann man kontrollieren, ob eine VC ihr richtiges Ende erreicht. Dazu wird vom Anfang eine Zelle zum Ende geschickt, welche dort eine Markierung von der End-Station und das Zurückschicken zum Anfang verlangt. Somit kann kontrolliert werden, bei wem die VC endet.

6 Aussichten

Es läßt sich sagen, daß das Netzwerkmanagement ein Schlüssel zum Erfolg von ATM ist. Nur dadurch ist es möglich, die verschiedenen Anforderungen (siehe Abschnitt 3) an das ATM-Netzwerk in den Griff zu bekommen. Deshalb wurde zuerst darauf geachtet, daß ausführbare Standards für das Konfigurationsmanagement geschaffen wurden, um den Notwendigkeiten der existierenden ATM-Netzwerken gerecht zu werden, was z.B. auch zur Einführung der ILMI führte. Die Schnittstelle M4 wurde auch bevorzugt behandelt, da sie der Schlüssel zum Betrieb der öffentlichen Netzwerke ist, und ohne diese Netzwerke eine weiträumige Nutzung von ATM-Netzwerken (insbesondere ist hier B-ISDN zu nennen) nicht möglich ist. Deswegen mußten die anderen Schnittstellen erstmal zurückstecken. Die OAM-Zellen bieten dem Modell die Möglichkeit, mehr an Managementfähigkeiten anzubieten, als SNMP oder CMIP, was sich bei der weiteren Entwicklung des Modells bemerkbar machen wird.

Wegen der großen Komplexität dauert es aber noch lange, bis das Modell fertig definiert ist. Sollten aber in Zukunft alle Schnittstellen soweit fertig sein, dann läßt sich durchaus sagen, daß das ATM-Protokoll ein gutes Management besitzt.

ATM–Management mittels SNMPv2

Rudi Somen

Kurzfassung

ATM ist aus der Notwendigkeit heraus entstanden, verschiedene Dienste wie Audio und Video gemeinsam übertragen zu können. Zuerst wird die ATM-Topologie beschrieben und das Architekturmodell kurz aufgezeigt. Anschließend werden die Grundeigenschaften des ATM-Netzes etwas näher beschrieben (Mini-Pakete, Virtuelle Pfade, Virtuelle Kanäle). Im zweiten Teil wird zuerst erläutert, was die Aufgaben eines Netz-Management-Systems sind. Danach wird kurz die SMI beschrieben. Danach schließt sich im dritten Teil die Beschreibung der MIB-II an. Dabei wird vor allem die Konfigurationsgruppe inklusive der Interface-Tabelle näher erläutert. Die Ausarbeitung endet mit einer kurzen Bewertung.

1 Eigenschaften eines ATM-basierten Netzwerks

1.1 Heutige Netzwelt und ATM

Zielfunktionen für ein Universalnetz Die heutige Netzwelt ist gekennzeichnet durch eine Vielzahl von Netzvarianten im öffentlichen und privaten Bereich. Diese decken aber oft gleiche Territorien ab und sind im wesentlichen für anwendungsspezifische Dienstarten ausgelegt und daher zueinander disjunkt. Folglich gibt es eine Vielzahl von übereinanderliegenden Kommunikationssystemen mit jeweils speziellen Übertragungsaufgaben und damit verbunden unterschiedlichen Zugriffs- und Vermittlungsprinzipien.

Die erste Abhilfe aus diesem Dilemma versucht das dienstintegrierende digitale Netz ISDN zu schaffen, über das zumindest mehrere Dienste mit niedriger Datenrate — nämlich 64 kbit/s - gleichzeitig angeboten werden. Daher gibt es Bestrebungen, für alle gegenwärtigen und möglichst auch zukünftigen Anforderungen ein geeignetes Universalnetz zu konzipieren. Dabei tritt unmittelbar ein Hauptproblem einer solchen Strategie zutage, nämlich die Vereinigung sich scheinbar ausschließender Behandlungsverfahren für die verschiedenen zu transportierenden Informationsarten. Als Beispiel soll hier die Erwähnung von Leitungs- und Paketvermittlung genügen. Das Ergebnis all dieser Betrachtungen heißt nun ATM:

- Zerlegung der Informationsströme in kleine, separat zu behandelnde Einheiten (siehe Abbildung 60).
- Multiplexen verschiedener Dienste bzw. verschiedener Kommunikationsbeziehungen mit asynchronem Zeitmultiplex (um alle Ruhepausen einer Verbindung anderweitig zu nutzen).
- schnelle verbindungsorientierte Paketvermittlung dieser Zellen.
- hochratige Übertragung der Zellen, um beim Empfänger auch isochrone Datenströme wieder kontinuierlich aus dem Eingangspuffer auslesen zu können.

Zellkopf 5 Byte	Nutzlast 48 Byte
--------------------	---------------------

Abbildung 60. Grundstruktur einer ATM-Zelle

Aufbau des Zellenkopfes: Die 5 Byte teilen sich wie folgt auf:

- CLP — Cell Loss Priority (Zellverlustpriorität): 1 Bit
- HEC — Header Error Control (Bitfehler-Kontrolle): 8 Bit
- PT — Payload Type (Nutzlasttyp): 3 Bit
- VCI — Virtual Channel Identifier (VC-Identifikator): 16 Bit
- VPI — Virtual Path Identifier (VP-Identifikator): 8 Bit
- GFC — Generic Flow Control (Zugriffssteuerung an UNI): 4 Bit

ATM-Topologie Der Asynchrone Transfermodus als Übermittlungsprinzip für das Breitband-ISDN (B-ISDN) wurde mit der Absicht entwickelt, die Menge der zu unterstützenden Dienst- bzw. Verkehrsarten gegenüber konventionellen Weitverkehrsnetzen beträchtlich zu erweitern und auch für zukünftige Forderungen offen zu halten. Dies betrifft:

- die Integration von Diensten beliebiger Datenrate
- den Transport von Datenströmen mit variabler Datenrate
- die Gewährleistung möglichst aller Varianten von Partnerbeziehungen, wie zum Bsp.
 - simplex, Punkt-zu-Punkt (z.B. Datei-Transfer)
 - simplex, Multicast (z.B. Nachrichtendienst)
 - duplex, Punkt-zu-Punkt, symmetrisch bzgl. der Datenrate (z.B. Remote Procedure Call, Telefon)
 - duplex, Punkt-zu-Punkt, asymmetrisch bzgl. der Datenrate (z.B. Video on Demand)
- die Gewährleistung verbindungsorientierter und verbindungsloser Kommunikation
- die Emulation typisch paket- als auch typisch leitungsvermittelter Verbindungen

Dieses Aufgabenspektrum unter besonderer Berücksichtigung einer (zumindest für die erste Aufgabe erforderlichen) hohen Datenrate hat unmittelbare Auswirkungen auf die Topologie eines dafür geeigneten Netzes. Für das B-ISDN ist die Verwendung des topologischen Prinzips des geschalteten (switched) Mediums prädestiniert; d.h., das Netz ist

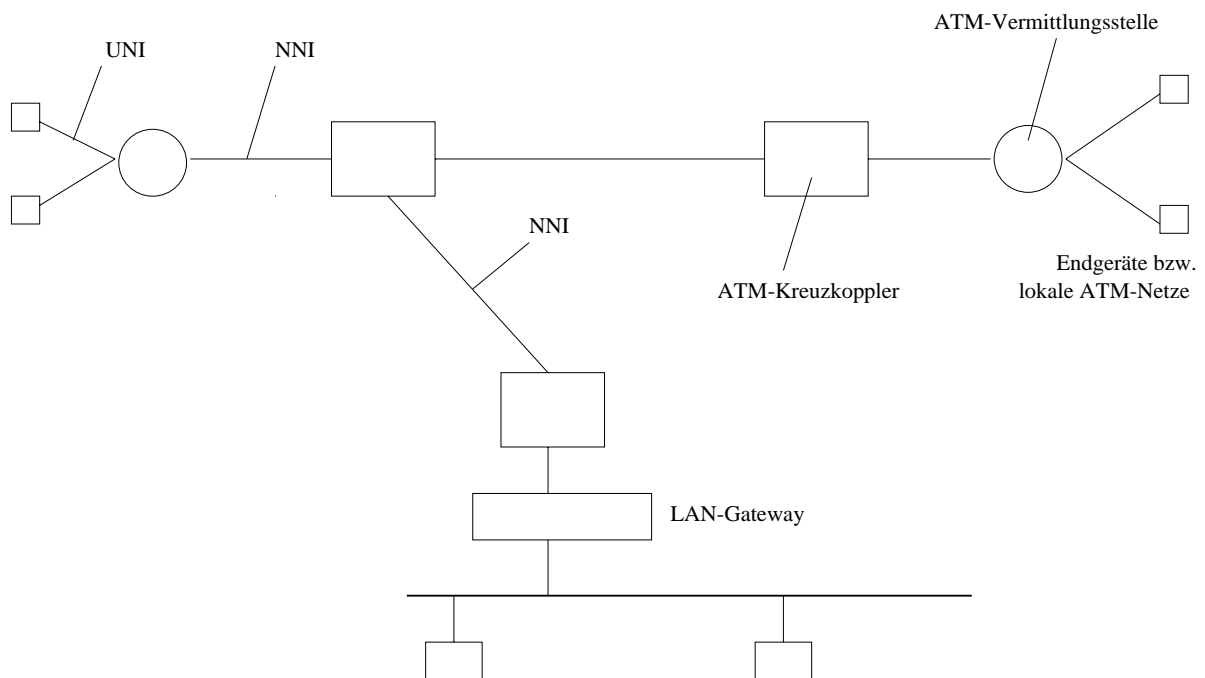


Abbildung 61. Topologische Struktur eines ATM-Netzes

physikalisch zweistufig strukturiert und besteht aus den vermaschten Netzknoten, die die Wegeverschaltung vornehmen, und den daran sternförmig angeschlossenen Endsystemen. Die angegebene Topologie enthält auch zwei für ein ATM-Netz typische Schnittstellen, nämlich die Nutzer-Netz-Schnittstelle (UNI — User Network Interface) und die Netzknoten-Schnittstelle (NNI — Network Node Interface), die sich hinsichtlich der dort übertragenen Datenstrukturen und ihrer Behandlungsprinzipien leicht unterscheiden.

Es gibt zwei verschiedene Arten von Netzknoten,

- die ATM-Kreuzkoppler und
- die sogenannten ATM-Vermittlungsstellen (siehe auch Kapitel 2)

ATM-Architektur Da der Asynchrone Transfermodus für den Weitverkehrsbereich entwickelt wurde, ist auch die logische Struktur diesem Einsatzfall angepaßt. Sie umfaßt also eine hierarchische Schichtenarchitektur, welche nur übertragungsspezifische und keine anwendungsspezifischen Aufgaben einer Netzarchitektur beinhaltet.

Die wesentliche Funktionalität des Asynchronen Transfermodus ist die Realisierung spezieller Multiplex- und Vermittlungsfunktionen. Bezüglich des Zeitmultiplex betrifft das die Herstellung eines kontinuierlichen Zellenstroms durch asynchronen Zeitteilung (ATD — Asynchronous Time Division). Anders als bei der synchronen Zeitmultiplex (STD — Synchronous Time Division) ist hier die Zuordnung eines Multiplexkanals zu einer Kommunikationsbeziehung nicht durch die Lage in einem Übertragungsrahmen bestimmt, sondern durch spezielle Identifikatoren im Zellkopf. Damit ergibt sich — allerdings unter Akzeptanz eines größeren Overheads für Steuerungsfunktionen — der entscheidende Vorteil gegenüber STD, auch Kanäle mit verschiedenen und sogar variablen Datenraten

multiplexen zu können. Darüber hinaus ist durch dieses Prinzip auch die Anpassung an stoßartige — burst type — Informationsströme effektiv gelöst, da nur bei tatsächlich vorliegendem Datenaufkommen Zellen generiert und in den Multiplexstrom eingeordnet werden. Ist dagegen die Übertragungsrate des Zellstroms größer als das Informationsaufkommen der gemultiplexten Quellen, werden speziell gekennzeichnete Leerzellen in den Strom eingefügt.

Zur Gewährleistung dieses Multiplexens müssen die Datenströme der betroffenen Quellen in Abschnitte der Nutzlastgröße einer Zelle zerlegt, mit einem Zellkopf versehen, entsprechend ihrer Ankunftszeitpunkte in eine gemeinsame Warteschlange eingeordnet und mit konstanter Zellrate ausgelesen werden. Die Leitweglenkung in ATM-Netzen beruht auf dem Prinzip der Paketvermittlung. Durch die Segmentierung der zu transportierenden Datenströme in sehr kleine Pakete fester Länge ist einerseits die Gewährleistung traditionell paketvermittelter Dienste problemlos gesichert, aber andererseits auch die Durchschaltung isochroner (kontinuierlicher) Datenströme, die bisher nur leitungsvermittelt übertragen wurden, mit akzeptablen Leistungseinschränkungen (z.B. hinsichtlich der Schwankung der Verzögerungszeit oder der Zellverlustrate) möglich.

Wenn man die genannten Multiplex- und Vermittlungsfunktionen von ATM einem Trichter zuordnet, so ist für die Konstruktion einer ATM-Architektur noch die Datensegmentierung (und Reassemblierung auf Empfängerseite) oberhalb und der Zellentransport unterhalb des Trichters einzubeziehen, wodurch eine Dreiteilung der Aufgaben in vertikaler Richtung entsteht. Eine Aufgabenteilung in horizontaler Richtung ist dadurch bedingt, daß wie beim ISDN auch für ATM-Netze eine Outband-Signalisierung (Übertragung von Steuerungs- und Nutzdaten über separate Kanäle) vorgesehen ist. Unter zusätzlicher Berücksichtigung der in allen Netzen üblichen Verwaltungsaufgaben (Management) ergibt sich in logischer Konsequenz ein Architekturmodell:

Einige grundsätzliche Bemerkungen zu diesem Modell seien hier angefügt:

1. Die Managementebene läßt sich in Komponenten zum Schichtenmanagement und in ein schichtenübergreifendes Netzmanagement separieren.
2. Das Nebeneinander von Nutzer- und Steuerungs- bzw. Signalisierungsebene reicht bis in die ATM-Anpassungsschicht (AAL — ATM Adaption Layer) hinein. Das Zusammenfassen (Multiplexen) beider Ebenen erfolgt wie auch für unterschiedliche Nutzdatenströme durch die ATM-Schicht.
3. Die Aufgaben der höheren Schichten sind von der jeweiligen Anwendungen bzw. von der Art des Datenstroms abhängig und nicht ATM-spezifisch.
4. Die für den Asynchronen Transfermodus relevante Struktur aus drei Schichten ist nicht unmittelbar an das OSI-Modell der ISO anpaßbar. Begründung: Wegen der Vermittlungsfunktion der ATM-Schicht enthält diese schon Aufgaben der OSI-Netzwerkschicht. Außerdem sind auch Netzwerkfunktionen (z.B. Routing von Paketen) auf die ATM-Schichtenstruktur aufsetzbar.

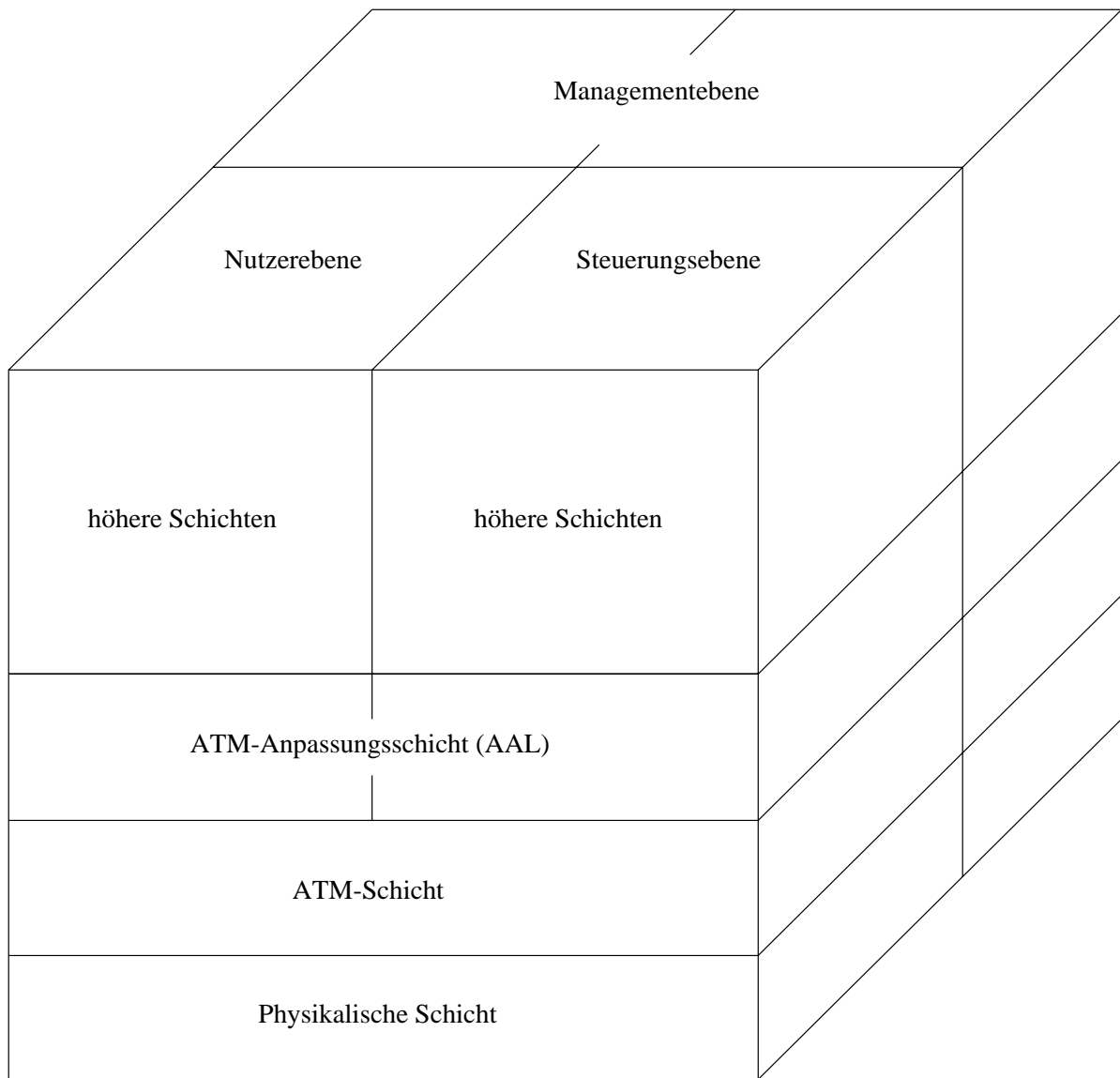


Abbildung 62. Breitband-ISDN-Referenzmodell

1.2 Grundeigenschaften des ATM

Minipakete fester Größe Die ausschließliche Verwendung von Minipaketen bildet die Voraussetzung für eine erfolgreiche Realisierung stark zeitabhängiger Übertragungsdienste durch Paketvermittlungsmnetze. ATM nutzt wie gesagt Minipakete fester Größe, sogenannte ATM-Zellen. Sie besitzen einheitlich eine Größe von nur 53 Byte. Davon entfallen fünf Byte auf die Steuerdaten im Paketkopf und 48 Byte auf das Nutzinformationsfeld der Zelle.

Zeitrelevanz in Paketvermittlungsnetzen: In Paketnetzen setzt sich die Übertragungsverzögerung für das Element eines Datenstroms (d.h. für die kleinste Datenstruktur innerhalb eines Informationsstromes) aus mehreren Zeitkomponenten zusammen:

- Paketiervverzögerung im Sender bzw. Depaketiervverzögerung im Empfänger
- konstante Signalverzögerung

- Paketverzögerung an den Zwischenknoten im Netz (z.B. an den Vermittlungsstellen)

Asynchrones Zeitmultiplex

Zeitrelevanz in Paketvermittlungsnetzen Bei ATM teilt sich das Übertragungsmedium in einzelne Übertragungsabschnitte (Links) auf, die innerhalb des Netzes jeweils von Vermittlungseinrichtungen begrenzt und durch diese miteinander verknüpft werden.

An den Netzgrenzen sind die Links mit Endeinrichtungen verbunden. Diese sind entweder netzkonforme Nutzerstationen oder übernehmen Netzanpassungsfunktionen für nichtkonforme Nutzerstationen. Jede Dateneinheit wird von Vermittlungsstation zu Vermittlungsstation über die Links zur Empfängerstation weitergeleitet. Eine Vermittlungsstation hat mehrere Eingänge und Ausgänge (Ports), um die vielen ankommenden und abgehenden Links bedienen zu können. Der Zugriff der konkurrierenden Verkehrsströme auf die Vermittlungsstelle, bzw. auf eine bestimmte Ausgangsleitung beim Verlassen der Station, muß daher geregelt werden.

ATD-Prinzip Die WAN-Paketvermittlungsnetze, z.B. auf der Basis von X.25 oder ATM, nutzen das asynchrone Zeitmultiplex (Asynchronous Time Division ATD — ATD) zur Steuerung des konkurrierenden Ressourcenzugriffs. Es regelt den Zugriff, wenn Übertragungswünsche vorliegen, jeweils zu stochastischen Zeitpunkten (Multiplexen nach Anforderungsprinzip). Das herausragende Merkmal des ATD besteht in seiner effektiven Ressourcenauslastung. Der ATD-Multiplexer ist als einfaches Bedienmodell mit Wartespeicher vor einer Bedieneinrichtung (der Ressource) beschreibbar. Die Ankunftsleitungen der konkurrierenden Verkehrsströme enden in einem gemeinsamen Paketpuffer. Wenn die Bedieneinrichtung nicht belegt ist, entnimmt sie jeweils das Paket mit dem frühesten Ankunftszeitpunkt aus dem Wartespeicher (First-Come-First-Serve-Prinzip — FCFS). Um die Pakete nachher wieder ihrer Quelle zuzuordnen zu können, wird ein Identifikator für jedes Paket notwendig.

Zellenvermittlung über virtuelle Kanäle Das Basiskonzept für die Vermittlung von ATM-Zellen beruht auf den virtuellen Kanälen. Zwischen zwei ATM-Endsystemen wird während des Verbindungsaufbaus in einer Richtung eine logische Verbindung erzeugt. Gleichzeitig erfolgt deren physikalische Abbildung auf eine feste Route durch das ATM-Netzwerk. Jeder Teilstrecke zwischen zwei Vermittlungsstellen wird dabei ein logischer Kanal zugeordnet, ein sogenannter virtueller Kanal (Virtual Channel — VC). Da sich mehrere logische Verbindungen gleichzeitig einen physikalischen Streckenabschnitt teilen können, existiert auf jedem Übertragungsabschnitt eine entsprechende Anzahl unterschiedlicher virtueller Kanäle. Als Unterscheidungsmerkmal besitzt jeder virtuelle Kanal einen eindeutigen Identifikator (Virtual Channel Identifier — VCI). Eine (Simplex-) Verbindung zwischen zwei ATM-Endsystemen ist deshalb durch die Abfolge der virtuellen Kanäle eindeutig bestimmt. Unmittelbar bevor eine ATM-Zelle einen virtuellen Kanal durchquert, wird dessen Identifikator in ein Steuerfeld (siehe auch Kap. 1.3) im ATM-Zellkopf eingetragen. Eine ATM-Zelle kann damit auf jeder Teilstrecke anhand

des VCI-Eintrages im Zellkopf einer konkreten Verbindung zugeordnet werden. Die VC-Vermittlung der ATM-Zellen in einer ATM-Vermittlungsstelle beschränkt sich auf die Aktionen: Aufnehmen der ATM-Zelle am Eingangsport, Lesen des VCI, Suchen des VCI-Eintrages in einer Tabelle, Feststellung des neuen VCI-Wertes für die nächste Teilstrecke der Verbindung, entsprechende Änderung im VCI-Feld der Zelle und Weiterleitung der Zelle im Vermittlungsnetzwerk zum zugehörigen Ausgangsport.

Konzept der virtuellen Pfade Das Konzept der virtuellen Pfade (Virtual Path — VP) dient der Vorvermittlung eines Bündels von virtuellen Kanälen und erleichtert die Netzwerkorganisation. Eine Pfadverbindung kann als unidirektionale logische Verbindung zwischen zwei, nicht notwendigerweise benachbarten Vermittlungseinrichtungen aufgefaßt werden. Es existiert für jede Pfadverbindung eine vordefinierte Route über dem physikalischen Netz. In Analogie zur Verbindung von ATM-Endnutzern über die virtuellen Kanäle setzt sich die Pfadverbindung aus einer Verkettung von virtuellen Pfaden entlang der physikalischen Teilstrecken zusammen. Jeder dieser virtuellen Pfade besitzt einen auf seiner physikalischen Teilstrecke eindeutigen Identifikator (Virtual Path Identifier — VPI). Jede ATM-Zelle enthält im Zellkopf eine Verbindungskennung, die sich aus der Kennung für den genutzten virtuellen Pfad (VPI) und dem VCI zusammensetzt. In den ATM-Kreuzkopplern werden alle Zellen mit einem gleichen VPI, unabhängig von ihrem VCI, gleichartig weitergeleitet. Die sogenannte 'Feinvermittlung' anhand des VCI bleibt den ATM-Vermittlungsstellen vorbehalten. Im Gegensatz zu den ATM-Kreuzkopplern werten diese sowohl den VPI als auch den VCI aus. Das VP-Konzept beschleunigt den Aufbau von VC-Verbindungen. Zellen unterschiedlicher VC-Verbindungen, die gleiche (Teil-)Strecken im Übertragungsnetzwerk überbrücken, lassen sich zusammenfassen. Dadurch vereinfachen sich auch die Routingprozeduren, da auf schon existierende Pfad-Routen zurückgegriffen werden kann.

Verkehrs-Verwaltungs-Parameter

Zellen-Verlust-Priorität Um während eines Verkehrsstaus diesen besser regeln zu können, gibt es zwei Prioritätsstufen: CLP=0,1. (CLP — Cell Loss Priority). CLP=0 hat eine höhere Priorität als CLP=1. D.h. während eines Verkehrsstaus werden CLP=1-Zellen eher fallengelassen als CLP=0-Zellen.

Dienstqualitäts (QoS -Quality of Service) Klassen Eine VCC oder VPC wird mit einer Anzahl von QoS-Klassen in Verbindung gebracht. Es gibt vier verschiedene Dienst-Klassen:

- Dienst-Klasse A: Konstante Bitrate Video
- Dienst-Klasse B: Variable Bitrate Video/Audio
- Dienst-Klasse C: Verbindungsorientierte Daten
- Dienst-Klasse D: Verbindungslose Daten

Vier QoS-Klassen (1 — 4) wurden mit dem Ziel spezifiziert, die Dienst-Klassen A-D zu unterstützen. Die verketteten VCLs (oder VPLs), die eine VCC (bzw. VPC) bilden, haben alle diesselbe QoS-Klasse wie die VCC (bzw. VPC). Der Zellen-Verlust-Verhältnis (Cell Loss Ratio — CLR), die Zellen-Verzögerungs-Variation (Cell Delay Variation — CDV) und die Ende-zu-Ende-Zellenverzögerung (End-to-End Cell Delay — CD) sind definiert als Teil der QoS-Klassendefinition. Zusätzlich gibt es eine unspezifizierte QoS-Klasse, Nummer 0, welche die gerade eben mögliche Leistung anzeigt.

1.3 Abbildung von Kommunikationsdiensten auf ATM

Notwendigkeit einer Anpassungsschicht Der reine ATM-Schicht-Dienst (ATM bearer service) arbeitet verbindungsorientiert, kennt nur die starre ATM- Zellenstruktur und behandelt keine Störeinflüsse, die u.a. aus der konkurrierenden Existenz mehrerer Datenströme im ATM-Netz resultieren. Es besteht daher die Aufgabe, auf dem universellen ATM-Verfahren verschiedene komfortable Kommunikationsdienste anzubieten, die von ATM-Nutzern unkompliziert für ihre speziellen Aufgaben verwendet werden können. Zwischen der ATM-Schicht und dem ATM-Nutzer residiert deshalb eine Anpassungsschicht (ATM Adaptation Layer — AAL). Sie soll den ATM-Schicht-Dienst ergänzen bzw. vervollkommen. Die AAL realisiert notwendige Adaptionfunktionen bei der Datenübertragung und bei der Verbindungssteuerung sowie zugehörige Managementaufgaben. Insgesamt gibt es fünf AAL-Typen. Nummer 1, 3/4 und 5 sind vom CCITT in den Empfehlungen I.362 und I.363 spezifiziert. Der Typ 3/4 ging aus den ursprünglich getrennten Typen 3 (verbindungsorientiert) und 4 (verbindungslos) hervor. Der Typ 5 basiert auf Vorschlägen zur Vereinfachung des Typs 3. AAL-Typ 2 befindet sich noch in der Entwurfsphase.

Strukturierung der Anpassungsschicht Unabhängig vom jeweiligen Typ besitzt die AAL zwei Hauptaufgaben:

- Abbildung von Nutzerdateneinheiten auf das ATM-Zellenformat am Sender und Wiederherstellung der Dateneinheiten am Empfänger
- Realisierung direkt anwendungsspezifischer und allgemein dienstklassenspezifischer Funktionen

Entsprechend dieser Aufgaben ist die AAL logisch in sich gegliedert.

Sie besteht aus einer Segmentierungs- und Reassemblierungsschicht (Segmentation and Reassembly Sublayer — SAR), die sich direkt über der ATM-Schicht befindet, und aus einer Konvergenz-Subschicht (Convergence Sublayer — CS) darüber. Diese ist ggf. nochmals logisch in einen dienstklassenspezifischen gemeinsamen Teil, die Common Part Convergence Sublayer (CPCS), und einen anwendungsspezifischen Teil, die Service Specific Convergence Sublayer (SSCS), unterteilt. Ein AAL-Nutzer besitzt nur über die jeweils oberste Teilschicht Zugriff auf die AAL. Die internen Teilschichten sind nicht frei zugänglich. Die Teilschichten tauschen untereinander hierarchisch Aufträge und Meldungen mit funktionsbezogenen Parametern aus.

AAL-Typ 1 Der AAL-Typ 1 unterstützt höhere Dienste mit konstanter Bitrate. Im Mittelpunkt steht die Realisierung einer Sender-Empfänger-Synchronisation mit dem Hauptaugenmerk auf spezielle Verzögerungs- und Zeittaktforderungen.

AAL-Typ 2 Der AAL-Typ 2 soll die universellen Fähigkeiten des ATM vorteilhaft ausnutzen und zur Geltung bringen. Einerseits wird die AAL für eine Sender-Empfänger-Synchronisation während des Zeitraums einer Verbindung sorgen, wie es die Nutzer von der Leitungsvermittlung her kennen, andererseits wird sie die Generierung von ATM-Zellen nur entsprechend dem tatsächlichen Nutzerdatenaufkommen initiieren. Die letztgenannte Eigenschaft ist typisch für die effektive Datenübertragung in Paketvermittlungsnetzen. Die AAL 2 ist deshalb speziell für die Unterstützung solcher höheren Dienste prädestiniert, die einen Sender- Empfänger-Gleichlauf erfordern. Zu dieser Kategorie gehören z.B. die Audio- und Videodienste mit variabler Bitrate.

AAL-Typ 3/4 Die CPCS der AAL 3/4 stellt den erwähnten unsicheren Transportdienst von Nutzerdatenframes bereit (Länge bis zu 65535 Bytes).

AAL-Typ 5 Der AAL-Typ 5 stellt eine stark vereinfachte AAL dar und ist als Reaktion auf den komplexen AAL-Typ 3/4 entstanden. Die SAR der AAL5 generiert quasi keinen Steuerdaten-Overhead.

1.4 Übertragung der ATM-Zellen

Fehlersicherung und Zellratenanpassung Die Übertragung der ATM-Zellen ist eine Aufgabe der physikalischen Schicht des Referenzmodells. Diese ist in zwei Subschichten geteilt. Bedeutung des Fehlerkontrollfeldes (HEC — Header Error Control): Sicherungsmechanismen gegen Übertragungsfehler sind vorrangig auf Methoden des CRC (Cyclic Redundancy Check) von Dateneinheiten oder ihren Abschnitten beschränkt. Die ersten vier Byte des Zellkopfes werden mittels eines CRC-Generatorpolynoms geschützt und das Divisionsergebnis im HEC-Feld eingetragen. Die Übertragung des senderseitig generierten Zellstromes kann sowohl ohne weitere Strukturierung als reines ATM-Zellmultiplex (cell based ATM) oder durch Einschachteln der Zellen in spezielle Übertragungsrahmen (framed ATM) erfolgen.

2 Struktur der Management-Information (SMI)

2.1 Objekt-Definitionen

Auf verwaltete Objekte wird durch einen virtuellen Informations-Speicher, genannt MIB (Management Information Base) zugegriffen. Objekte in der MIB sind in der Untermenge der Abstract Syntax Notation One (ASN.1) definiert. Jeder Objekttyp wird von einem OBJECT IDENTIFIER benannt. Der Objekttyp zusammen mit einer Objektinstanz identifiziert eindeutig eine spezielle Instanziierung des Objekts. Aus Bequemlichkeit benutzt man oft einen Textstring, genannt den Deskriptor, um sich auf den Objekttyp zu beziehen.

2.2 Informations-Module

Ein Informations-Modul ist ein ASN.1 Modul, welches die Informationen, die sich auf das Netz-Management beziehen, definiert. Die SMI beschreibt, wie man eine Untermenge von ASN.1 benutzt, um so ein Informations-Modul zu definieren. Es gibt drei Arten von Informations-Module:

1. MIB-Module, die Definitionen von miteinander in Beziehung stehenden Verwaltungs-Objekten enthalten. Diese machen Gebrauch von den Makros namens OBJECT-TYPE und NOTIFICATION-TYPE.
2. Darstellungen für die MIB-Module, welche die Übereinstimmungen betreffen. Diese benutzen die Makros mit Namen MODULE-COMPLIANCE und OBJECT-GROUP.
3. Darstellungen für die Agenten-Implementierungen, welche die allgemeinen Fähigkeiten betreffen. Diese benutzen Makros wie das AGENT-CAPABILITIES-Makro.

Alle Informations-Module starten mit genau einem Aufruf des MODULE-IDENTITY-Makros. Dies stellt den Kontakt her und liefert die Überarbeitungs-Historie.

2.3 Namens-Hierarchie

Die Wurzel für den Unterbaum, die von der Internet Assigned Numbers Authority (IANA) verwaltet wird, ist:

internet OBJECT IDENTIFIER ::= (iso 3 6 1)

d.h. der Internet-Unterbaum des OBJECT IDENTIFIERS beginnt mit dem Präfix: 1.3.6.1 Für das Netz-Management werden unterhalb dieses Unterbaums verschiedene Zweige benutzt:

mgmt OBJECT IDENTIFIER ::= (internet 2)
experimental OBJECT IDENTIFIER ::= (internet 3)
private OBJECT IDENTIFIER ::= (internet 4)
enterprises OBJECT IDENTIFIER ::= (private 1)

Trotzdem verbietet das SMI die Definition von Objekten in anderen Teilen des Objekt-Baumes nicht. Der mgmt-Unterbaum (2) identifiziert „Standard“-Objekte. Der experimental-Unterbaum (3) identifiziert Objekte, welche von der Arbeitsgruppe der IETF definiert werden. Der private-Unterbaum (4) identifiziert Objekte, welche einseitig definiert werden. Der enterprises-Unterbaum (1) unterhalb des private-Unterbaums wird unter anderem dazu benutzt, um Providern von Netz-Komponenten zu ermöglichen, Modelle ihrer Produkte registrieren zu lassen.

3 Definition der Management-Objekte

3.1 Einführung

Im folgenden wird ein Teil der Management Information Base (MIB) definiert. Hauptsächlich handelt es sich um Objekte, welche zum Verwalten von ATM-basierten Schnittstellen, Geräten, Netzen und Diensten dienen.

3.2 Das SNMPv2 Netz-Management-Grundgerüst

Das SNMPv2 Netz-Management-Grundgerüst besteht aus vier Hauptkomponenten:

- RFC (Request for comments) 1442, das die SMI definiert.
- STD 17, RFC 1213; definiert MIB-II, die Grundmenge der Verwaltungsobjekte.
- RFC 1445; definiert die administrativen und architektonischen Aspekte des Grundgerüsts.
- RFC 1448; definiert das Protokoll, welches zum Zugang zu den Verwaltungsobjekten benutzt wird.

3.3 Überblick

Im folgenden wird hauptsächlich die Verwaltung von ATM-PVCs gezeigt (weniger ATM-SVCs).

Struktur der MIB Die verwalteten ATM-Objekte sind in sechs Gruppen angeordnet:

- ATM-Schnittstellen-Konfigurationsgruppe
- ATM-Schnittstellen-DS3-PLCP-Gruppe
- ATM-Schnittstellen-TC-Unterschicht-Gruppe
- ATM-Schnittstellen-Virtuelle Link- (VPL/VCL) Konfigurationsgruppen
- ATM-VP/VC-Kreuzverbundene Gruppen
- AAL5: Statistikgruppen für die Verbindungsleistung

ATM-Schnittstellen-Konfigurationsgruppe Diese Gruppe enthält Information der ATM-Zellenschicht-Konfiguration.

ATM-Schnittstelle DS3 PLCP und TC-Schicht Diese Gruppen liefern Leistungs-Statistiken der DS3 PLCP und der TC-Unterschicht von lokalen ATM-Schnittstellen.

Virtueller ATM Link und Kreuzverbundene Gruppen Diese Gruppen modellieren bidirektionale virtuelle ATM-Links und ATM-Kreuzverbindungen. Die ATM VP/VC Link-Gruppen sind in einem ATM-Host, einem ATM-Switch und in einem ATM-Netz implementiert. Die virtuellen ATM Link-Gruppen werden benutzt, um virtuelle ATM-Links in einem ATM-Host (switch, Netz) zu erzeugen bzw. löschen oder ändern.

3.4 Anwendung von MIB II auf ATM

Die Schnittstellengruppe Die Schnittstellengruppe definiert Objekte für die Verwaltungsschnittstellen. Im folgenden werden die medienspezifischen Erweiterungen auf die Schnittstellengruppen zum Verwalten von ATM-Schnittstellen aufgezeigt. Im folgenden wird angenommen, daß die Interpretation der Schnittstellengruppe genau mit der Schnittstellentabelle (Interface Table — IfTable) übereinstimmt. Die IfTable enthält Information über die verwalteten Ressourcenschnittstellen. Außerdem wird jede Unterschicht unterhalb der Zwischennetzschicht einer Netzschnittstelle als Schnittstelle betrachtet. D.h. die ATM-Zellschichtschnittstelle wird als ein Eintrag in der IfTable angesehen. Dieser Eintrag betrifft die ATM-Zellschicht als Ganzes.

Unterstützung der ATM-Zellschicht durch die ifTable Im folgenden werden einige Interpretationen der IfTable für die ATM-Zellschicht angezeigt:

Objekt	Nutzen für die generische ATM-Schicht
IfIndex	Jeder ATM-Port wird durch einen IfEintrag repräsentiert
IfDescr	Beschreibung der ATM-Schnittstelle
IfType	37 (der Wert der für ATM zugewiesen wurde)
IfSpeed	Die gesamte Bandbreite in Bits pro Sekunde
IfPhysAddress	Die Schnittstellenadresse an der ATM-Protokollunterschicht

3.5 Definitionen

Die folgende Gruppe enthält ATM-spezifische Konfigurations-Information in Verbindung mit einem ATM- Interface:

```
atmInterfaceConfEntry OBJECT-TYPE
SYNTAX AtmInterfaceConfEntry
MAX-ACCESS not-accessible
STATUS current
DESCRIPTION
"Diese Liste enthaelt ATM-Schnittstellen-Konfigurations-Parameter
und Statusvariablen"
INDEX { IFIndex  }
 ::= { atmInterface ConfTable 1 }
```

Einige Erläuterungen:

SYNTAX: Gibt den Datentyp an

STATUS: „current“ oder „historic“

MAX-ACCESS: Gibt an, ob man die Instanz eines Objektes lesen, schreiben und/oder erzeugen kann

```
AtmInterfaceConfEntry ::= SEQUENCE {
atmInterfaceMaxVpcs INTEGER,
atmInterfaceMaxVccs INTEGER,
atmInterfaceConfVpcs INTEGER,
atmInterfaceConfVccs INTEGER,
atmInterfaceMaxActiveVpiBits INTEGER,
```

```

atmInterfaceMaxActiveVciBits INTEGER,
atmInterfaceMaxIlmiVpi INTEGER,
atmInterfaceMaxIlmiVci INTEGER,
atmInterfaceMaxAddressType INTEGER,
atmInterfaceMaxAdminAddress OCTET STRING,
atmInterfaceMyNeighborIpAddress IpAddress,
atmInterfaceMyNeighborIfName DisplayString,
}

```

```

atmInterfaceMaxVpcs OBJECT-TYPE
SYNTAX INTEGER (0..4096)
MAX-ACCESS read-write
STATUS current
DESCRIPTION
"The Maximum number of VPCs supported at this ATM interface. At the
ATM UNI, the maximum number of VPCs ranges from 0 to 256 only."
::= { atmInterfaceConfEntry 1 }

```

Erläuterung: Gibt die maximale Anzahl der VPCs (virtuelle Pfadverbindungen) an.

```

atmInterfaceConfVpcs OBJECT-TYPE
SYNTAX INTEGER (0..4096)
MAX-ACCESS read-only
STATUS current
DESCRIPTION
"The number of VPCs configured for use at this ATM interface.
At the ATM UNI, the configured number of VPCs can range from 0 to 256
only."
::= { atmInterfaceConfEntry 3 }

```

Erläuterung: Gibt die Anzahl der VPCs (virtuelle Pfadverbindungen) an.

```

atmInterfaceMaxActiveVpiBits OBJECT-TYPE
SYNTAX INTEGER (0..12)
MAX-ACCESS read-write
STATUS current
DESCRIPTION
"The maximum number of active VPI bits configured for use at the ATM
interface. At the ATM UNI, the maximum number of active VPI bits
configured for use ranges from 0 to 8 only."
::= { atmInterfaceConfEntry 5 }

```

Erläuterung: Gibt die maximale Anzahl der aktiven VPI-Bits an.

```

atmInterfaceAddressType OBJECT-TYPE
SYNTAX INTEGER {
private(1),
nsapE164(2),
nativeE164(3),
other(4)
}
MAX-ACCESS read-only
STATUS current

```

DESCRIPTION

```
"The type of primary ATM address configured for use at this ATM  
interface."  
::= { atmInterfaceConfEntry 9 }
```

Erläuterung: Gibt die primäre ATM-Adresse an.

4 Bewertung

Das ATM-Protokoll wird in den nächsten Jahren sicherlich sowohl im LAN- als auch im Weitverkehrsbereich eine immer wichtigere Stellung erlangen. Daher ist das Management eines ATM-Netzes von immenser Wichtigkeit. Mit SNMPv2 liegt nun ein neues Managementprotokoll vor: Mit dem Modell der SMI scheint ein Weg gefunden zu sein, die Verwaltung der vielen Objekten in den Griff zu bekommen. Die Objekte sind in einige wenige Gruppen angeordnet, deren Verwaltung logisch erscheint. Daher erscheint es empfehlenswert, sich mit dem Management-Konzept, wie in SNMPv2 mit dem ATM-Netz verfolgt, weiter zu beschäftigen.

Vorteile von SNMP v2 gegenüber v1:

- erweitertes Kommunikationsmodell: Manager-Agent
- neue Sicherheitskonzepte
- Übertragung größerer Datenmengen in einer SNMPv2-Nachricht
- Manager-zu-Manager-Kommunikation

Nachteile: SNMP v2 nicht mehr einfach, wie es der Name sagt.

Fazit: Es wird momentan an einer Erweiterung von SNMP v2 gearbeitet, da sich v2 seit einiger Zeit überlebt hat.

ATM-LAN Emulation und seine Verwaltung

Daniel Schempf

Kurzfassung

Der asynchrone Transfermodus bildet die Basis für die kommende Breit-band-Kommunikation. Deswegen wird ATM in den nächsten Jahren mit „herkömmlichen“ Netzen kooperieren müssen. Hierzu sind Migrationswege gefragt, um diesen Leistungsaufstieg möglich zu machen und die bisherigen Investitionen in LANs zu schützen. Darum muß es möglich, sein vorhandene Protokollstacks und Software auf den Endgeräten ohne Änderungen der Hard- oder Software auf ATM-Netze übernehmen zu können. Einen umfassenden Vorschlag hierzu bietet die „LAN-Emulation over ATM Specification Version 1.0“ des ATM-Forums, welche ein Client/Server-Modell definiert. Die LAN-Emulation und dessen Verwaltung am Beispiel des LAN-Emulation Clients soll in diesem Beitrag vorgestellt werden.

1 Einführung

1.1 Grundlagen der LAN-Emulation

Damit vorhandene Applikationen und Protokolle ohne Änderung weiter benutzt werden können, muß das ATM-Netz eine Schnittstelle bereitstellen, die das Verhalten eines LANs wie z.B. Ethernet oder Token-Ring emuliert. Dies betrifft hauptsächlich den Unterschied in dem Verbindungsaufbau und der Adressierung. Die LAN-Emulation [Epp95, Inc95] wird dabei entweder im ATM-Adapter für ein Endgerät oder im Cell-/Frame-Switch für alle Endgeräte der angeschlossenen LANs implementiert.

Verbindungslos vs. verbindungsorientiert. Da bei ATM für jeden Sendewunsch eine dedizierte Verbindung aufgebaut wird, die heutigen Netze aber üblicherweise verbindungslos arbeiten, muß die Verwendung von Broad- und Multicasts auf ATM-Strukturen abgebildet werden, d.h. das Verhalten eines geteilten Mediums LAN muß emuliert werden. Mehr dazu bei der Behandlung der Broadcast-Mechanismen.

Adressierung. Die LAN-Emulation muß auch den Unterschied der Adressierung bei LAN und ATM überbrücken, da der Adreßraum bei ATM groß und hierarchisch organisiert ist, bei LANs hingegen nach IEEE 802 eine flache 48-Bit Adresse darstellt, welche die Station schon bei der Herstellung zugewiesen bekommt. Auch müssen Endgeräte, die via ATM kommunizieren möchten, automatisch eine ATM-Verbindung erhalten.

Einen umfassenden Lösungsvorschlag, bei dem alle ATM-Vorteile zur Geltung kommen, bietet die „LAN-Emulation over ATM Specification Version 1.0“ [New95] des ATM-Forums. Die LAN-Emulation unterstützt alle Protokolle, und es wird definiert, wie Ethernet- oder Token-Ring-Stationen mit ATM-Stationen kommunizieren können und wie Stationen gleicher Netzzugangstechnik dies auch untereinander tun können.

Emulation von Broadcast. Die Verbindungen, welche bei dem ATM Forum User-Network Interface specification 3.0/3.1 festgelegt wurden, enthalten sowohl Punkt-zu-Punkt wie auch Punkt-zu-Mehrpunkt Verbindungen. Eine Verbindung kann so gesetzt werden, daß eine Station mit allen anderen Hosts bzw. mit einigen anderen Hosts im emulierten LAN (ELAN) verbunden ist. Auf diese Weise kann die Broadcast- bzw. Multicast-Möglichkeit emuliert werden. Über ATM ist der Multicast-Verkehr sogar viel effizienter, weil nur die adressierten Stationen angesprochen werden und nicht wie beim LAN der ganze Verkehr auf das Netz gelegt wird und dann von nicht angesprochenen Stationen herausgefiltert werden muß. Somit ergeben sich zwei Ansätze für die Emulation von Broadcasts.

Der eine Ansatz verbindet jeden Host mit jedem anderen Host (Abb. 63). Der Vorteil ist natürlich die Ausfallsicherheit auf Grund der hohen Redundanz, da bei Ausfall einer Verbindung nur ein Host betroffen ist. Der Nachteil ist, daß jedesmal, wenn ein Host dem LAN beitrtritt oder es verläßt, alle Verbindungen auf den neuesten Stand gebracht werden müssen.

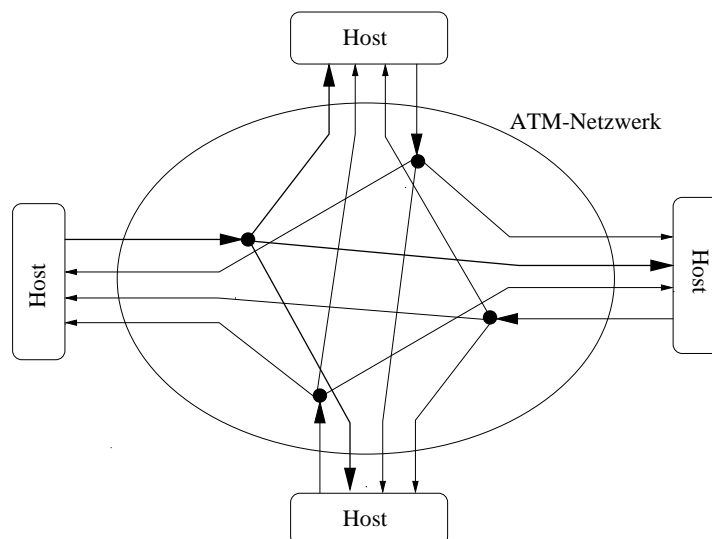


Abbildung 63. Viele Punkt-zu-Mehrpunkt-Verbindungen.

Der andere Ansatz benutzt einen zentralen Multicast-Server, welcher mit jeder anderen Station im ELAN verbunden ist (Abb. 64). Somit brauchen jeweils nur zwei Verbindungen geändert zu werden, wenn sich die Mitgliederzahl des LANs ändert. Wird ein Broadcast von einem Host gewünscht, so sendet dieser sein Paket an den Multicast-Server, der dieses wiederum weiterleitet. Damit der Server nicht zum Flaschenhals wird, kann man diesen verteilt über das Netzwerk implementieren oder ihn replizieren.

1.2 Das Client/Server-Modell der LAN-Emulation

Das Client/Server-Modell besteht aus zwei Komponenten

- dem LAN-Emulation Client (LEC)
- und dem LAN-Emulation Service.

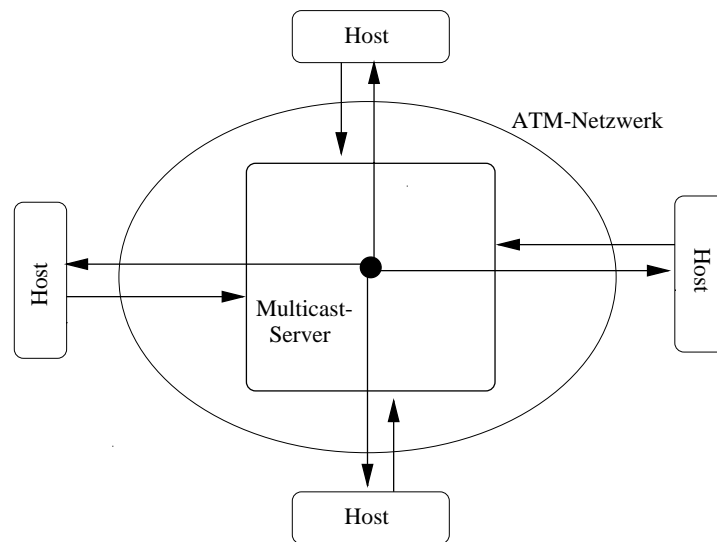


Abbildung 64. Eine Punkt-zu-Mehrpunkt-Verbindung.

Der LAN-Emulation Client. Die LEC-Funktionalität wird in der Hardware oder als Teil der Treibersoftware implementiert und befindet sich entweder im ATM-Adapter des Endgeräts oder im LAN/ATM-Switch. Die Hauptaufgabe des LEC ist die Adreßauflösung, d.h. die Zuordnung von MAC-Zieladresse zu ATM-Zieladresse.

Der LAN-Emulation Service. Der LAN-Emulation Service setzt sich aus drei Server-Bestandteilen zusammen:

- dem LAN-Emulation Configuration Server (LECS)
- dem LAN-Emulation Server (LES)
- und dem Broadcast and Unknown Server (BUS),

wobei nicht festgelegt ist, wo die Funktionen zu implementieren sind. D.h. der LAN-Emulation Service könnte eine Applikation sein, oder aber auch verteilt über das Netzwerk implementiert werden. Hier im Client-/Server Modell werden die Server als zentrale Multicast-Server implementiert, was dem ersten Ansatz entspricht, der weiter oben in Abschnitt 1.1 beschrieben ist. Es gibt im ELAN einen logischen LECS, LES und BUS.

Da die LAN-Emulation auf OSI-Ebene 2 angesiedelt ist, ist sie transparent für höhere Protokolle und kann somit routbare Protokolle wie TCP/IP oder nicht routbare Protokolle wie Netbios bedienen.

Zwischen den LECs und den Servern gibt es zwei Arten der Verbindung. Einmal existiert die Kontrollverbindung, die für administrative Zwecke eingesetzt wird, wie z.B. das Auffinden der Adresse eines anderen Client. Zum anderen gibt es für die Datenübertragung spezielle Datenverbindungen.

Der LAN-Emulation Configuration Server (LECS). Der LECS unterhält die Datenbank von Konfigurationsinformation für jedes ELAN. Wird ein LEC initialisiert, baut

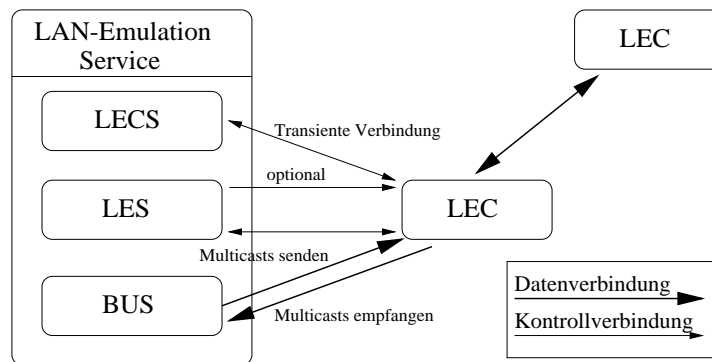


Abbildung 65. Komponenten der LAN-Emulation.

er eine Verbindung zum LECS auf, um die Adresse des LES zu erhalten. Die Datenbank in dem LECS wird üblicherweise vom Netzwerkverwalter initialisiert und über SNMP-Verwaltungsapplikationen kontrolliert.

Der LAN-Emulation Server (LES). Der LES ist verantwortlich für die Adreßauflösung im ELAN. Wenn ein LEC ein Packet verschicken will und weder eine Verbindung noch eine ATM-Adresse hat, so sendet er dem LES ein LE_ARP (LAN-Emulation Adress Resolution Protocol) auf der Kontrollverbindung, um die ATM-Adresse zu erhalten, welche mit der Ziel-MAC-Adresse assoziiert ist. Der LES kann einen Cache besitzen, damit er zu einer MAC-Adresse schnell die ATM-Adresse finden kann und dann direkt dem LEC antworten.

Kann der LES die Adresse nicht finden, so gibt er den LE_ARP auf seinen Kontrollverbindungen an alle Clients weiter. Wenn der LEC ein LE_ARP empfängt, überprüft er seine Tabelle, ob er mit der MAC-Adresse gemeint ist. Ist dies der Fall, so sendet er eine LE_ARP Rückantwort (reply back) an den LES, der diese wiederum an den ursprünglichen LEC weitergibt.

Broadcast and Unknown Server (BUS). Der BUS ist der Multicast-Server für ein ELAN. Wenn ein LEC dem ELAN beitrifft, wird eine Punkt-zu-Punkt Verbindung von jedem LEC zum BUS aufgebaut. Pakete, die der LEC per Rundruf sendet, werden zum BUS gesendet. Dieser sendet sie zurück zu allen LECs über eine Punkt-zu-Mehrpunkt Verbindung (Vergleiche dazu Abbildung 63).

Es gibt potentiell zwei Pfade, nämlich via den BUS und via die direkte Verbindung, über die Pakete zum gewünschten LEC gelangen können. D.h. während der Client eine direkte Verbindung aufbaut, kann er schon über den BUS Daten versenden, bis er dann die restliche Daten über die direkte Verbindung schicken kann. Damit bei solchen Umschaltvorgängen keine Pakete doppelt verschickt werden oder gar verloren gehen und auch die Reihenfolge erhalten bleibt, definiert die LAN Emulation ein „Flush-Protokoll“. Vor dem Umschalten auf die neue Verbindung versendet der LEC auf der gerade benutzten Verbindung eine „Flush-Nachricht“. Wenn der Empfänger ein „Flush“ zurücksendet, so weiß der Sender, daß er alle Daten vor der „Flush-Nachricht“ erhalten hat, und kann nun auf die neue Verbindung umschalten.

1.3 Funktionsweise der LAN-Emulation

Die Funktionsweise der LAN-Emulation läßt sich in fünf Schritten erklären:

1. Initialisierung
2. Konfiguration
3. Eintritt in den LAN (Join Phase)
4. Registrierung und BUS-Initialisierung
5. Datentransfer

Initialisierung. Zuerst muß der Client Mitglied eines emulierten LANs werden, wenn er Daten versenden möchte. Dazu benötigt er die Adresse des zuständigen LES. Die Spezifikation sieht hierfür mehrere Möglichkeiten vor. Der Client erhält die Adresse vom LECS, dessen Adresse er aus eine Tabelle im Switch bekommen hat. Wenn der LECS nicht erreichbar ist, wird in der Switch-Tabelle nach einem anderen LECS oder gleich nach einem LES gesucht. Hat der Client immer noch keinen Erfolg, verbindet er sich mittels einer „Well-known ATM Address“ mit dem LECS, der für jedes ATM-Netzwerk nach LUNI spezifiziert ist.

Eine andere Möglichkeit ist eine permanente virtuelle Verbindung zum BUS über einen „Well-known Virtual Path Identifier/Virtual Channel Identifier (VPI/VCI)“.

Konfiguration. Die Konfigurationsphase bereitet den Client zum Eintritt in das ELAN vor. Hierzu teilt er dem LECS seine MAC-Adresse, seine ATM-Adresse und den gewünschten LAN-Typ (z.B. Ethernet oder Token-Ring) sowie die „Maximum Frame Size“ (MFS) mit. Der Client erhält im Gegenzug die gewünschte LES-Adresse und die Bestätigung des LAN-Typs und der MFS.

Eintritt. Der LEC baut nun eine bidirektionale Kontrollverbindung zum LES und sendet ein „Join Request“. Dieses beinhaltet seine ATM-Adresse, den LAN-Typ und die MFS. Zudem teilt er mittels Multicast-Adressen mit, welchen Gruppen er angehört und ob er Daten mit nicht registrierten Zieladressen empfangen möchte.

Der LES seinerseits sendet ein „Join Response“ zurück, welches den Client zum ELAN zuläßt oder ihn abweist.

Registrierung und BUS-Initialisierung. Nachdem der Client die BUS-ATM Adresse vom LES erhalten hat, baut er eine unidirektionale Verbindung zum BUS auf, der diesen in seine Punkt-zu-Mehrpunkt-Verbindung integriert.

Datentransfer. Nun kann der eigentliche Datentransfer stattfinden. Will ein LEC Daten versenden, prüft er zuerst, ob er die zur MAC-Zieladresse korrespondierende ATM-Zieladresse schon kennt und ob bereits eine virtuelle Verbindung besteht. Kennt er die ATM-Zieladresse nicht, so sendet er ein LE-ARP an den LES. Während dieser Zeit kann der Client schon Daten über den BUS aussenden, bis er die ATM-Adresse geliefert bekommt und dann eine direkte Datenverbindung zur Zielstation aufbauen kann. Damit es bei diesen Umschaltvorgängen zu keiner Inkonsistenz der Daten kommt, findet das „Flush-Protokoll“ Verwendung, das weiter oben schon vorgestellt wurde.

2 Ziele der LAN Emulation Client Management Specification

2.1 Migration und Leistungssteigerung

Damit vorhandene Hard- und Software weiter benutzt werden können sind Migrationswege gefragt, die die ATM-Netzwerktechnik mit den „normalen“ LANs benutzbar machen, d.h. also, daß auf Rechnern bestehende Programme ohne Änderung in ATM-Netze übernommen werden können.

Dies fördert die Akzeptanz von ATM und dadurch ist es möglich, trotz alter Hard- und Software von den Vorzügen von ATM zu profitieren, beispielsweise von der höheren Bandbreite, des besseren Netzwerkmanagements, der einfacheren Skalierbarkeit und der Möglichkeit, über die selbe physikalische Schnittstelle mehrere virtuelle LANs zu betreiben. Nicht zu vergessen ist außerdem die Multimediafähigkeit aufgrund von Isochronität und Zeittransparenz.

2.2 Transparenz

Der Lösungsansatz, wie er in der LAN-Emulation Client Management Specification festgelegt wird, sieht die LAN-Emulation auf der MAC-Schicht (Schicht 2 im OSI Protokoll-stack) vor, so daß Transparenz für die höherliegenden Protokolle und die Anwendungen gewährleistet ist und somit auf höherer Ebene keine Änderungen notwendig sind (s. Abb. 66). Somit kann die LAN-Emulation auch routbare Protokolle wie TCP/IP oder nicht routbare Protokolle wie Netbios bedienen.

2.3 Verbindungen und Adressierung

Da die heutigen Netze üblicherweise verbindungslos arbeiten, aber bei ATM für jeden Sendewunsch eine dedizierte Verbindung aufgebaut wird, muß es eine Schnittstelle geben, die diesen Unterschied ausgleicht. D.h. mit anderen Worten, daß das Verhalten eines geteilten Mediums LAN emuliert werden muß.

Der Unterschied zwischen beiden Technologien macht sich auch in der Adressierung bemerkbar. Nach IEEE 802 hat man bei LANs eine „flache“ 48-Bit Adresse, die jeder Station bei der Herstellung fest zugewiesen wird. Im Gegensatz dazu ist der Adreßraum

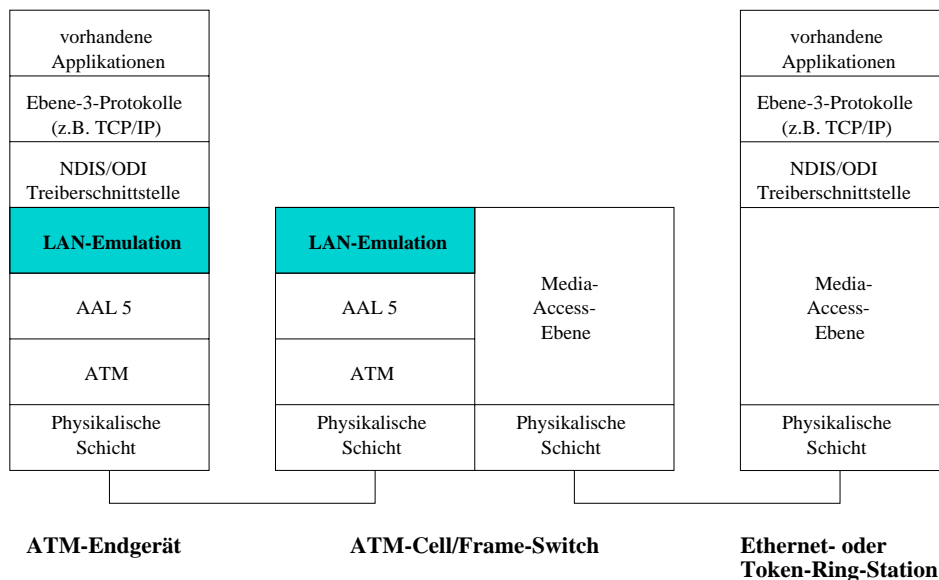


Abbildung 66. ATM-Protokoll Stack.

bei ATM groß und hierarchisch, wobei die Adresse der Station von dem Verteiler oder Netzwerk, an das sie angehängt wurde, abgeleitet wird. Somit sind hier zwei Hauptziele in der Abbildung von:

- LAN-Verbindungen auf ATM-Verbindungen
- und von MAC-Adressen auf ATM-Adressen.

3 LAN-Emulation Client Management

3.1 Einführung

Aufgrund der in einem System unterschiedlichen enthaltenen Ressourcen wurde der Begriff des Managed Objects eingeführt, um eine einheitliche Beschreibung zu erhalten. Dabei umfaßt die Management Information Base (MIB) alle in einem System enthaltenen Managed Objects mit ihren Attributen und Operationen. Sie wird von den am Management beteiligten Instanzen gemeinsam unterhalten und über Managementprotokolle angesprochen und verändert.

3.2 Management Funktionen

Netzwerkmanagement kann man in die Gebiete

- Konfigurationsmanagement,
- Leistungsmanagement,
- Fehlermanagement,
- Sicherheitsmanagement,

- und Abrechnungsmanagement

unterteilen, wobei die letzten beiden nicht Inhalt der LAN-Emulation Spezifikation sind.

Konfigurationsmanagement. Das Konfigurationsmanagement befaßt sich mit dem Einrichten der Netzwerk-Elemente für normale Operationen. Dabei versucht es, die aktuellen operationellen Parameter herauszufinden und Netzwerk-Topologien sowie Erweiterungen festzuhalten.

Diese Netzwerkmanagement-Spezifikation spricht mehrere Aufgaben an:

- Alle der aktuellen LECs Identifizieren.
- LECs Erzeugen und Zerstören für alle Hosts, welche dies unterstützen.
- Das Zulassen und Verweigern des Zutritts eines LEC in einen ELAN. (Wird ein LEC nicht zugelassen, so wird er wieder in seinen Initialzustand versetzt, in dem seine Konfiguration untersucht und verändert werden kann.)
- Die Untersuchung und Veränderung der initialen Zustandsparameter.
- Die Identifikation von Konfigurations-, Kontroll- und Multicast-VCCs.

Leistungsmanagement. Das Leistungsmanagement befaßt sich mit der Qualität und der Effizienz, mit denen der Dienst dem Benutzer geleistet wird.

Eine wichtige Komponente des Leistungsmanagement ist die Möglichkeit, das Verhalten des Netzwerkes zu beobachten. Mehrere Faktoren machen ELANs schwerer zu beobachten als traditionelle LANs:

- Der Datenverkehr ist nicht in einem physikalischen Netzwerksegment konzentriert, sondern über viele VCCs verteilt.
- Diese VCCs können beliebig häufig gesetzt oder abgebrochen werden und erschweren somit eine Beobachtung.
- Die Leistung von jedem VCC kann von Faktoren betroffen sein, die außerhalb der Kontrolle der Emulation liegen.
- Zu dem Datenverkehr kommt auch noch der LAN-Emulation Kontrollverkehr, der besser separat untersucht werden sollte.
- Manager können nicht so einfach Protokollanalyatoren hinzufügen, um herauszufinden, was passiert.

Unter diesen Umständen gibt es nur einen Weg für ein Netzwerkmanagement, um den Datenverkehr zu einem bestimmten Host zu untersuchen:

1. Die LECs müssen helfen diese Informationen zu sammeln.

2. Sammeln und Aggregieren von Statistiken über individuelle VCCs.
3. Sammeln von Leistungs-Statistiken auf ATM-Port-Ebene mit Hilfe der AToM MIB.
4. Abhören der Kommunikation zwischen zwei LAN-Emulation Hosts.

Diese Spezifikationen erfordert von den LECs die Durchführung von grundsätzlichen Datenverkehrsmessungen. Zudem müssen diese über die SNMP MIBs zugänglich gemacht werden.

Fehlermanagement. Das Fehlermanagement befaßt sich mit der Verhütung, Entdeckung und Korrektur von Fehlern in einem ELAN, welche durch das Versagen der Netzwerkelemente verursacht werden.

Auf der LEC Ebene gibt es deshalb folgende Objekte:

- Das MIB-II Objekt „ifOperStatus“ zeigt an, ob die emulierte Schnittstelle in Aktion ist oder nicht.
- Das LEC MIB Objekt „lecInterfaceState“ liefert zusätzliche Information über den Status des LEC.
- Die LEC MIB Objekte „lecLastFailureRespCode“ und „lecLastFailureState“ liefern Informationen zu dem letzten Konfigurationsfehler oder Beitrittsfehler.

3.3 Managementrahmen

Jeder LAN-Emulation Client sollte die LAN-Emulation Client MIB implementieren, um durch das Netzwerkmanagement verwaltbar zu sein. Dies geschieht entweder über das SNMP Netzwerkmanagementprotokoll oder über ein anderes Protokoll. Alle SNMP Agenten, welche LECs unterstützen, müssen die MIB-II implementieren, welche die Grundmenge der Managed Objects für die Protokolle der Internet Suite beinhaltet. Sie sollten auch den RFC 1573 implementieren, der die Evolution der Schnittstellengruppen der MIB-II festlegt.

Das Ziel der LAN-Emulation ist, ein Modell von einem oder mehreren ATM-Ports zu präsentieren, die als einer oder mehrere 802.* LAN-Ports behandelt werden können. Echte LAN-Ports haben Einträge in der MIB-II Schnittstellen-Tabelle und in der RFC 1573 Schnittstellen-Erweiterungstabelle. Um einen konsistenten Managementrahmen zu gewährleisten, ist es für diese emulierten Ports empfehlenswert, auch Einträge in diesen Tabellen vorzusehen.

3.4 LAN-Emulation Client Adressen

Da die LAN-Emulation eine Brücke zwischen der IEEE 802.*- und der ATM-Welt schafft, haben LECs MAC- und ATM-Adressen, die gespeichert und konfiguriert werden müssen. Für die MAC-Adressen benutzt man Objekte, die in der MIB-II und RFC 1573 definiert sind. Da auch jeder LEC mehrere ATM-Adressen haben kann, die verschiedenen oder

dem gleichen ATM-Port entsprechen können, kann man nicht die Objekte in der MIB-II und RFC 1573 benutzen, um die ATM-Adressen zu speichern, da sie schon für die MAC-Adressen belegt sind. Deshalb werden die Adressen in der LAN-Emulation Client MIB plaziert.

3.5 Beschreibung der LAN-Emulation Client MIB

MIB Organisation. Die LAN-Emulation Client MIB ist in einer Anzahl von Gruppen organisiert, von denen jede mit einer Tabelle korrespondiert. Diese enthalten:

1. Schnittstellen Gruppen
 - (a) Emulierte LAN-Gruppe
Konfiguration und Status
 - (b) Statistik-Gruppe
Zähler für den LAN Emulation Kontrollverkehr und SVCs
 - (c) Server Verbindungs-Gruppe
Identifiziert Kontroll- und Multicast VCCs
2. ATM Adressen-Gruppe
3. Registrierte LAN Ziel-Gruppen
 - (a) MAC Adressen-Gruppe
 - (b) Routendeskriptoren-Gruppe
4. LE_ARP Cache-Gruppe
 - (a) MAC Adreßübersetzungen
 - (b) Routendeskriptorübersetzungen

In Abb. 67 ist die Organisationsstruktur noch einmal in einer Übersicht dargestellt.

MIB Gruppen. Dieser Abschnitt beschreibt kurz jede Gruppe in der LEC MIB.

1. Schnittstellen-Gruppen

Emulierte LAN-Gruppe: Diese obligatorische Gruppe besteht aus einer Tabelle mit einer Zeile für jeden LEC. Sie ist eine leichte Erweiterung der MIB-II/RFC 1573 Schnittstellen-Tabelle. Viele der Objekte in dieser Tabelle korrespondieren mit den initialen Zustandsparametern in der LAN-Emulation Spezifikation.

Statistik-Gruppe: Diese Gruppe ist obligatorisch und besteht aus einer Tabelle, welche die Statistik für den Datenverkehr für alle LECs enthält (eine Zeile für jeden LEC).

LAN Emulation Client Management Information Base

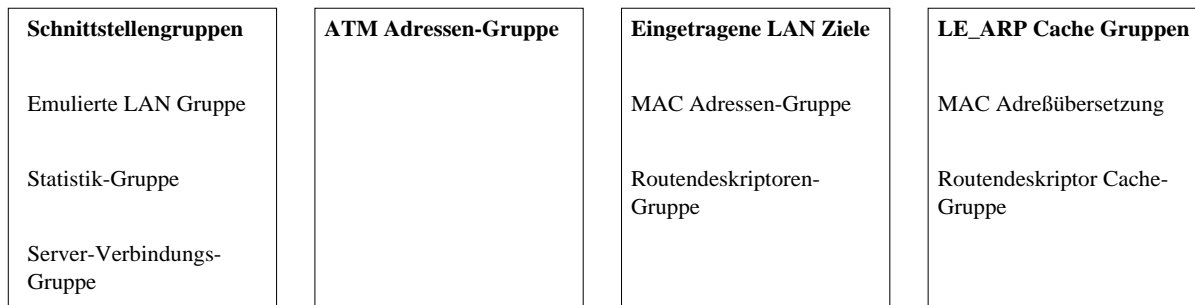


Abbildung 67. MIB Organisation.

Server Verbindungs-Gruppe: Diese Gruppe ist auch obligatorisch und besteht aus einer Tabelle. Jede Zeile in dieser Tabelle beschreibt die Kontroll-VCCs und Multicast-VCCs für einen LEC.

2. ATM Adressen-Gruppe

ATM Adressen-Gruppe: Diese obligatorische Gruppe listet alle ATM-Adressen für dessen Host LECs auf. Jede Zeile in der Tabelle beschreibt eine der ATM Adressen des Clients.

3. Registrierte LAN Ziel-Gruppen

MAC Adressen-Gruppe: Diese obligatorische Gruppe besteht aus einer Tabelle, die Einträge für alle eingetragenen MAC-Adressen hat, die zu einem Agenten des LEC gehören. Dabei beschreibt jede Zeile der Tabelle ein Paar (MAC Adresse und ATM Adresse).

Routendeskriptoren-Gruppe: Sie ist obligatorisch für Hosts, die emulierte IEEE 802.5 Token-Ring-LANs unterstützen, und optional für Hosts, die das nicht tun. Die Gruppe besteht aus einer Tabelle, die alle Routendeskriptoren, welche von dem LEC des Hosts registriert sind, auflistet. Jede Zeile enthält ein Paar, den Routendeskriptor und die ATM-Adresse, das für einen bestimmten Client eingetragen ist.

4. LE_ARP Cache-Gruppe

MAC Adreßübersetzungen: Durch diese obligatorische Gruppe wird der Zugang zu dem LAN-Emulation Clients MAC-zu-ATM-ARP-Cache ermöglicht. Die Tabelle enthält Einträge für unicast- und broadcast-Adressen, aber nicht für multicast-Adressen. Der LE_ARP Cache selbst ist eine Tabelle von Einträgen, die einer MAC-Adresse eine ATM-Adresse zuordnet.

Routendeskriptor Cache-Gruppe: Diese Gruppe ermöglicht den Zugang zu den Routendeskriptor-zu-ATM ARP Cache des IEEE 802.5 LAN-Emulation Client.

Die Implementierung dieser Tabelle ist optional für Hosts, die keinen emulierten IEEE 802.5 Token-Ring unterstützen. Ein ATM-LAN-Emulation ARP Cache-Eintrag enthält Informationen über die Zuordnung eines IEEE 802.5 Routenskriptors zu einer ATM-Adresse.

Beispiele für einige „Managed Objects“.

„*lecAtmAddressTable*“: Die ATM-Adressentabelle listet alle ATM-Adressen auf, die zu den LECs dieses Hosts gehören. Jeder Eintrag („*LecAtmAddressEntry*“) beschreibt eine der ATM-Adressen des Clients.

„*lecMacAddressTable*“: Die MAC-Adressentabelle enthält alle Einträge der eingetragenen MAC-Adressen, die zu den LECs dieses Agenten gehören. Jeder Eintrag („*LecMacAddressEntry*“) besteht aus der ATM- und MAC-Adresse für einen bestimmten Client.

„*leArpTable*“: Diese Tabelle sorgt für den Zugang zu dem MAC-zu-ATM ARP Cache eines LECs, wobei ein Eintrag neben einer MAC-Adresse – die auch eine Broadcast-Adresse sein kann – und einigen Zusatzinformationen, aus der ATM-Adresse des BUS oder aus der ATM-Adresse des LECs, besteht.

4 Möglichkeiten der Management Information Base

Da die Einträge der Schnittstellen-Tabelle für jeden LEC nicht alles abdecken, was ein Netzwerkmanager wissen können sollte, kann man mit dem LEC-MIB

- initiale Zustandsparameter untersuchen, wie den LAN Namen und die LES-ATM-Adresse für jeden operierenden/aktiven Client. Das macht es für den Netzwerkmanager leichter, andere Clients zu lokalisieren.
- Monitor-Statistiken für den LAN-Emulation Kontrollverkehr und SVC-Fehler erstellen. Dies ermöglicht die Meldung, falls ein LEC keine LE_ARPs empfangen kann oder von ihnen überflutet wird.
- LE_ARP-Caches untersuchen.

Optionale Möglichkeiten der LEC-MIB sind

- das Erzeugen und Zerstören von LECs.
- das Einrichten der initialen Zustandsparameter des Clients.
- das Erzeugen und Zerstören der LE_ARP Cache Einträgen.

Mittels der Managementfunktionen kann man auch einfach definieren, welche Endgeräte einem emulierten LAN angehören sollen. Die Switches stellen dabei sicher, daß nur Endgeräte des gleichen virtuellen Netzes Daten austauschen können und Broadcasts innerhalb dieses virtuellen Netzes bleiben.

5 Abkürzungsverzeichnis

AAL	ATM Adaption Layer
ATM	Asynchroner Transfermodus
BUS	Broadcast- and Unknown Server
ELAN	Emuliertes LAN
LAN	Local Area Network
LE	LAN Emulation
LE_ARP	LAN Emulation Address Resolution Protocol
LEC	LAN Emulation Client
LECID	LAN Emulation Client Identifier
LECS	LAN Emulation Configuration Server
LES	LAN Emulation Server
LUNI	LAN Emulation User-Network Interface
MAC	Medium Access Control
MFS	Maximum Frame Size
RFC	Request For Comment (Dokumentenreihe)
SNMP	Simple Network Management Protocol
SVC	Switched Virtual Circuit
UNI	User-Network Interface
VCC	Virtual Channel Connection
VCI	Virtual Channel Identifier
VPI	Virtual Path Identifier

Abbildungsverzeichnis

1 Ein Modell eines Transportdienstes.	20
2 OSI95 Definition des Sendedurchsatzes	21
3 Klassischer 4-Primitiven-Austausch	21
4 Dreiecksverhandlung zum Informationsaustausch	22
5 Dreiecks Aushandlung eines Vertragswertes	23
6 Betroffener Bereich im Transferstatus	27
7 Zeitdiagramm einer a) erfolgreichen und b),c) nicht erfolgreichen Neuverhandlung	27
8 Eine Folge von T-UNITDATA Primitiven	28
9 Einige request/response Szenarien	29
10 QOS-Parameter auf den unterschiedlichen Systemebenen	33
11 Ressourcenmanagement bei multimedialen Anwendungen	34
12 Der QOS-Broker innerhalb eines Kommunikationssystems	37
13 Aufbau einer QOS-Broker-Instanz	37
14 Die AQOSM-Architektur	39
15 Das QOS-A Schichtenmodell	42
16 Verhandlungsvorgehensweise von QOS-A	43
17 Vergleich der drei Ansätze	45
18 Einordnung des SMT in das ISO/OSI-Schichtenmodell.	48
19 Interaktion der Management-Komponenten.	50
20 SMT Frameformat.	51
21 SMT Management Model	52
22 Allgemeines CCE-Interface für einen Port/MAC	58
23 Die Tenet Real-Time Protocol Suite.	65
24 Ohne Multicast-Unterstützung.	68
25 Mit Multicast-Unterstützung.	69
26 Architektur eines GSM-Netzes	77
27 geographische Anordnung der Funkzellen	78
28 mögliche Netzkonfigurationen	79
29 Frequenzaufteilung	81
30 Mehrfachzugriff auf Zeit- und Frequenzebene	82
31 Burst	82
32 Kodierung von Nicht-Sprachdaten	88
33 Mobiles Netzwerk.	95
34 Basis-Schema.	96

35	Triangle-Routing.	97
36	Redirection-Tunneling.	99
37	Patron-Host-Konzept.	100
38	Erneuerung der Forwarding-Pointer.	101
39	Mobility-Support-Router-Konzept.	103
40	IMHP-Tunneling-Header.	106
41	Kommunikations-Formen	109
42	Netzauslastung	110
43	MBone und Tunneln	111
44	Beispielnetz und Spannender Baum	113
45	RPF Bäume einer Gruppe mit verschiedenen Quellen	114
46	Pruning eines Baumes	115
47	CBT Baum	117
48	Beitreten des Host H zum MC-Baum	121
49	Virtuelle Pfade und Kanäle	141
50	IP-Adreßklasse	144
51	Zugriffszeiten bei dynamischen Auswahlstrategien	152
52	Entfernungen zu WWW-Servern im Internet	154
53	Erklärung der <i>AVG</i> -Funktion	154
54	Paketumlaufzeiten zu WWW-Servern im Internet	155
55	Einteilung der Techniken zur Auffindung von Servern	157
56	CBT Architektur in der Anycasttechnik	162
57	Beispiel zur Triangulation	164
58	Modell eines Managementobjekts	169
59	Das ATM-Management-Modell	173
60	Grundstruktur einer ATM-Zelle	186
61	Topologische Struktur eines ATM-Netzes	187
62	Breitband-ISDN-Referenzmodell	189
63	Viele Punkt-zu-Mehrpunkt-Verbindungen.	200
64	Eine Punkt-zu-Mehrpunkt-Verbindung.	201
65	Komponenten der LAN-Emulation.	202
66	ATM-Protokoll Stack.	205
67	MIB Organisation.	209

Tabellenverzeichnis

1 Zusammenfassung aller Dienste	7
2 Verhalten des Betreibers in Abhängigkeit der Dienstqualitäts-Semantik .	24
3 Das Verhalten des Betreibers in Abhängigkeit der QoS Semantik	26
4 ISO/OSI Transportprotokollklassen.	63
5 Entfernungen und Koordinaten zu Bild 57	165
6 Zusammenfassender Vergleich der vorgestellten Techniken	166

Literatur

- [AB94] E. Hoffmann, O. Knauer, A. Badach. *High Speed Internetworking*. Addison-Wesley, 1994.
- [AC95] P. Alexander und K. Carpenter. ATM Net Management: A Status Report. *Data Communications International* **24**(12), September 1995, Seite 110 – 116. The Network Technology Magazine of the McGraw-Hill Companies.
- [ADL94] Oliver Bonaventure, André Danthine und Guy Leduc. The QoS Enhancement in OSI95. *Research Reports ESPRIT* Band 1, Jul 1994, Seite 125–151.
- [All95] A. Alles. ATM Internetworking. Technischer Bericht, Cisco Systems, Inc., Mai 1995.
- [AT94a] M. Ahmed und K. Tesink. Definitions of Managed Objects for ATM Management version 8.0. RFC 1695, IETF, August 1994. ATOM MIB.
- [AT94b] M. Ahmed und K. Tesink. Definitions of Managed Objects for ATM Management Version 8.0 using SMIV2. Request for Comments 1695, Network Working Group, August 1994.
- [ATM94a] ATM Forum. *ATM Network Management*, Oktober 1994. Vortragsfolien.
- [ATM94b] ATM Forum. *Customer Network Management (CNM) for ATM Public Network Service (M3 Specification)*, Ausgabe 1.04, Oktober 1994. ATM Forum 94-0342.
- [ATM94c] ATM Forum. *M4 Interface Requirements and Logical MIB: ATM Network Element View*, Ausgabe 1.0, Oktober 1994. ATM Forum 94-0388R4.
- [Bag94] Yves Baguette. The OSI95 Connectionless-Mode Transport Services. *Research Reports ESPRIT* Band 1, Jul 1994, Seite 181–198.
- [Bal95] T. Ballardie. *Core Based Tree (CBT) Multicast*. Internet-Draft, Oktober 1995.
- [BFC93] T. Ballardie, P. Francis und J. Crowcroft. Core Based Trees (CBT). *Computer Communication Review* **23**(4), Oktober 1993.
- [BFMM94] A. Banerjee, D. Ferrari, B.A. Mah und Mark Moran. The Tenet Real-Time Protocol Suite: Design, Implementation, and Experiences. Tr-94-059, The Tenet Group, November 1994.
- [Boo95] Booz. *Mobilfunk - Vom Statussymbol zum Wirtschaftsfaktor*. Booz, Allen & Hamilton, 1995.
- [BS95] L. Breslau und S. Shenker. A Proposal for Accomodating Heterogeneity. Internet-Draft, Internet Engineering Task Force, „draft-ietf-intserv-hetero-00.txt“, November 1995.
- [Byn] M. Byne. An Introduction to IP Multicast. <http://ganges.cs.tcd.ie/4ba2/multicast>.
- [CA93] Garcia F., Cambell A., Coulson G. Integrated Quality of Service for Multimedia Communications. *IEEE INFOCOM '93*, March 1993.

- [CA94] Hutchinson D. Cambell A., Coulson G. A Quality of Service Architecture. *ACM SIGCOMM '94* **24**(2), April 1994.
- [CC95] M. E. Crovella und R. L. Carter. *Dynamic Server Selection in the Internet*. Third IEEE Workshop on High Performance Communication Subsystems (HPCS'95), Mystic, Connecticut, USA, August 1995.
- [Cid92] I. Cidon. A Critique of ATM from Data Communications Perspective. *Journal of High Speed Networks* **Vol. 1**(No. 4), 1992, Seite 315–336.
- [CM95] Gihwan Cho und Lindsay F. Marshall. An Efficient Location and Routing Scheme for Mobile Computing Environment. *IEEE Journal on selected Areas in Communications* **13**(5), Januar 1995.
- [CMRW93] J. Case, K. McCloghrie, M. Rose und S. Waldbusser. Structure of Management Information for Version 2 of the Simple Network Management Protocol (SNMPv2). Request for Comments 1442, Network Working Group, April 1993.
- [Dan94] André Danthine. The Networking Environment of the Nineties and the Need of New Standards. *Research Reports ESPRIT* Band 1, Jul 1994, Seite 1–13.
- [Dei95] J. Deibert. Stand der Dinge. *Gateway 5/95*, 1995, Seite 94–97.
- [Epp95] K. Eppele. Verbindungslose Verbindung. *Gateway — Magazin für Daten- und Telekommunikation* (4), April 1995, Seite 114–117.
- [FBJ⁺93] W. Fuhrmann, V. Brass, U. Janßen, F. Kühl und W. Roth. Digitale Mobilkommunikationsnetze. *Kommunikation in Verteilten Systemen*, März 1993.
- [Fis] Kay Uwe Fischer. Multicast Backbone - Gruppenkommunikation im Internet. Seminar Netzwerkmanagement und Hochgeschwindigkeits-Kommunikation im WS 1994/95.
- [FJM⁺95] S. Floyd, V. Jacobson, S. McCanne, C. Liu und L. Zhang. A Reliable Multicast Framework for Light-weight Sessions an Application Level Framing. In *Proceedings of ACM SIGCOMM '95*. Cambridge, Massachusetts, 1995.
- [GM95] M. Germuska und G. Morgan. Comparison of XTP with TCP and TP4. *Transfer* **8**(3), Mai/Juni 1995.
- [Gro95] E. Gronert. Datenübertragung im GSM. *Funkschau* (16), 1995, Seite 48–51.
- [GS95] J. D. Guyton und M. F. Schwartz. *Locating Nearby Copies of Replicated Internet Servers*. Proceedings of ACM SIGCOMM'95, Boston, Massachusetts, USA, August 1995.
- [HA93] H.-G. Hegering und S. Abeck. *Integriertes Netz- und Systemmanagement*, Kapitel OSI-Netzmanagement, Seite 129 – 155. Addison-Wesley Publishing Company. 1993.
- [Har95] P. Harrschar. Kommunikationsunterstützung für multimediale Anwendungen. Diplomarbeit, Institut für Telematik, Universität Karlsruhe, Juli 1995.
- [HAU83] J. E. Hopcroft, A. V. Aho und J. D. Ullman. *Data Structures and Algorithms*. Addison-Wesley. 1983.

- [HD95] R. Hinden und S. Deering. *IP Version 6 Addressing Architecture*. Internet-Draft, Juni 1995.
- [Hof94] Markus Hofmann. Zuverlässige Gruppenkommunikation in heterogenen Netzen. In *Diplomarbeit am Institut für Telematik*. Karlsruhe, 1994.
- [Hof96] Markus Hofmann. A Generic Concept for Large-Scale Multicast. In *Proceedings of International Zurich Seminar on Digital Communications*. Zurich, 1996.
- [HSC95] H. Holbrook, S. Singhal und D. Cheriton. Log-Based Receiver-Reliable Multicast for Distributed Interactive Simulation. In *Proceedings of ACM SIGCOMM '95*. Cambridge, Massachusetts, 1995.
- [Hui95] Christian Huitema. *Routing in the Internet*. Prentice Hall. 1995.
- [HW95] M. Hochmuth und F. Wildenhain. *ATM-Netze — Architektur und Funktionsweise*. Nr. 10 der Reihe „Thomson's Aktuelle Tutorien“ (TAT). International Thomson Publishing, Bonn. Februar 1995.
- [IETFa] Internet Engineering Task Force. IETF-Homepage. WWW-Seite. <http://www.ietf.cnri.reston.va.us/home.html>.
- [IETFb] Internet Engineering Task Force. IETF Overview. WWW-Seite. <ftp://ds.internic.net/ietf/1ietf-description.txt>.
- [IETFc] Internet Engineering Task Force. Integrated Services (intserv). WWW-Seite. <http://www.ietf.cnri.reston.va.us/html.charters/intserv-charter.html>.
- [Inc95] ZeitNet Inc. ATM LAN Emulation in Workgroup Networks, 1995.
- [Kau94] D. Kauer. Durch den Äther. *iX* (3), 1994, Seite 150–154.
- [KM95] R. Kosak und T. Mersevole. ATM as the Well Managed Protocol. ATM Forum Newsletter, Vol. 3, Issue 3, Juli 1995. Aus der WWW-Seite des ATM-Forums.
- [Kön93] K. Köntopp. Einheitliche Sicht. *c't* 3/93, März 1993, Seite 232–236.
- [Krü95] G. Krüger. Praktikum Telematik I + II. Praktikumsunterlagen, Sommersemester 1995. Institut für Telematik, Universität Karlsruhe.
- [Krü95] G. Krüger. Telematik I. Unterlagen zur Vorlesung, Wintersemester 1994/95. Institut für Telematik, Universität Karlsruhe.
- [Lau94] M. Laubach. IP Over ATM. *ConneXions* 7/94, 1994, Seite 20–25.
- [LKK93] P.C. Lockemann, G. Krüger und H. Krumm. *Telekommunikation und Datenhaltung*. Hanser Studienbücher der Informatik. ISBN 3-446-17465-6, 1993.
- [Mey95a] E. Meyer. GSM-World-Congress: Mit GSM ins 21. Jahrhundert. *Funkschau* (11), 1995, Seite 50–51.
- [Mey95b] E. Meyer. PCS in USA mit europäischem Standard? *Funkschau* (3), 1995, Seite 52–55.
- [Mil95] W. Milliken. Host Anycasting, Oktober 1995. E-Mail an Markus Hofmann.

- [MJP95] Andrew Myles, David B. Johnson und Charles Perkins. A Mobile Host Protocol Supporting Route Optimization and Authentication. *IEEE Journal on Selected areas in communications* **13**(5), Juni 1995.
- [MMU95] MMU. D2 - Privat. *Handbuch von Mannesmann Mobilfunk GmbH* (1), Jan 1995.
- [MR92] G. Milligan und F. Ross (Hrsg.). *FDDI Station Management (SMT)*. Draft Proposed American National Standard. ANSI. Dec. 1992.
- [New95] T. Newton. LAN Emulation Client Management: DRAFT Straw Ballot Specification, Februar 1995.
- [NK95a] Smith J. M. Nahrstedt K. The QoS Broker. *IEEE Multimedia* **2**(1), Spring 1995.
- [NK95b] Steinmetz R. Nahrstedt K. Resource Management in Networked Multimedia Systems. *IEEE Computer*, May 1995.
- [PC93] David M. Piscitello und A. Lyman Chapin. Introduction to Routing. *ConneXions* **7**(9), September 1993.
- [Per95] Charles Perkins. IP Mobility Support draft-ietf-mobileip-protocol-12. August 1995.
- [PMM93] C. Partridge, T. Mendez und W. Milliken. *Host Anycasting Service*. RFC 1546, November 1993.
- [Ram95] R. Ramanathan. Multicast Support for Nimrod: Requirements and Solution Approaches. Februar 1995. Internet-Draft.
- [Sch95] H. Schulzrinne. *ATM: Dangerous AT Any Speed?*, Band 9, No. 4. IEEE Network. 1995.
- [Sei94] Jochen Seitz. *Netzwerkmanagement*, Band 2. International Thomson Publishing. Thomson's aktuelle Tutorien, 1994.
- [SH95] C. Schmidt und M. Hofmann. Hochleistungskommunikation. Folienvorlagen zur Vorlesung, Sommersemester 1995. Institut für Telematik, Universität Karlsruhe.
- [SP95a] S. Shenker und C. Partridge. Specification of Guaranteed Quality of Service. Internet-Draft, Internet Engineering Task Force, „draft-ietf-intserv-guaranteed-svc-01.txt“, Juni 1995.
- [SP95b] S. Shenker und C. Partridge. Specification of Predictive Quality of Service. Internet-Draft, Internet Engineering Task Force, „draft-ietf-intserv-predictive-svc-01.txt“, Mai 1995.
- [SPW95] S. Shenker, C. Partridge und J. Wroclawsky. Specification of Controlled Delay Quality of Service. Internet-Draft, Internet Engineering Task Force, „draft-ietf-intserv-control-del-svc-01.txt“, Juli 1995.
- [SSK94] P. Sanoz, K. Sabani und D. Kristol. Multicast Transportprotocols for High Speed Networks. In *Proceedings of International Conference on Networkprotocols*. Boston, Massachusetts, 1994.

- [Ste94] J. Sterbenz. Protocols for high Speed Networks: Life After ATM? In *Protocols for High Speed Networks IV*, Seite 3–18. G. Neufeld and M. Ito. eds., Chapman & Hall, London, 1994.
- [SW] W.T. Strayer und A.C. Weaver. Is XTP Suitable for Distributed Real-Time Systems? Internal Report. University of Virginia, Charlottesville Virginia.
- [TW94] Horlait E. Tawbi W., Fladenmuller A. End-System QoS Management of Multimedia Applications. In *Architecture and Protocols for High-Speed Networks*. Kluwer Academic Publishers, 1994.
- [WD93] B. Walke und P. Decker. Mobile Datenkommunikation - Eine Übersicht. *it+ti - Oldenbourg Verlag* (5), 1993, Seite 12–25.
- [Wea94] A.C. Weaver. Xpress Transport Protocol Version 4. *Transfer* 7(6), November/Dezember 1994.
- [Wro95a] J. Wroclawsky. Specification of General Characterization Parameters. Internet-Draft, Internet Engineering Task Force, „draft-ietf-intserv-charac-00.txt“, November 1995.
- [Wro95b] J. Wroclawsky. Specification of the Controlled-Load Network Element Service. Internet-Draft, Internet Engineering Task Force, „draft-ietf-intserv-ctrl-load-svc-00.txt“, November 1995.
- [YBD94] Guy Leduc Yves Baguette, Luc Léonard und André Danthine. The OSI95 Connection-Mode Transport Service. *Research Reports ESPRIT* Band 1, Jul 1994, Seite 212–225.
- [YGS95] R. Yavatkar, J. Griffioen und M. Sudan. A Reliable Dissemination Protocol for Interactive Collaborative Applications. In *Proceedings of ACM Multimedia '95*. 1995.
- [Yua95] Ruixi Yuan. Traffic pattern based mobile Routing Scheme. *Computer Communications* 18(1), Januar 1995.
- [ZS95] M. Zitterbart und C. Schmidt. *Internetworking: Bruecken, Router und Co*. Thomson's Aktuelle Tutorien. Thomson Publishing. 1995.