IP-Telefonie über Differentiated Services

U. Thürmann, M. Zitterbart
Institut für Betriebssysteme und Rechnerverbund, TU Braunschweig
{thuerman|zit}@ibr.cs.tu-bs.de

Abstract— **Neben klassischen Datendiensten entwickeln sich zur Zeit eine Reihe von Echtzeitdiensten imInternet. Insbesondere hat die IP-Telefonie in den letzten Jahren an Bedeutung gewonnen. Es fehlt aber immer noch eine geeignete Infrastruktur, um diesen Anwendungen die geforderte Dienstgüte zur garantieren zu können. Der zur Zeit bei der IETF entwickelte Ansatz der Differentiated Services wird in dieser Arbeit bezüglich seiner Eignung für IP-Telefonie untersucht.**

Keywords—**IP-Telefonie, Differentiated Services, Quality-of-Service, Simulationen**

I. MOTIVATION

Das Internet wird derzeit als globale Kommunikationsinfrastruktur für verschiedenste Anwendungen gesehen. Einer der prominentesten Vertreter aus Sicht der Anwendungen ist sicherlich das WWW. Neben reinen Datenanwendungen sollen aber auch Anwendungen mit Audio- und Videokomponenten unterstützt werden, vom „einfachen" Telefonieren bis zu Tele-Konferenzen und virtuellen Welten. Hieraus resultieren Anforderungen hinsichtlich einer geeigneten Dienstgüteunterstützung für solche Anwendungen, die im Internet kaum bereitgestellt wird.

In Abbildung 1 ist ein Beispiel für eine IP-Telefonie-Anwendung dargestellt, basierend auf H.323 [1]. IP-Telefone in verschiedenen lokalen Netzen sollen über das Internet bzw. über ein IP-basiertes Intranet verbunden werden und einen vom Telefonnetz her gewohnten Telefondienst anbieten. Hierzu muß unter anderem eine Sprachqualität entsprechend derjenigen im Telefonnetz angeboten werden. Dies erfordert die Unterstützung von Dienstgüte im Intranet bzw. im Internet.

Der Beitrag untersucht den Ansatz der Differentiated Services [2] zur Dienstgüteunterstützung von IP-Telefonie und ist wie folgt strukturiert. In Abschnitt II wird zunächst das den Untersuchungen zugrunde liegende Simulationsmodell skizziert. Ergebnisse der simulativen Experimente diskutiert Abschnitt III. Abschnitt IV schließt den Beitrag mit einer Zusammenfassung und einem Ausblick auf weitere Arbeiten.

II. BASISSZENARIO

Für die simulativen Experimente wurde ein Simulationsmodell in ns-2 realisiert, das die geforderten DiffServ-Mechanismen unterstützt. Die damit aufgebauten Netztopologien bestehen aus einer Reihe von DiffServ-Domänen, an die jeweils vier lokale Netze mit Endsystemen (u.a. IP-Telefone) angeschlossen sind (siehe Abb. 2). Die DiffServ-Domänen bestehen jeweils aus einem Ring von fünf inneren Routern und fünf Border-Routern. Diese ringförmige Struktur ist typisch für große ISPs wie z.B. den DFN. Jeder Border-Router ist mit genau einem der inneren Router und mit einem Router eines benachbarten Netzes verbunden. In den Experimenten wurden in den lokalen Netzen jeweils vier IP-Telefonie-Datenströme etabliert und in Richtung des Border-Routers der nächsten Domäne weitergeleitet (B14 in Abbildung 2). Insgesamt wurde ein Netz mit drei DiffServ-Domänen simuliert (siehe Abb. 3). Die Domänen sind dabei jeweils über einen Link mit einer Datenrate von 10 Mbit/s verbunden.

In allen Routern der DiffServ-Domänen kommen DiffServ-Mechanismen zum Einsatz [2], [3], [4]. In den Border-Routern wird eine Zugangskontrolle durchgeführt. Ankommende Pakete werden vom Classifier für den ihnen zugeordneten Service markiert. Hier wird den IP-Paketen einer Telefonverbindung ein entsprechender DiffServ-Codepoint zugeordnet, der deren weitere Behandlung im Netz bestimmt. Der Traffic-Meter überprüft die von den einzelnen Service-Klassen benutzten Ressourcen und markiert gegebenfalls Pakete als out-of-profile. Für die Warteschlangen werden Class Based Queueing (CBQ) [5] und RIO-Warteschlangen [6] verwendet.

In den Endsystemen des LANs dagegen werden einfache FIFO-Warteschlangen ohne Differenzierung zwischen Paketen verschiedener Verkehrsströme verwendet. Die Router im LAN sind nicht DiffServ-fähig.

Für die Simulation des *Hintergrundverkehrs* wurden ON/OFF-Quellen mit Pareto-Verteilung benutzt [7], um selbstähnlichen Verkehr zu generieren. Jedes lokale Netz sendet und empfängt 40 dieser burstartigen Datenströme. Insgesamt werden also 480 Pareto-verteilte Hintergrundverbindungen etabliert, wobei sich davon 120 an der Stelle der höchsten Konzentration zwischen R14 und B14 überlagern. In den Simulationen mit Differentiated Services wird der Hintergrundverkehr verschiedenen Service-Klassen (Assured Service und Best Effort) zugeordnet. Es wird mehr Hintergrundverkehr erzeugt, als in den Traffic-

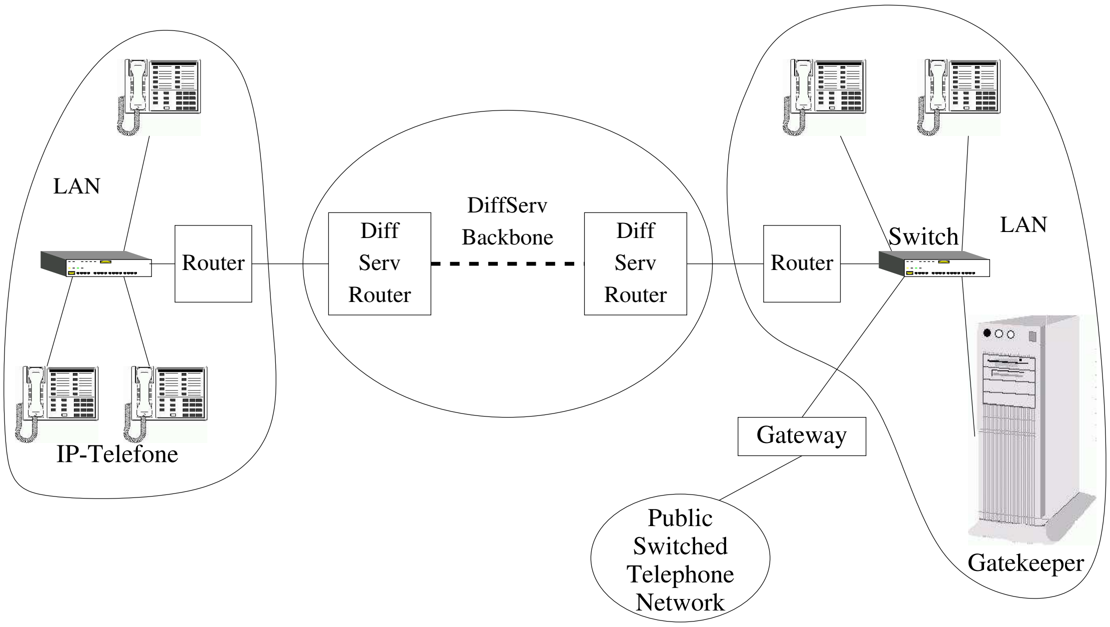


Fig. 1. IP-Telefonie über das Internet

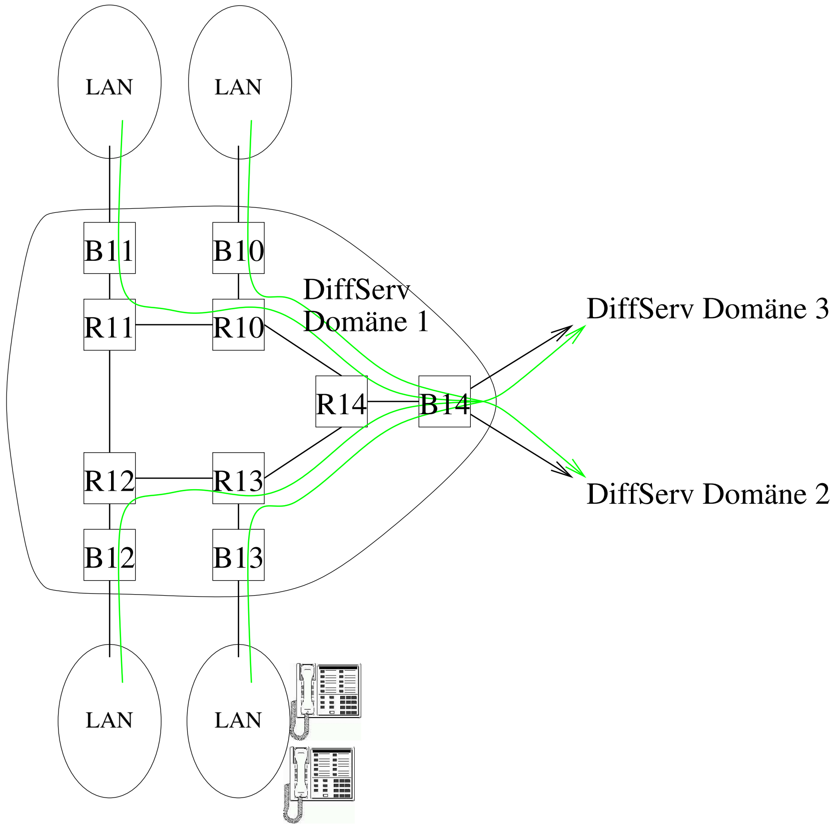


Fig. 2. Aufbau einer DiffServ-Domäne

Metern erlaubt ist, so daß ein Teil der Pakete als out-ouf-profile markiert wird und damit eine höhere Drop-Precedence erfährt. Jedes LAN darf nur jeweils 800 kbit/s in-profile-Verkehr in den Klassen AF1 und AF2 an seine DiffServ-Domäne senden. Der Rest wird als out-of-profile markiert. Zwischen R14 und B14 aggregiert sich dieser Verkehr zu einem Strom von jeweils 3.2 Mbit/s.

Jedes LAN erzeugt neben dem bereits erwähnten Hintergrundverkehr vier Datenströme, die aus den Paketen der Sprachverbindungen bestehen. Von den vier Daten-

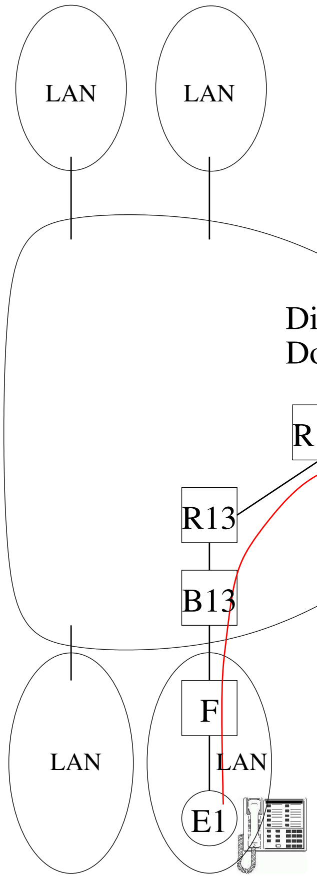


Fig. 3. Simulationsszenario

strömen, die von einem LAN erzeugt werden, gehen jeweils 2 Datenströme zu einem LAN in einer der beiden benachbarten DS-Domänen (vgl. Abbildung 2).

Die höchste Konzentration von Telefonieverkehr tritt in den DiffServ-Domänen zwischen dem Border-Router zu den beiden anderen DiffServ-Domänen und dem daran angeschlossenen inneren Router auf, also z.B. in der DS-Domäne 1 zwischen den Routern B14 und R14 (vgl. Abbildung 3). Es handelt sich um 16 einzelne Verkehrsströme, die zusammen als aggregierter Strom ca. 1.1 Mbit/s Bandbreite benötigen. Der Anteil der Link-Kapazität, die für Premium Service mit Hilfe des CBQ-Verfahrens reserviert wird, ist so bemessen, daß der Telefonieverkehr ohne Verluste und große Verzögerungsschwankungen fließen kann, falls er als Premium-Service-Verkehr klassifiziert wird. Außerdem wird der Premium-Service-Verkehr in den CBQ-Warteschlangen mit hoher Priorität behandelt, so daß die Warteschlangen für Premium Service nie mehr Pakete enthalten sollten als die Anzahl der an dem Router ankommenden Links. Assured Service und Best Effort besitzen die gleiche Priorität, die allerdings niedriger ist als diejenige von Premium Service. Die Zugangskontrolle in den Border-Routern verhindert, daß der Premium Service mehr als die ihm zugeordneten Ressourcen erhalten kann.

Die Zuordnung von Bandbreite auf den 10Mbit/s-Links innerhalb und zwischen den DS-Domänen auf die einzelnen PHBs ist wie folgt konfiguriert:

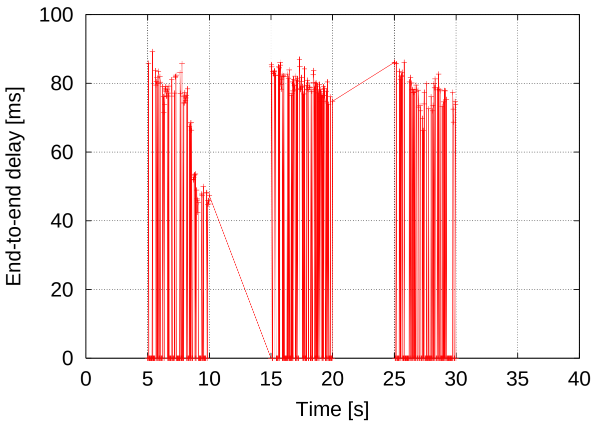
EF (Premium Service)	11%
AF1 (Assured Service)	40%
AF2 (Assured Service)	30%
BE (Best Effort)	5%

Die restliche nicht zugeteilte Bandbreite, sowie ungenutzte, einem PHB zugeteilte Bandbreite, kann durch die CBQ-Mechanismen auf aktive PHBs verteilt werden.

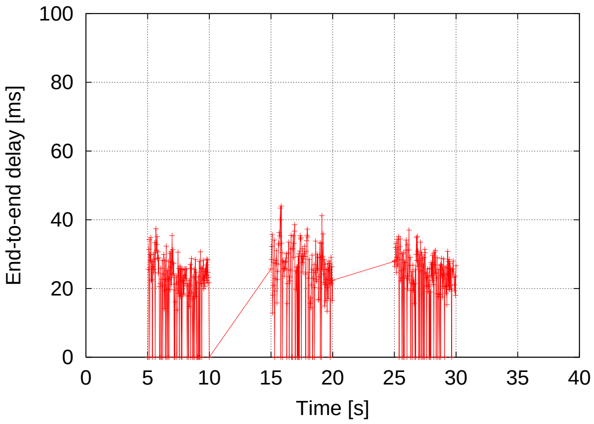
Zur Beurteilung der Qualität von IP-Telefonie-Datenströme wurden drei verschiedene *Sprachverkehrsmuster* untersucht:

- 64 kbit/s CBR.
- gemessener LiveLan-Verkehr
- gemessener NetMeeting-Verkehr

Für den *CBR-Verkehr* werden in Intervallen von 32 ms Pakete einer Länge von 256 Bytes gesendet. Die IP-Telefoniedatenströme senden in drei Intervallen von jeweils 5 s Länge mit Unterbrechungen von 5 s. Es soll hier-

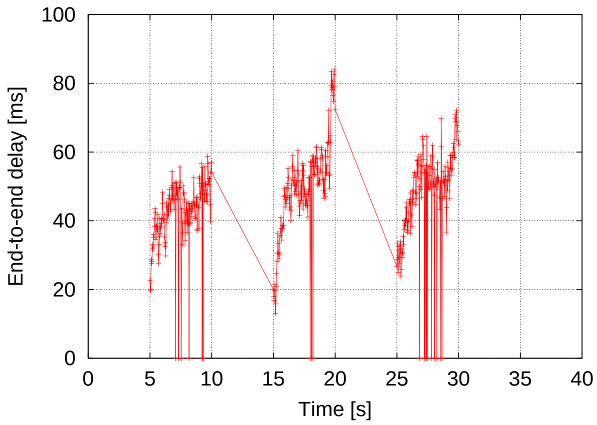


(a) Best Effort mit DropTail/FIFO

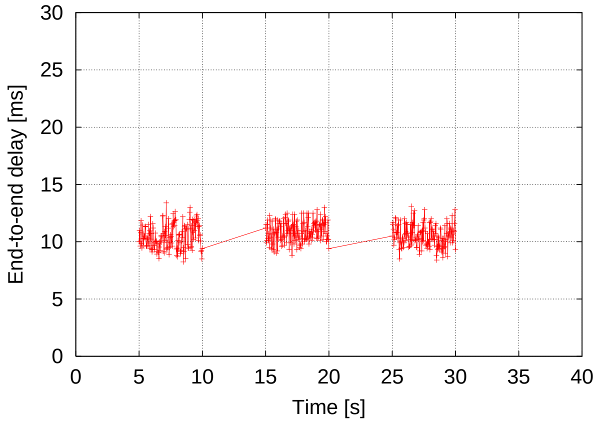


(b) Best Effort mit RED

Fig. 4. Verzögerung mit Best Effort für CBR-Verkehr



(a) Assured Service



(b) Premium Service

Fig. 5. Verzögerung mit DiffServ für CBR-Verkehr

mit das Verhalten des Verkehrs direkt nach dem Einschalten bei evtl.[1] Silence Suppression beobachtet werden.

Bei *LiveLan* werden Pakete von 566 Bytes benutzt. Die Sendeintervalle schwanken zwischen 50 ms und 80 ms, so daß sich eine mittlere Bandbreite von ca. 70 kbit/s ergibt.

Der *NetMeeting-Verkehr* ist deutlich burstartiger als LiveLan. Die Paketgröße beträgt 78 Bytes, die Intervalle zwischen den Paketen liegen zwischen 3.5 ms und 200 ms. Dabei treten meist zwischen 2 und 5 Pakete in sehr kurzen Abständen von 3.5 ms auf, gefolgt von einem Intervall von 70-200 ms. Die mittlere Bandbreite des gemessenen NetMeeting-Verkehrs beträgt ca. 16 kbit/s.

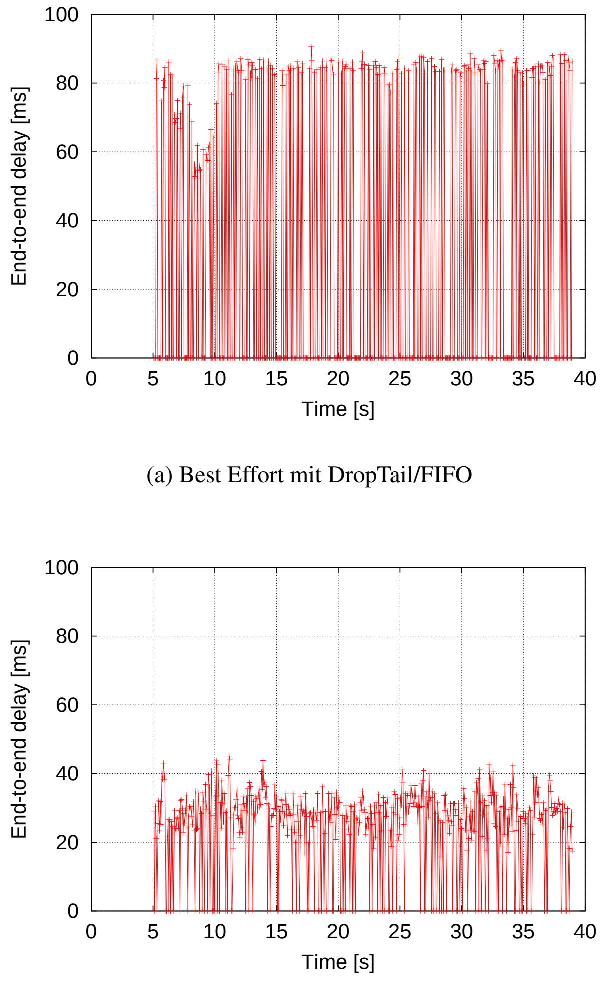
Die Pakete eines IP-Telefonie-Datenstroms werden in den Simulationen mit dem DS-codepoint für Best Effort, Assured Service bzw. Premium Service markiert. Im Falle von Assured Service werden die Pakete alle als in-profile markiert.

[1] H.323 und G.711 spezifizieren keine Silence Suppression für Audio-Datenströme

III. Simulative Experimente

Mit dem in Abschnitt II kurz skizzierten Simulationsmodell wurde eine Reihe von Experimenten durchgeführt zur Evaluierung der Eignung von DiffServ für die Unterstützung von IP-Telefonie.

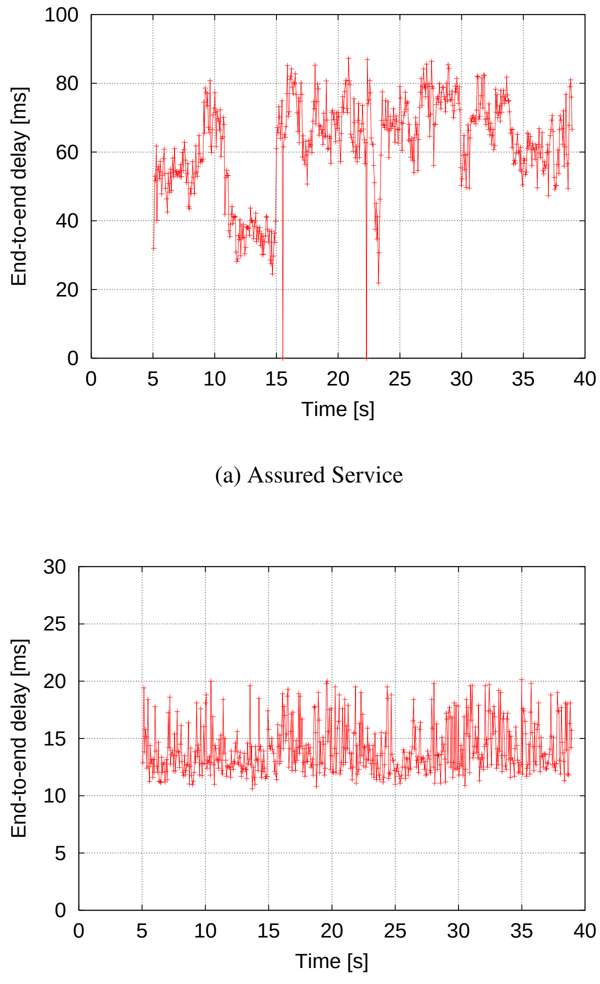
Um einen Vergleich der mit DiffServ erzielten Dienstgüte gegenüber derjenigen ohne DiffServ zu haben, wurden zunächst Experimente ohne spezielle Dienstgüteunterstützung im Netz durchgeführt. Der durch IP-Telefonie verursachte Verkehr wird somit einfach als Best-Effort behandelt. Die Warteschlangen der Router arbeiten gemäß dem FIFO-Prinzip bzw. werden zusätzlich durch RED unterstützt. Die Warteschlangen sind dabei auf eine maximale Länge von 50 Paketen konfiguriert. Exempla-



(a) Best Effort mit DropTail/FIFO

(b) Best Effort mit RED

Fig. 6. Verzögerung mit Best Effort für LiveLan-Verkehr



(a) Assured Service

(b) Premium Service

Fig. 7. Verzögerung mit DiffServ für LiveLan-Verkehr

risch wird in den folgenden Untersuchungen ein dedizierter IP-Telefonie-Datenstrom herausgegriffen.

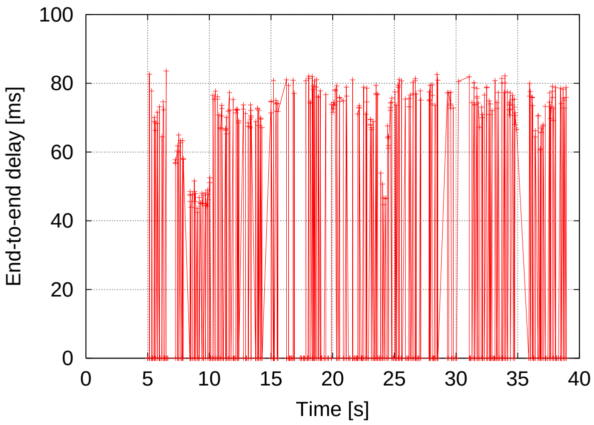
A. *CBR-basierter Telefonverkehr*

Der CBR-Datenstrom ist jeweils dreimal für eine Zeitdauer von jeweils 5 Sekunden aktiv, wobei je 471 Pakete erzeugt werden. In den Diagrammen ist jeweils die Ende-zu-Ende Verzögerung dargestellt. Paketverluste sind durch eine Verzögerung von 0 repräsentiert.

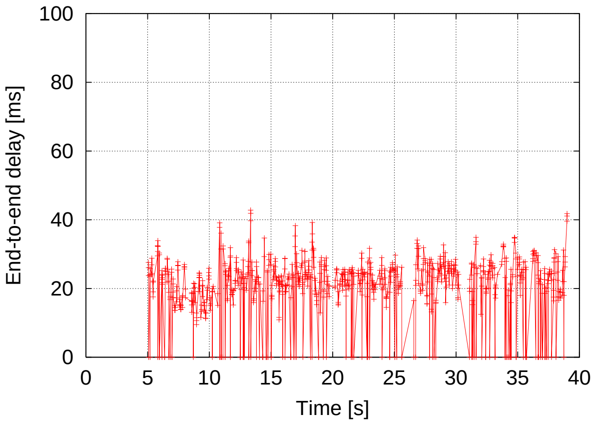
Bei der Verwendung von *Best-Effort* sind hohe Paketverluste von 59% zu verzeichnen. Die Verzögerungen liegen im Interval $[42.5\text{ms}, 89.2\text{ms}]$. Die mittlere Verzögerung beträgt 75.6 ms. Die hohen Verzögerungen und Verluste kommen zustande, da die Quellen für den Hintergrundverkehr mehr Verkehr erzeugen, als das Netz transportieren kann. Dadurch wachsen die Warteschlangen in den Routern bis zur ihrer maximalen Länge. Ab diesem Zeitpunkt werden ankommende Pakete verworfen. Telefondaten werden hier genauso behandelt wie der Hintergrundverkehr auch.

Bereits die Verwendung von *RED* führt zu einer sichtbaren Verbesserung der Qualität. Durch den RED-Algorithmus wird die mittlere Länge der Warteschlange durch frühzeitiges Verwerfen von Paketen gering gehalten, so daß sich die Verzögerung in der Warteschlange reduziert. Burstartiger Verkehr kann die Länge der Warteschlange und damit die Verzögerung jedoch für kurze Zeitintervalle erhöhen. Die Verzögerung reduziert sich durch RED auf das Interval $[12.9\text{ms}, 44.0\text{ms}]$ mit einem Mittelwert von 25.9 ms. Die Paketverluste verbessern sich ebenfalls erheblich, sind aber mit 18% nach wie vor sehr hoch. Die Dienstqualität ist damit für den Benutzer zwar besser aber noch nicht akzeptabel.

Mit DiffServ-Mechanismen läßt sich die Qualität der Telefonie nochmals deutlich verbessern. Bei Verwendung von *Assured Service* verringert sich allerdings die

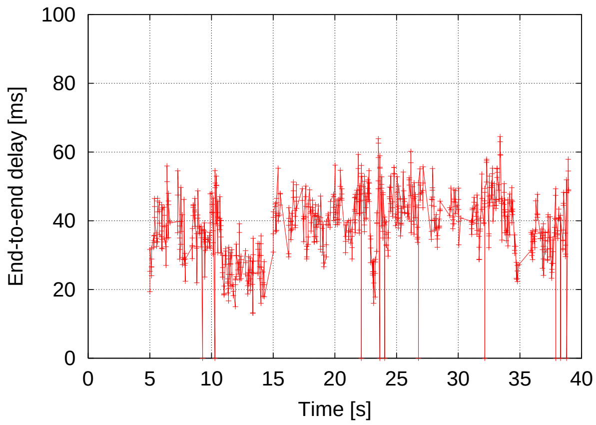


(a) Best Effort mit DropTail/FIFO

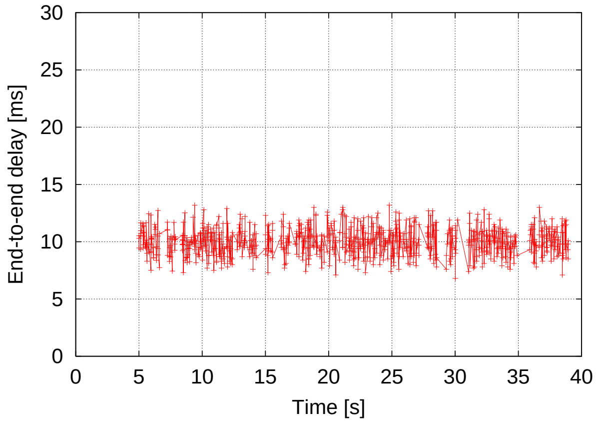


(b) Best Effort mit RED

Fig. 8. Verzögerung mit Best Effort für NetMeeting-Verkehr



(a) Assured Service



(b) Premium Service

Fig. 9. Verzögerung mit DiffServ für NetMeeting-Verkehr

Dienstgüte für den Telefonieverkehr, wenn nicht die Länge der Warteschlangen begrenzt wird. Die Pakete der IP-Telefonie-Datenströme werden hier zusammen mit einem Teil des Hintergrundverkehrs durch den Assured Forwarding PHB (AF) weitergeleitet. Der Rest des Hintergrundverkehrs ist Best Effort. Die Telefoniepakete sind alle als in-profile markiert. Die hohe Verzögerung gegenüber einfachem RED ist darin begründet, daß der AF1-Klasse nur 40% der Link-Bandbreite zur Verfügung stehen und die RIO-Warteschlange für AF1 dadurch langsamer abgearbeitet wird. Als Out-of-profile markierte Pakete werden zwar mit höherer Wahrscheinlichkeit verworfen, dennoch bewegt sich die Länge der Warteschlange an der oberen Grenze, da sie schon mit den in-profile Paketen (3.2 Mbit/s Hintergrundverkehr und 1.1 Mbit/s Telefonieverkehr) leicht überlastet ist. Die Paketverlustrate ist aber gegenüber RED merkbar auf 4% verbessert. In einem überlasteten Netz, in dem der Assured Service keine ungenutzte Bandbreite von anderen Klassen borgen kann, kann eine Übertragung ohne Paketverluste nur erreicht werden, wenn die für Assured Service reservierte Bandbreite ausreichend dimensioniert ist.

Durch Verkürzung der Warteschlangenlänge für jede Warteschlange innerhalb der CBQ-Klassen lassen sich ähnliche Ergenisse wie mit RED erreichen.

Die Verzögerungen der Telefoniepakete in den Warteschlangen kommen einerseits durch Pakete derselben AF-Klasse, die in der Warteschlange weiter vorne stehen, zustande, andererseits aber auch durch Pakete in den Warteschlangen anderer PHBs, also auch Best-Effort. Dies liegt daran, daß der AF PHB zwar bestimmte Ressourcen besitzt, die nicht von Best-Effort-Verkehr genutzt werden können, solange AF-Verkehr vorhanden ist, aber AF besitzt keine höhere Priorität gegenüber Best-Effort. AF garantiert also eine bestimmte Bandbreite mit geringer Verlustwahrscheinlichkeit aber keine geringe Verzögerung.

	Verlustrate	min. Verzögerung	max. Verzögerung	mittl. Verzögerung
Best Effort	59%	42.5	89.2	75.6
Best Effort mit RED	18%	12.9	44.0	25.9
Assured Service	4%	13.0	83.9	47.8
Premium Service	0%	8.2	13.4	10.6

TABLE I
UNTERSCHIEDLICHE DIENSTE FÜR CBR-VERKEHR IM VERGLEICH

	Verlustrate	min. Verzögerung	max. Verzögerung	mittl. Verzögerung
Best Effort	60%	52.8	90.7	82.0
Best Effort mit RED	22%	16.0	45.1	29.7
Assured Service	0.4%	21.9	87.3	62.3
Premium Service	0%	10.6	20.2	14.1

TABLE II
UNTERSCHIEDLICHE DIENSTE BEIM EINSATZ VON LIVELAN-VERKEHR

	Verlustrate	min. Verzögerung	max. Verzögerung	mittl. Verzögerung
Best Effort	65%	42.5	83.6	70.3
Best Effort mit RED	14%	9.5	42.8	23.5
Assured Service	1%	13.1	64.5	38.9
Premium Service	0%	6.8	13.2	10.0

TABLE III
UNTERSCHIEDLICHE DIENSTE BEIM EINSATZ VON NETMEETING-VERKEHR

Über den Assured Service hinaus bietet DiffServ den *Premium Service* an. Die Pakete werden durch den Expedited Forwarding PHB (EF) mit höherer Priorität bearbeitet als alle Pakete der anderen Klassen. Es wird von den Ingress-Routern nur soviel Premium-Service-Verkehr in das Netz gelassen, daß an jedem Router die Summe der Bandbreiten der aggregierten Premium-Service-Ströme kleiner gleich der Bandbreite des ausgehenden Links für diesen Verkehr ist. Dadurch kann die Warteschlange für Premium Service nie mehr Pakete enthalten als die Anzahl der eingehenden Links mit Premium-Service-Verkehr. Zusammen mit der höchsten Priorität ist dadurch minimale Verzögerung gewährleistet.

B. LiveLan-Verkehrsmuster

Die Ergebnisse der Simulationen mit LiveLan-Verkehr sind in den Abbildungen 6 und 7 zu sehen. Die Abbildungen zeigen wieder jeweils die Ende-zu-Ende-Verzögerung für die Verbindung zwischen den Knoten E1 und E2.

Wie beim CBR-Verkehr treten ohne DiffServ-Mechanismen hohe Verzögerungen, Verzögerungsschwankungen und Paketverluste auf. Abbildung 6(a) zeigt Verzögerungen zwischen 52.8 ms und 90.7 ms. Die Paketverlustrate beträgt 60%. Durch das RED-Warteschlangenverfahren kann auch hier die Qualität verbessert werden. Die mittlere Verzögerung sinkt von 82.0 ms auf nur noch 29.7 ms mit Schwankungen im Bereich von $[16.0\text{ms}, 45.1\text{ms}]$. Die Paketverlustrate ist mit 22% aber immer noch sehr hoch. Mit Assured Service kann die Paketverlustrate erheblich verringert werden, wobei allerdings die Verzögerung wie bei CBR steigt. Mit Premium Service gehen dagegen keine Pakete mehr verloren. Die mittlere Verzögerung kann durch Premium Service auf 14.1 ms verbessert werden. Abbildung 7 zeigt, daß mit Premium Service auch die Verzögerungsschwankungen gegenüber Assured Service deutlich verbessert werden.

C. NetMeeting-Verkehrsmuster

In den Abbildungen 8 und 9 sind die Ergebnisse derselben Simulationen mit NetMeeting-Verkehr statt LiveLan dargestellt.

Der Vergleich sowohl der Abbildungen 6 bis 9 als auch der Tabellen II und III zeigt ein sehr ähnliches Verhalten von LiveLan-Verkehr und NetMeeting-Verkehr für alle Dienstklassen.

Der Vergleich von LiveLan- und NetMeeting-Verkehr zeigt, daß die Ende-zu-Ende-Verzögerung von NetMeeting-Verkehr fast durchgehend niedriger liegt als die des LiveLan-Verkehrs. Dies ist auf die geringere Paketgröße von nur 78 Bytes gegenüber 566 Bytes zurückzuführen. Die reine Übertragungs- und Ausbreitungszeit zwischen den Endsystemen E1 und E2 beträgt 6.7 ms für 78 Bytes große Pakete und 10.3 ms für 566 Bytes große Pakete. Dieser Vorteil der kleineren Pakete wird aber teilweise wieder durch den stärker burstartigen Charakter des NetMeeting-Verkehrs zunichte gemacht.

IV. Zusammenfassung und Ausblick

Im Rahmen dieser Arbeit wurde ein ns-2-Simulationsmodell für IP-Telefonie über Differentiated Services implementiert. Mit diesem Modell wurde die erreichbare Dienstgüte für IP-Telefonie bei hoher Hintergrundlast mit verschiedenenen DiffServ-Service-Klassen untersucht. Die Ergebnisse der Simulationen zeigen, daß sich bei der gewählten Hintergrundlast IP-Telefonie mit Best Effort Service nicht in akzeptabler Qualität implementieren läßt. Mit RED ohne Differenzierung der Pakete verschiedener Verkehrsquellen können Pakteverluste und Verzögerung zwar reduziert werden, bleiben jedoch in den durchgeführten Simulationen bei Werten, die nicht für eine gute Sprachqualität ausreichen. Die durchgeführten Simulationen zeigen, daß durch Differentiated Services wie zu erwarten eine erhebliche Verbesserung sowohl der Paketverluste als auch der Ende-zu-Ende-Verzögerung möglich ist, wenn Assured Service oder Premium Service für IP-Telefonie verwendet wird. Dabei zeigt sich erwartungsgemäß, daß sich Premium Service gegenüber Assured Service neben der geringeren mittleren Verzögerung vor allem durch niedrigere Verzögerungsschwankungen auszeichnet.

Weitere Arbeiten werden sich mit der Untersuchung anderer Warteschlangen-Algorithmen als CBQ beschäftigen. Außerdem werden die Simulationen mit burstartigeren Verkehrsmodellen, wie sie z.B. für Videoquellen typisch sind, durchgeführt werden. Zur Zeit wird außerdem an der Implementierung eines Testbetts zur Demonstration der praktischen Umsetzbarkeit der Ergebnisse gearbeitet.

V. Danksagungen

Die Untersuchungen entstanden im Rahmen eines von der Siemens AG, ICN geförderten Projektes.

Für die Bereitstellung der aufgezeichneten Verkehrsströme für LiveLan und NetMeeting bedanken wir uns herzlich bei der Gruppe von Prof. Dr. J. Eberspächer an der TU München.

References

[1] International Telecommunication Union, “Visual telephone systems and equipment for local area networks which provide a non-guaranteed quality of service,” Recommendation H.323, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, May 1996.

[2] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, “An architecture for differentiated service,” Request for Comments (Informational) 2475, Internet Engineering Task Force, Dec. 1998.

[3] J. Heinanen, F. Baker, W. Weiss, and J. Wroclawski, “Assured forwarding PHB group,” Request for Comments (Proposed Standard) 2597, Internet Engineering Task Force, June 1999.

[4] V. Jacobson, K. Nichols, and K. Poduri, “An expedited forwarding PHB,” Request for Comments (Proposed Standard) 2598, Internet Engineering Task Force, June 1999, URL: ftp://ftp.ee.lbl.gov/papers/ef_phb.pdf.

[5] Sally Floyd and Van Jacobson, “Link-sharing and resource management models for packet networks,” *IEEE/ACM Transactions on Networking*, vol. 3, no. 4, Aug. 1995.

[6] K. Nichols, V. Jacobson, and L. Zhang, “A two-bit differentiated services architecture for the internet,” Request for Comments (Informational) 2638, Internet Engineering Task Force, July 1999.

[7] Mark. E. Crovella and Bestavros Azer, “Self-similarity in world wide web traffic: Evidence and possible causes,” *IEEE/ACM Transactions on Networking*, vol. 5, no. 6, pp. 835–846, Dec. 1997.