

Verteilte Multimedia-Anwendungen und ihre Anforderungen an die Netzinfrastruktur

Thomas R. Schmidt

Institut für Nachrichtenvermittlung und Datenverarbeitung

Pfaffenwaldring 47

70569 Stuttgart

Tel. 0711/685-7965, Fax 0711/685-7983, schmidt@ind.uni-stuttgart.de

Zusammenfassung: Multimedia-Anwendungen stellen sehr unterschiedliche Anforderungen an Netze und Architekturen. Bei der Übertragung von Multimediadaten zwischen Rechnern muß die Anwendung versuchen, die vom Benutzer gewünschte Dienstqualität mit den momentan zur Verfügung stehenden Mitteln möglichst gut zu erreichen. Hierzu ist eine Konvertierung von Parametern zwischen verschiedenen Ebenen nötig, z. B. aus der Benutzersicht in die Netzsicht.

Ziel dieses Projekts ist die Entwicklung einer Multimedia-Architektur, die diese Dienstqualität möglichst effizient bereitstellt. Hierfür ist es notwendig, Aussagen über die Zusammenhänge zwischen Anforderungen und Parametern auf den verschiedenen Ebenen der Architektur und bessere Strategien der Reaktion auf Veränderungen in der beobachteten Dienstgüte zu finden. Dies soll einerseits durch Modellierung und Simulation, andererseits durch experimentelle Untersuchung verschiedener Techniken geschehen.

1. Einleitung

Durch die Steigerung der Leistungsfähigkeit der digitalen Übertragungsnetze wurde es möglich, neue verteilte Anwendungen zu realisieren. So können z. B. in einem Unternehmen mit Hilfe von *Computer Supported Collaborative Work* (CSCW) Ingenieure an verschiedenen Standorten an einem gemeinsamen Projekt arbeiten, oder es kann eine Vorlesung digital in viele Hörsäle übertragen werden.

Diese Anwendungen nennt man *Verteilte Multimedia-Systeme*. Sie zeichnen sich durch die integrierte, computergesteuerte Verwaltung unabhängiger, diskreter und kontinuierlicher Medien aus [6]. Z. B. stellen sich im Falle der Übertragung einer Vorlesung folgende Anforderungen an die Anwendung: Das Bild des Vortragenden muß synchron zu seiner Stimme abgespielt und eventuelle Hilfsmittel wie Folien oder Demonstrationen von allen Hörern optisch erkannt werden können. Andererseits ist es keine große Beeinträchtigung der Leistung, wenn zwischen den Ausspielzeitpunkten in verschiedenen Hörsälen eine Verzögerung von einer Sekunde liegt.

Somit können aus der Art der Anwendung Anforderungen wie Bildqualität, Verzögerung und Synchronität abgeleitet werden. Um diese Anforderungen zu erfüllen, muß die Anwendung sie auf die ihr zur Verfügung stehenden Mittel abbilden. Je nach verwendeter Technik stehen nur eine begrenzte Anzahl von z. B. Bildformaten oder Übertragungsmethoden zur Verfügung, aus denen ausgewählt werden kann. Aus diesen muß die Anwendung diejenigen auswählen, die zusammen den gewünschten Dienst – unter Berücksichtigung von Rahmenbedingungen wie Kosten oder Benutzerwünsche – am besten erbringen.

Handelt es sich beim Übertragungsmedium zudem um ein Netz mit veränderlicher Dienstgüte, so muß die Anwendung außerdem auf Veränderungen reagieren können, indem sie ihre Parameter an die neue Situation anpaßt, bzw. den Benutzer darüber informiert, daß sich die gewünschte Dienstqualität zur Zeit nicht erbringen läßt.

Im Rahmen dieser Arbeit soll nun eine Architektur definiert werden, in der die verschiedenen Abbildungsfunktionen lokalisiert, ihre Parameter beschrieben und in der Mechanismen zur Anpassung an Veränderungen definiert werden. Durch theoretische und experimentelle Untersuchungen soll ihre Funktionsweise verstanden bzw. verifiziert werden.

2. Dienstgüte – Quality of Service

Beim Aufbau einer Multimedia-Verbindung muß eine Zuteilung von Ressourcen stattfinden, um eine bestimmte Dienstgüte initial zu erreichen. Die gewünschte Dienstgüte wird dabei vom Benutzer vorgegeben (z. B. „Fernsehqualität“, oder auf einer Skala von 1 bis 10), bzw. ist durch die Anwendung festgelegt (z. B. „Dokumentenübertragung“) und muß nun mit den verfügbaren Mitteln (Hardware, Software, Netze) realisiert werden. Ein weiterer wichtiger Aspekt ist die Transparenz für den Benutzer. Besonders wenn er für den Dienst bezahlen muß, will er über die erbrachte Dienstqualität und den Ressourcenverbrauch informiert werden.

Um eine vorgegebene Dienstgüte zur Benutzung zur Verfügung zu stellen, müssen aus ihr Anforderungen an die verwendete Hardware, Software und Netzinfrastruktur abgeleitet werden. So bedeutet der Wunsch nach „Fernsehqualität“, daß Bild und Ton der Übertragung nicht mehr als 80 ms [3] gegeneinander verschoben sein dürfen, während bei einer „Dokumentenübertragung“ der Ton keine große Rolle spielt, jedoch das Bild sehr detailreich sein muß.

Diese Anforderungen werden von der verwendeten Hard- und Software in dieser Form nicht verstanden. Somit müssen sie zunächst transformiert werden, man spricht von „QoS-Mapping“. In diese Funktionen fließt sehr viel Erfahrungswissen, um die Zuordnung von Parametern zu Anforderungen in sinnvoller Weise durchführen zu können. Um dieses Wissen zu erhalten, wurden viele Studien unternommen, deren Ergebnisse nun in eine Umsetzung einfließen sollen.

Es reicht aber nicht aus, diese Zuordnung in einer Richtung (von den Anforderungen zu den Parametern) vorzunehmen. In den heutigen Datenübertragungsnetzen kann sich die Dienstgüte mit der Zeit ändern. Als Beispiel sei hier das Internet erwähnt, dessen Leistung

stark variieren kann. Da aber auch das Anwendungssystem seine Leistung verändern kann, z. B. kann die Bildrate erniedrigt werden, besteht nun die Möglichkeit, die Anwendung an die veränderte Netzsituation dergestalt anzupassen, daß die gewünschte Dienstgüte noch zu einem gewissen Maß erreicht wird.

Für diese Anpassung der Anwendung an das Netz muß nun eine Umkehrabbildung der Parameter auf die Anforderungen erfolgen, um dann zu entscheiden, welche der Anforderungen abgeschwächt werden kann. Aus den neuen Anforderungen werden dann neue Netzparameter abgeleitet, die die augenblickliche Situation besser berücksichtigen.

Beispiel. Der Benutzer einer Anwendung gibt als Wunsch „Fernsehqualität“ ein. Dies entspricht hier einer Rate von 25 Bildern/s bei einer Auflösung von 720 x 576 Punkten und einem Stereoton mit 32 kHz und 16 bit Auflösung; Bild und Ton dürfen maximal 80ms voneinander abweichen. Die Anwendung muß nun zunächst eine Konvertierung auf die ihr zur Verfügung stehenden Ressourcen vornehmen: die verwendete Hardware und die implementierten Kodierer können nur 20 Bilder/s bei 640 x 480 Bildpunkten und Stereoton mit 24 kHz, 16 bit liefern. Durch die Kodierung reduziert sich die zu übertragende Datenmenge auf 3 Mbit/s für Video und 350 kbit/s für Audio, die zeitliche Abweichung kann bis zu 40 ms betragen.

Als nächstes muß nun vom Netzwerk diese Übertragungskapazität angefordert werden. Hier sei nun das Netz nur in der Lage, 2.5 Mbit/s insgesamt zu übertragen. Somit muß die Anwendung diese Kapazität den beiden Medienströmen so zuweisen, daß sich beide nicht zu stark verschlechtern. Die Hardware sei lediglich in der Lage, die Bildauflösung auf 320 x 240 Punkte zu reduzieren; somit benötigt die Videoübertragung nur noch etwa 0.75 Mbit/s und die Audioqualität muß nicht verringert werden.

An dieser Stelle muß auch der Benutzer darüber informiert werden, daß sein Wunsch nach „Fernsehqualität“ zur Zeit nicht erfüllt werden kann. Er hat nun die Wahl, entweder die reduzierte Qualität zu akzeptieren oder die Übertragung ganz abzubrechen. Hier entscheide sich der Benutzer für ersteres.

Nach einiger Zeit steht im Netz für die Anwendung 3.2 Mbit/s Kapazität zur Verfügung.

Somit kann die Anwendung auf die höhere Bildauflösung umschalten, wenn sie zugleich die Audioqualität reduziert, d. h. die Frequenz auf 18 kHz einstellt.

Abb. 1 zeigt den Ablauf dieser Dienstgütereinstellung und -regelung: Der Benutzer übergibt einen Wunsch an die Anwendung, diese fordert vom Netz¹ die Ressourcen an, die sie benötigt. Das Netz teilt der Anwendung mit, welche Ressourcen es zur Verfügung stellen kann und die Anwendung vergleicht dies mit der damit erbringbaren Qualität. Da nur eine schlechtere Dienstqualität erbracht werden kann, fragt die Anwendung zunächst beim Benutzer nach, der sie jedoch akzeptiert. Somit kann die Anwendung die Ressourcen belegen und mit der Übertragung beginnen.

Nach einiger Zeit stellt das Netz nun eine Veränderung in der erbringbaren Dienstgüte fest, teilt dies der Anwendung mit, die daraufhin ihre Ressourcenbelegung verändert. Da hier die vom Benutzer gewünschte Dienstqualität besser erreicht wird, muß keine Bestätigung eingeholt werden; eine Mitteilung reicht aus.

1. Hier ist als Netz der Zugang zur Übertragungsschicht gemeint.

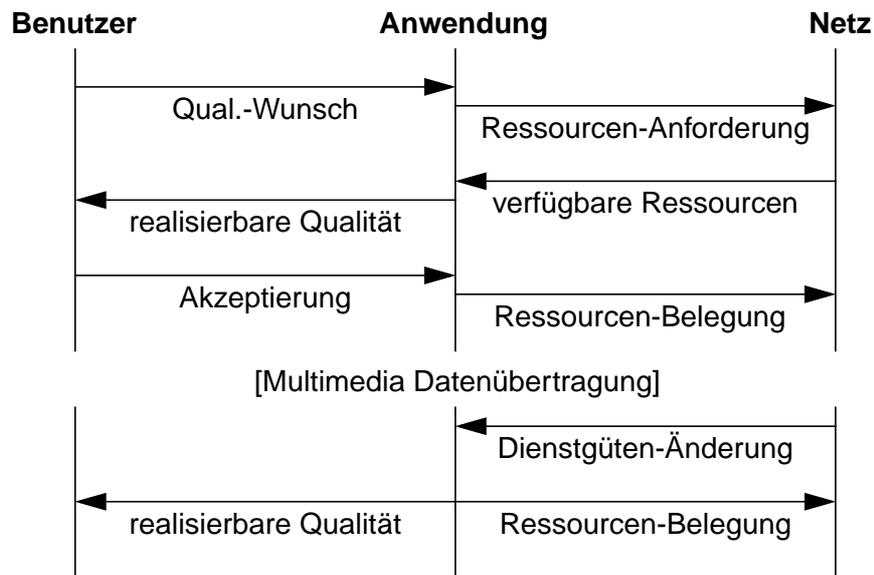


Abb. 1: Ablauf der Dienstgüteneinstellung

3. Ziele der Arbeit

Um die Zuordnung zwischen Benutzer-Qualität und Parametern der Anwendung zu beschreiben, müssen Studien mit möglichst vielen Testpersonen durchgeführt werden, die ihr subjektives Empfinden einer Parametereinstellung mitteilen. Dies wurde an verschiedenen Stellen dokumentiert (z. B. [4]).

Hier soll nun aber verstärkt die Auswirkung der Parameterauswahl zwischen Anwendung und Netz auf die beobachtete Dienstgüte untersucht werden. Es sollen dabei die Parameter identifiziert werden, die den größten Einfluß auf die Qualität der Übertragung haben; als sinnvolle Ansatzpunkte für neue Entwicklungen und Strategien. Diese Untersuchung wird quantitativ vorgenommen, da nur so eine Gewichtung der Bedeutung der Parameter möglich ist.

Sobald eine Gewichtung der Parameter gefunden wurde, kann ein Modell einer Multimedia-Übertragung entwickelt werden, das sich auf die wesentlichen Aspekte konzentriert und so mit vertretbarem Aufwand Ergebnisse liefern kann. Diese theoretischen Ergebnisse werden dann mit Hilfe einer Experimentierumgebung/Architektur verifiziert, neue Strategien implementiert und ihre Funktion im Versuch bewertet.

4. Vorgehensweise

Im Rahmen des Projekts ist die erste Phase die Auswertung der bisher durchgeführten Untersuchungen. Hierbei ist besonders die Verwendung existierender Daten über die quantitativen Zusammenhänge von Anforderungen und Parametern zwischen Benutzer und Anwendung zu nennen. Solche Untersuchungen benötigen eine große Anzahl von Testpersonen, die verschiedene Zuordnungen subjektiv bewerten, und sind somit sehr aufwendig.

Eine qualitative Beschreibung der Anforderungen und bekannter Strategien auf den verschiedenen Ebenen einer Multimedia-Architektur wird benötigt, um im weiteren Verlauf Hinweise für zu untersuchende Parameter zu erlangen. Hier sollen zunächst durch Simulation quantitative Aussagen über Zusammenhänge zwischen Parametern und Dienstgüte gefunden werden, die eine Gewichtung der Parameter erlauben.

Die in der Simulation gewonnenen Aussagen sollen durch Experimente verifiziert werden, wobei ebenfalls das Simulationsmodell kritisch auf seine Realitätsnähe zu untersuchen ist. Anschließend kann eine Untersuchung der Leistungsfähigkeit bekannter Protokolle vorgenommen und aus ihnen ggf. bessere Strategien zur Sicherstellung der Dienstgüte entwickelt, simuliert und im Experiment verifiziert werden.

4.1 Simulation

Beispiel. In einer Multimedia-Anwendung werden sowohl ein Audio- als auch ein Videostrom übertragen. Um Sprache lippensynchron zum Videobild eines Sprechers zu übertragen, darf die zeitliche Verschiebung („Skew“) zwischen diesen Strömen nicht mehr als 80 ms betragen.

Es wird nun ein Simulationsmodell entworfen (Abb. 2), in dem die Quellen, das Übertragungsnetz, die Empfangspuffer und die Datensenke modelliert sind. An den Puffern kann anhand von Zeitstempeln in den Datenpaketen gemessen werden, wie groß der Skew zwischen ihnen ist.

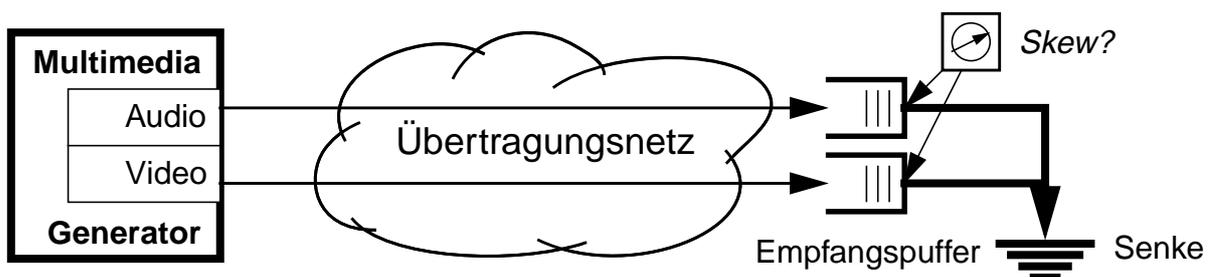


Abb. 2: Skew-Simulation

Für den Aufbau solcher Simulationen steht am Institut für Nachrichtenvermittlung und Datenverarbeitung (IND) bereits eine umfangreiche Bibliothek („SimLib“) verkehrstheoretischer Modellkomponenten, wie Warteschlangen, Bedieneinheiten usw., zur Verfügung, durch die die Modellierung eines Multimedia-Systems sehr erleichtert wird.

Diese Simulationen werden in einem weiteren Schritt verfeinert, indem u. a. das Übertragungsnetz genauer betrachtet wird (Abb. 3). So besteht es normalerweise aus mehreren seriellen Warteschlangen und Bedieneinheiten; in einem IP-Netz sind dies die Router, die den zu übertragenden Verkehr verzögern und außerdem Datenpakete anderer Übertragungen auf denselben Pfaden vermitteln. Im Modell kann dies durch zusätzlichen Verkehr in die Puffer und abgehenden Verkehr aus den Bedieneinheiten modelliert werden.

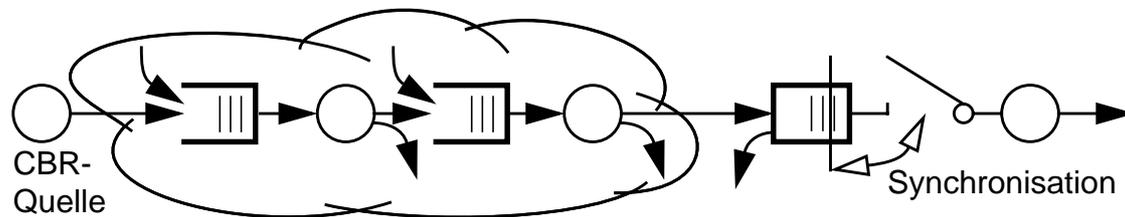


Abb. 3: Modellierung des Übertragungsnetzes

Hier ist es nun möglich, in einer Parameterstudie die Auswirkungen auf die Gleichlaufkorrektur zu untersuchen. Um vor allem Audiodaten mit einer hohen Qualität auszugeben, muß das abspielende Gerät die Verzögerungsschwankungen der Übertragung ausgleichen, indem es empfangene Daten nicht sofort ausspielt, sondern zunächst in einem (endlichen) Puffer zwischenspeichert. Je nach Größe der Varianz („Jitter“) wird sich der Puffer füllen bzw. leeren. Nun kann untersucht werden, unter welchen Umständen Daten aufgrund eines vollen Puffers verworfen werden, bzw. durch einen leeren Puffer eine hörbare Pause oder ein Knacksen entstehen würde. Als weiterer Schritt können dann Protokolle zur Verhinderung dieser Zustände untersucht werden, z. B. das „Adaptive Synchronization Protocol“ (ASP, [5]).

Protokolle wie ASP verwenden Informationen, die in einem Endgerät über die Übertragung erhoben werden können, um Parameter der Übertragung zu beeinflussen. Hier wird anhand der Füllstände der Empfangspuffer entschieden, ob die Rate der Datenentnahme (z. B. die Bildrate bei Videodaten) verändert werden muß, um den Füllstand wieder in den gewünschten Bereich zu steuern. Gleichzeitig wird hier eine Verringerung des Skews zwischen verschiedenen Ausspielpunkten angestrebt, die in verteilten Multimedia-Systemen an verschiedenen Stellen im Netz sitzen können. Um dieses Verhalten zu untersuchen, muß das Modell des Systems um eine Abbildung der Protokolle erweitert werden.

4.2 Experimentelle Verifikation

Die durch die Simulation gewonnenen Aussagen über Zusammenhänge zwischen Parametern und Dienstgüte, sowie neue Methoden zur Sicherstellung derselben, sollen durch Experimente verifiziert werden. Hierzu entsteht im „Netzlabor“ des IND eine Meßumgebung, die z. B. eine „Experimentierplattform“ enthält. Sie ermöglicht es, durch gezielte und streng lokalisierte Veränderungen einer Multimedia-Anwendung (z. B. den Austausch einer Prioritätsstrategie zur Aufteilung von Betriebsmitteln) schnell Aussagen über deren Auswirkungen zu erhalten.

Das Netzlabor ist eine Einrichtung des IND, in dem neben PC auch Netzinfrastruktur für Lehr-, Demonstrations- und Forschungszwecke installiert wurde. Es gibt ein ATM-Netz mit

Vermittlungsknoten (Switches), Ethernet-Segmente können über eine Steckschalttafel einfach eingerichtet werden und es gibt die Möglichkeit, Rechner und andere Endgeräte per ISDN zu verbinden. Außerdem können Geräte zur Überwachung und Beeinflussung der Datenübertragung eingesetzt werden, namentlich ATM-, LAN- und ISDN-Tester.

Mit dieser Infrastruktur können existierende Multimedia-Anwendungen untersucht werden. Sie eignet sich aber auch für die Entwicklung neuer Anwendungen, wobei besonders die modulare Multimedia-Experimentierplattform zu nennen ist, an der zur Zeit gearbeitet wird. Dabei handelt es sich um eine Spezifikation und Schnittstellendefinition, mit der es möglich wird, Funktionen in verschiedenen Modulen unterschiedlich zu implementieren, den Rest der Anwendung jedoch zu übernehmen. Hier wird es also möglich sein, gezielt an einer sehr begrenzten Stelle Veränderungen vorzunehmen, ohne daß dies Auswirkungen auf den Rest der Anwendung hat. Man kann dann davon ausgehen, daß beobachtete Veränderungen des Verhaltens des Systems primär durch die Änderung verursacht wurden. Dagegen ist es sonst meist nicht möglich, nur einen kleinen Teil einer Anwendung zu ändern, wodurch nicht nur beabsichtigte Änderung vorgenommen werden müssen und so die Aussagekraft der Ergebnisse beschränkt ist.

5. Zusammenfassung

Verteilte Multimedia-Systeme sollen für ihre Benutzer einen Dienst bestimmter Qualität erbringen. Da jedoch die heute verwendeten digitalen Übertragungsnetze entweder keine Garantie für die von ihnen erbrachte Leistung geben (z.B. IP) oder die Qualität ihrer Leistung mit den Kosten zusammenhängt (z.B. ATM), muß eine Anwendung darauf bedacht sein, die verfügbaren Ressourcen bestmöglich zu nutzen und ggf. auf Änderungen der beobachteten Netzgüte zu reagieren.

Somit ist es Aufgabe der Anwendung, die Wünsche des Benutzers auf Parameter der Hard- und Software abzubilden und sie so gut wie möglich zu erfüllen. Hierzu ist die Kenntnis der Zusammenhänge zwischen Anforderungen und Parameter nötig und es müssen Strategien bekannt sein, um auf Veränderungen der einen mit Anpassung der anderen zu reagieren. Dieser Zusammenhang soll im Rahmen einer Architektur untersucht, das Verhalten und die Leistungsfähigkeit verschiedener Ansätze sollen simuliert und neue Ansätze implementiert und experimentell verifiziert werden.

6. Verwandte Arbeiten

Im Rahmen des Graduiertenkollegs Parallele und Verteilte Systeme wurde von Walter Fiederer am Institut für Parallele und Verteilte Höchstleistungsrechner (IPVR) der Universität Stuttgart bereits das „Dienstgüte- und Ressourcenmanagement in verteilten Multimedia-Systemen“ untersucht [1]. Ergebnisse flossen u. a. in das „Extended Negotiation and Resource Reservation Protocol“ (XNRP, [2]) sowie das Systemkonzept „CINEMA“ [7] ein.

Betrachtungen zu QoS-Mapping und Scheduling-Strategien finden sich u. a. in Daniel Bauers Dissertation [8]. Ebenfalls an der ETH Zürich entwickelt wurde das System „Da CaPo“, das die erbrachte Dienstqualität überwacht und ggf. versucht, die gewünschte Leistung durch Rekonfiguration zu erbringen. Daneben gibt es viele weitere QoS-Architekturen; eine Übersicht kann z. B. [9] entnommen werden.

Die Auswirkungen der Netzwerkdienstgüte auf die erbrachte Qualität einer Multimedia-Übertragung wurde bereits für einige spezielle Fälle untersucht. So wurden z. B. in [10] die Auswirkungen von ATM-Zellverlusten auf eine Videoubertragung betrachtet und ein Simulationsmodell einer Multimedia-Applikation entwickelt. Versuche, durch Glättung eines MPEG-kodierten Video-Stroms die Verfügbarkeit einer bestimmten Dienstqualität zu erhöhen, wurden in [11] gemacht.

7. Literatur

- [1] Walter Fiederer: „*Dienstgüte- und Ressourcenmanagement in verteilten Multimedia-Systemen*“, 4. Berichtskolloquium des GK PVS, 1997.
- [2] G. Dermler, W. Fiederer, K. Rothermel: „*QoS Negotiation and Resource Reservation for Distributed Multimedia Applications*“, IEEE International Conference on Multimedia Computing and Systems, Ottawa 1997.
- [3] R. Steinmetz, C. Engler: „*Human Perception of Media Synchronization*“, IBM Technical Report no 439310, 1993.
- [4] K. Fukuda, N. Wakamiya, M. Murata, H. Miyahara: „*QoS Mapping between User's Preference and Bandwidth Control for Video Transport*“, in: A Campbell & K. Nahrstedt (Eds.), *Building QoS into Distributed Systems*, IFIP/Chapman & Hall, 1997.
- [5] K. Rothermel, T. Helbig: „*An Adaptive Stream Synchronization Protocol*“, URL <http://www.informatik.uni-stuttgart.de/ipvr/vs/Publications/Publications.html>, 1996.
- [6] K. Rothermel, S. Wirag: „*Distributed Multimedia Systems*“, IPVR Universität Stuttgart, Vorlesung, URL <http://www.informatik.uni-stuttgart.de/ipvr/vs/lehre/ss98/vorlesungen/VMMS/VMMS.html>, 1998.
- [7] W. Fiederer, G. Dermler, K. Rothermel: „*Dienstgüte-Management im Systemkonzept CINEMA*“, Informationstechnik und Technische Informatik (it+ti), August 1997.
- [8] D. Bauer: „*A Multipoint Communication Architecture for End-to-End Quality of Service Guarantees*“, Dissertation ETH Zürich No. 12163, 1997.
- [9] A. Campbell, C. Aurrecochea, L. Hauw: „*A Review of QoS Architectures*“, Proceedings of IWQoS'96, GMD-Studien Nr. 282, Sankt Augustin 1996.
- [10] M. Baumann, T. Müller, W. Ooghe, A. Santos, S. Winstanley, M. Zeller: „*Multi-layer modelling of a multimedia application*“, in: Paul J. Kühn, Roya Ulrich (Eds.), *Broadband Communications*, IFIP/Chapman & Hall, 1998.
- [11] G. Bai, B. E. Wolfinger: „*Possibilities and Limitations in Smoothing MPEG-coded Video Streams: A Measurement-based Investigation*“, in: *Messung, Modellierung und Bewertung von Rechen- und Kommunikationssystemen*, ITG/GI-Fachtagung 9, VDE-Verlag, 1997.