





Bereitstellung von Dienstgüte in IP- und ATM-Netzen als Voraussetzung für die Video-Übertragung mit Hardware Codecs

Ursula Hilgers¹, Susanne Naegele-Jackson¹,
Peter Holleczek¹, Richard Hofmann²

¹ Regionales Rechenzentrum der Universität Erlangen-Nürnberg, Martensstraße 1, 91058 Erlangen

² Universität Erlangen-Nürnberg, IMMD 7, Martensstraße 3, 91058 Erlangen

Zusammenfassung Multimedia-Anwendungen dringen zunehmend in den kommerziellen Bereich ein. Videodaten müssen daher nicht nur – wie bisher – über ATM-Strecken sondern auch über IP-Netzwerke übertragen werden. Beide Übertragungstechniken bieten für die interaktive Bildübertragung stark unterschiedliche Möglichkeiten, wobei ATM alle notwendigen Aspekte zur Realisierung von Dienstgüte im Netzwerk bereitstellt. Dem ATM-Ansatz wird die Differentiated Services Architektur gegenüber gestellt, ein Ansatz, der das IP-Protokoll um die Bereitstellung von Dienstgüte erweitert. Dieses Konzept wird mit dem ATM-Protokoll im Hinblick auf die Dienstgüteunterstützung verglichen. Da Hardware-Codecs für die interaktive Bildübertragung unerlässlich sind, werden die theoretischen Ergebnisse anhand von Messungen an Hardware Codecs veranschaulicht.



1 Einleitung

Multimedia-Übertragungen mit Video- und Audiodaten über Netzwerke erfordern zunehmend die Beachtung von Aspekten der Übertragungsqualität (Quality of Service (QoS)), denn multimedialen Anwendungen mit Echtzeitanforderungen reicht die Qualität des Best Effort Dienstes, den das IP-Protokoll bietet, nicht aus.

Im ATM-Protokoll sind verschiedene Dienstkategorien definiert, die sowohl Echtzeit als auch Best Effort Dienste umfassen [ATM96]. Damit die von den Kategorien angebotene Leistung vom Netzwerk tatsächlich erbracht wird, sind umfangreiche Funktionen zur Verkehrsregulierung, Ressourcenkontrolle und Überlastbehandlung im ATM-Protokoll implementiert.

Der Ansatz der Differentiated Services (DiffServ) erweitert das IP-Protokoll um die Möglichkeit, unterschiedliche Dienstqualitäten in Netzwerken bereitzustellen [Bla98]. IP-Pakete werden klassifiziert und zu Dienstklassen mit unterschiedlichen Qualitätscharakteristiken zusammengefasst. In den Netzwerkkomponenten werden die Pakete in Abhängigkeit von den Qualitätsanforderungen der zugehörigen Dienstklasse weitergeleitet. Die unterschiedliche Behandlung wird durch Scheduling- und Queue-Management-Verfahren, z. B. Mechanismen zur Verhinderung von Überlast, implementiert.

Diese Arbeit vergleicht die Bereitstellung von Dienstgüte im ATM- und im DiffServ-Ansatz. Die theoretischen Ergebnisse werden dabei anhand von Messungen an Hardware

Codecs veranschaulicht, da diese für interaktive Multimedia-Übertragungen unerlässlich sind.

2 Vergleich der Mechanismen

Um Pakete¹ unterschiedlichen Dienstklassen bzw. -kategorien zuordnen zu können, muss in den beteiligten Netzwerkkomponenten die Funktionalität der Klassifikation implementiert sein. Das ATM-Protokoll verwendet dazu die VP- und VC-Identifikationsnummern im Zellheader, die eine Verbindung eindeutig kennzeichnen. So wird der Verbindung für die gesamte Lebensdauer eine Dienstkategorie zugeordnet, die in einem Service Vertrag festgehalten ist. Damit kennt der ATM-Switch die zeitlichen Anforderungen bei der Weiterleitung dieser Zellen. Im DiffServ-Konzept erfolgt das Setzen des Type of Service Feldes im IP-Header zur Klassifizierung der IP-Pakete bei Eintritt in ein Netzwerk durch die Netzwerkkomponenten. Diese Klassifizierung erfordert es, in allen Netzwerkkomponenten eine konsistente Konfiguration herzustellen. Darüber hinaus konnte in [HH01] gezeigt werden, dass Interfaces von Routern mit einer Klassifizierung von IP-Paketen durch in Software realisierten Access-Listen (wie beispielsweise bei Cisco Routern der 7500 Serie) schon bei einer Anzahl von 50 Verkehrsverbindungen zu 100 % ausgelastet sein können. Normalerweise müssen Routerinterfaces aber Tausende von Verbindungen bewältigen können.

Wenn ein Konzept QoS im Netzwerk bereitstellen will, muss der Durchsatz von hochprioritären Strömen auch bei Überlast garantiert werden. Während ATM über diese Eigenschaft verfügt, stellt DiffServ nur relative Prioritäten bereit [HH01], [Hil00]. Experimente im Labor haben gezeigt, dass die im ATM-Protokoll für einen CBR VC reservierte Bandbreite auch in Überlastsituationen garantiert wird, nachdem die Zugangskontrolle (Call Admission Control) einen Verbindungswunsch akzeptiert hat und die Verbindung aufgebaut wurde. Im DiffServ-Konzept dagegen kann beispielsweise auf Überlast mit der Aktivierung des Verfahrens RED, bei dem ab einer vorher festgelegten Warteschlangenlänge Pakete nach einem Zufallsverfahren verworfen werden, reagiert werden. Dann werden aber sogar für Ströme mit hohen Prioritäten keine Durchsatzgarantien gegeben. Das ist der größte Unterschied zwischen ATM und DiffServ. Aufsetzend auf dem IP-Protokoll ist – zur Zeit – noch kein Mechanismus zur Ressourcen-Reservierung implementiert. Erweiterungen des DiffServ-Konzeptes sehen daher eine Reservierung mit dem *Resource Reservation Protocol* (RSVP) [Bra97] vor. Es hat sich allerdings herausgestellt, dass dieses Protokoll in WANs nicht skaliert [Man97].

Die unterschiedliche Behandlung von Paketen mit verschiedenen Prioritäten im Netzwerk wird durch Queuemanagement-Mechanismen, z. B. Scheduling Algorithmen, realisiert. Im ATM-Protokoll wird das Scheduling über strikte prioritätsgesteuerte Algorithmen durchgeführt. In einigen Implementationen werden die Zellen unterschiedlicher Dienstklassen in unterschiedlichen Pufferbereichen gespeichert. Die Puffer mit höherer Priorität, z. B. für die CBR und VBR Zellen, werden zuerst bedient.

Ein Scheduling-Mechanismus zur Realisierung verschiedener Dienstklassen im DiffServ-Ansatz ist das Weighted Fair Queueing (WFQ). Der Scheduling Mechanismus verwaltet

¹ Der Begriff Pakete wird hier als Überbegriff für die transportierte Dateneinheit verwendet. In ATM-Netzen entspricht sie Zellen, beim IP-Protokoll IP-Paketen.

mehrere Warteschlangen für unterschiedliche Dienstqualitäten. Jeder Warteschlange wird ein Gewicht zugeordnet, das den Anteil an ausgehender Bandbreite festlegt. Damit kann für jede Dienstklasse, der eine ausgehende Warteschlange zugeteilt wird, eine untere Grenze für den Durchsatz angegeben werden, die nicht unterschritten wird.

Wird der Verkehr einer Klasse zusätzlich durch einen Token Bucket reguliert, kann eine maximale obere Grenze für das Ende-zu-Ende-Delay angegeben werden [Par92]: Betrachtet man z.B. ein Token Bucket, das 1000 Pakete fassen kann, die im schlimmsten Fall die maximal zugelassene Paketgröße im Netzwerk annehmen (bei einem Packet over SONET Interface liegt die MTU mit Standardkonfiguration bei 4470 Byte), und wird das Bucket kontinuierlich mit einer Rate von 50 MBit/s gefüllt, dann kann sich das Delay im schlimmsten Fall, wenn alle Buckets dieses Verkehrsstroms in den Netzwerkkomponenten auf dem Weg durch das Netzwerk gefüllt sind, auf 715 ms belaufen.

3 Einsatz von Hardware Codecs im Netzwerk

Damit interaktive Videoübertragungen für den Benutzer attraktiv sind, müssen die Reaktionszeiten des Systems bestimmten Zeitanforderungen genügen. Im Folgenden werden daher die Anforderungen formuliert, die Codecs aufgrund ihres eigenen Zeitbedarfs an die Übertragungsleistung von Netzwerken stellen. Dazu wird zunächst die Verzögerung bestimmt, die die Codecs durch den Kodierungs- und Dekodierungsprozess selbst verursachen.

Die Untersuchungen werden mit zwei Geräten des Typs MAC500 für MPEG-2 (4:2:0) Komprimierung über ATM durchgeführt. Die Verzögerung, die sich durch die Kodierung bzw. Dekodierung ergibt, wird mit Hilfe eines zwei-Kanal-Speicheroszilloskops (Typ Tektronix 2220) gemessen (s. Abbildung 1). Als Eingangsquelle wird ein Videosignal verwendet, bei dem sich 25 schwarze Frames mit 25 weißen Frames in einer Endlosschleife abwechseln. Der erste Kanal des Oszilloskops wird mit der Videoquelle direkt verbunden; dem zweiten Kanal wird das kodierte und dekodierte Videosignal zugewiesen.

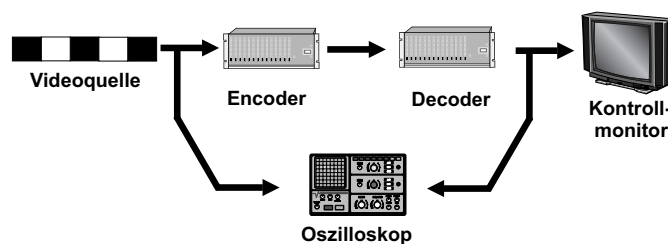


Abbildung 1. Testaufbau zur Messung der Verzögerung.

Die Messungen werden für drei verschiedene Kodierungen durchgeführt: Für eine Kodierung mit I-Frames only (Group of Picture Size = 1 (GOP)) ergibt sich ein Delay von 95 ms. Kompliziertere Algorithmen, die bei der Kompression auf bekannte Informationen

von bereits übertragenen Bildern zurückgreifen (IP Kodierung) oder zusätzlich durch Bewegungsvektoren auf zukünftige Bildinhalte schließen (IBBP Kodierung), weisen größere Delaywerte von 180 ms bis 250 ms auf. Die Ergebnisse der Tabelle 1 ergeben sich jeweils bei einer Video-Rate von 10 MBit/s im Format MPEG-2 (4:2:0).

Tabelle 1. Delaywerte bei drei Kodierungen.

Größe der GOP	Frame Encoding	Delay [ms]
1	I	95
13	IP	180
28	IBBP	250

Verbindet man dieses Wissen mit den Empfehlungen der ITU, die für die qualitativ gute Durchführung einer interaktiven Applikation ein Ende-zu-Ende Delay von kleiner 400 ms vorschlägt [ITU96a], muss das Round Trip Delay bei einer interaktiven bidirektionalen Multimedia-Übertragung, beispielsweise bei den hier betrachteten Codecs und einer I-Frames only Einstellung, kleiner als 210 ms sein, damit der Anwender eine gute Qualität erhält. Diesen zeitlichen Anforderungen werden ATM-Netzwerke gerecht, während durch die fehlende Ressourcen-Reservierung im DiffServ-Konzept Überlast in IP-Netzwerken zu unbestimmtem Zeitverhalten führt.

Im nächsten Schritt wird der Einfluss von Fehlern bei der Bildübertragung bewertet. Als Bildquelle wird eine Kamera verwendet, die auf ein sich bewegendes Metronom gerichtet ist. Das Bild wird zunächst in einer ersten Phase ohne äußere Störeinflüsse bewertet. Im zweiten Testabschnitt werden mit Hilfe eines Impairment Tools (Interwatch 95000) Fehlerraten simuliert und in die Bilddatenübertragung eingeschleift. Die Qualität des Bildes wird subjektiv bewertet.

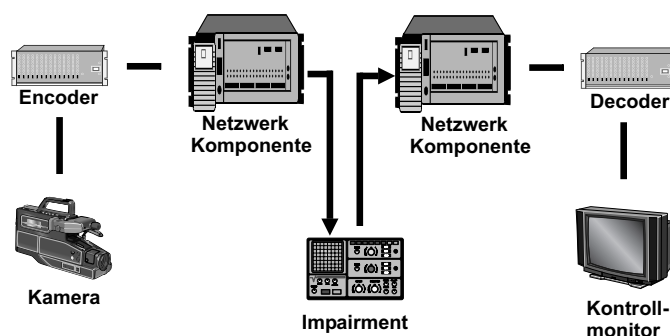


Abbildung 2. Testaufbau mit Impairment Tool.

In Abbildung 2 ist der Testaufbau dargestellt. Eine Kamera zur Aufnahme des Bildes ist mit einem Encoder verbunden, der über zwei ATM-Switches, einen LS1010 der Firma Cisco und einen ASX1000 der Firma Fore, an einem Decoder hängt mit einer zusätzlichen Verbindung zu einem Kontrollmonitor. In die ATM-Verbindung zwischen Encoder und Decoder ist der ATM-Monitor GN-Nettest IW95000 mit Impairment-Funktion eingeschleift. Dieser erzeugt Bitfehler auf der ATM-Verbindung zwischen den Switches.

Tabelle 2. Subjektive Bewertungen der Bildqualität.

Fehlerrate	subjektive Qualität
0	sehr gut
10^{-8}	sehr gut
10^{-7}	zitterndes Bild, deutliche Fehler
10^{-6}	unakzeptable Qualität, Blockfehler
10^{-5}	schwarzes Bild

In Tabelle 2 sind die Ergebnisse abgebildet. Für die subjektive Bewertung der Bildqualität wird die Bezeichnung des Mean Opinion Score (MOS) der ITU verwendet, eine Qualitäts-Skala, bei der Testpersonen die Übertragungsqualität in 5 Stufen [ITU96b] einteilen: Sehr gut, gut, ordentlich, störend und unakzeptabel. Bis zu einer Fehlerrate von 10^{-8} wird das Bild mit einer sehr guten Qualität übertragen. Für eine Fehlerrate von 10^{-7} sind schon deutliche Einbußen festzustellen, und bei einer Fehlerrate von 10^{-6} ist die Qualität nicht mehr annehmbar. Bei dem Experiment wird deutlich, dass bereits geringste Fehler auf der Übertragungsverbindung zu nicht akzeptablen Qualitätsverlusten führen.

4 Bewertung der Ergebnisse

Die Delaymessungen der Codecs haben gezeigt, dass die Verzögerungen, die im Netzwerk verursacht werden, kleiner als 210 ms sein müssen, damit noch eine für den Anwender ansprechende Qualität bei der Übertragung erreicht wird. Die Fehlerrate muss kleiner sein als 10^{-8} .

Da bereits geringste Übertragungsfehler deutliche Bildbeeinträchtigungen hervorrufen, sind hochqualitative Videoübertragungen nur dann möglich, wenn Paketverluste in Überlastsituationen in Netzwerkkomponenten vermieden werden. Dazu zählen auch Verlustraten, die sich durch ein zu großes Delay und damit eine zu späte Auslieferung eines Paketes ergeben, da Videosequenzen innerhalb strenger zeitlicher Grenzen beim Empfänger ausgespielt werden müssen.

Die beiden Konzepte, das ATM-Protokoll und der Ansatz der Differentiated Services, stellen unterschiedliche Mechanismen zur Verfügung, um die Anforderungen von multimedialen Datenübertragungen an Netzwerke zu garantieren.

Das ATM-Protokoll stellt sowohl den hochpriorien Echtzeit-Kategorien eine reservierte Bandbreite für die Übertragung zur Verfügung als auch eine garantierte Einhaltung von Zeitgrenzen. Dieses ist über ein prioritätenbasiertes Scheduling realisiert.

Der DiffServ-Ansatz dagegen kann diese Anforderungen nicht erfüllen. Zwar ermöglicht der Einsatz von Scheduling-Algorithmen wie WFQ eine garantierte obere Schranke für die zur Verfügung stehende Bandbreite pro Warteschlange [PG93], aber zur Zeit existiert noch keine Ressourcen-Reservierung mit Zugangskontrolle für IP. Auch sind die zeitlichen Zusagen, die dieses Verfahren erlaubt, für interaktive bidirektionale Applikationen nicht ausreichend. Bei dem erläuterten Beispiel liegen sie weit über dem in Abschnitt 3 angegebenen Wert von 210 ms [Sta98]. Leider sind Codecs, die Videoübertragungen über IP durchführen können, bisher noch nicht in der Lage, eine Reservierung von Bandbreite, beispielsweise über RSVP, zu initiieren [Nae01].

5 Zusammenfassung

In dieser Arbeit wird erläutert, dass IP-Netzwerke, die durch das DiffServ Konzept erweitert werden, noch nicht in der Lage sind, hochpriorien Verkehr mit garantierten Übertragungsanforderungen zu transportieren. Dies liegt zum einen an einem fehlenden Ressourcen-Reservierungsverfahren, aber auch daran, dass zeitliche Zusagen nicht in den Größenordnungen möglich sind, wie sie für interaktive Bildkommunikation notwendig sind.

Im Gegensatz zu dem DiffServ-Ansatz verfügt ATM über Ressourcen-Reservierungsmechanismen und garantierte Übertragungs-Charakteristiken und wird wohl deshalb weiterhin für Echtzeit-Dialoganwendungen die bevorzugte Kommunikationsinfrastruktur sein.

Literatur

- [ATM96] ATM Forum, *Traffic Management Specification 4.0*. The ATM Forum, Technical Committee, April 1996.
- [Bla98] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, *An Architecture for Differentiated Services*. Request for Comments 2475, Dezember 1998.
- [Bra97] B. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin, *Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification*. Request for Comments 2205, September 1997.
- [HH01] U. Hilgers, R. Hofmann, *QoS – ATM versus Differentiated Services*. In: J. Knop, P. Schirmbacher (Hrsg.), Proceedings EUNIS 2001, März 2001.
- [Hil00] U. Hilgers, R. Hofmann, P. Holleccek, *Differentiated Services – Konzepte und erste Erfahrungen*. In: Praxis der Informationsverarbeitung und Kommunikation, Februar 2000.
- [Man97] A. Mankin, F. Baker, B. Braden, S. Bradner, M. O'Dell, A. Romanow, A. Weinrib, L. Zhang, *Resource ReSerVation Protocol (RSVP) – Version 1 Applicability Statement – Some Guidelines on Deployment*. Request for Comments 2208, September 1997.
- [Nae01] S. Naegele-Jackson, U. Hilgers, P. Holleccek, *Evaluation of Codec Behavior in IP and ATM Networks*. In: J. Knop, P. Schirmbacher (Hrsg.), Proceedings EUNIS 2001, März 2001.
- [ITU96a] ITU-T Recommendation G.114, *One-Way Transmission Time*. Februar, 1996.
- [ITU96b] ITU-T Recommendation P.800, *Methods for Subjective Determination of Transmission Quality*. August, 1996.
- [PG93] A. K. Parekh, R. G. Gallager, *A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single-Node Case*. In: IEEE/ACM Transactions on Networking, Vol. 1, Nr. 3, Juni 1993.



- [Par92] A. K. J. Parekh, *A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks*. LIDS-TH-2089, MIT Laboratory for Information and Decision Systems, Cambridge, Mass., February 1992.
- [Sta98] W. Stallings, *High-Speed Networks*. Prentice Hall, 1998.

