

# **3V**

## **Verteilte Verkehrssimulation und Visualisierung**

**Abschlußbericht**

Gigabit Testbed West

1. November 1998 – 30. Juni 2000

**Universität zu Köln - Zentrum für Paralleles Rechnen (ZPR), Regionales Rechenzentrum (RRZK)**

**Deutsches Zentrum für Luft- und Raumfahrt (DLR)**

**Forschungszentrum Informationstechnik GmbH (GMD)**

**Echtzeit GmbH**

**Juli 2000**

**WWW:** <http://www.3V.Uni-Koeln.de/>



# Inhaltsverzeichnis

1	Einleitung .....	5
2	Universität zu Köln/Deutsches Zentrum für Luft- und Raumfahrt.....	7
2.1	Einleitung .....	7
2.2	Netzwerkschnittstelle .....	9
2.3	Simulation.....	9
2.3.1	Das Simulationsmodell.....	9
2.3.2	Abspalten der Visualisierung .....	10
2.3.3	Die Kommunikationsschnittstelle .....	11
2.3.4	Mehrprozessorsysteme .....	11
2.4	Visualisierung.....	12
2.4.1	2D-Visualisierung.....	13
2.5	Der Rückkanal.....	14
2.5.1	Umsetzung.....	14
2.5.2	Ein Beispiel: Sperrung einer Straße .....	15
2.6	Zusammenfassung und Ausblick.....	17
3	Echtzeit .....	19
3.1	Einleitung .....	19
3.2	Modellierungswerkzeug für Verkehrsnetze .....	19
3.3	VR-Visualisierungsclient .....	20
3.3.1	Laufzeitumgebung.....	20
3.3.2	Datenaustausch.....	21
3.4	Visualisierung.....	21
3.4.1	Steuerung.....	21
3.4.2	Stadtmodell.....	21
3.5	Resümee .....	22
3.6	Ausblick.....	22
4	GMD – Forschungszentrum Informationstechnologie.....	23
4.1	Übersicht.....	23
4.2	Methoden für Multicast-Verteilung.....	24
4.2.1	Native ATM.....	24
4.2.2	Gemeinsames IP-Netz / LANE.....	26
4.2.3	IP Multicast-Routing .....	28
4.2.4	Sonstige Verfahren.....	30
4.2.5	Fazit.....	30
4.3	Messungen.....	31

4.3.1	Testwerkzeug.....	31
4.3.2	Theoretische Übertragungsraten.....	31
4.3.3	Einfluss der Rechnerarchitektur.....	32
4.3.4	Einfluss der Paketgröße.....	34
4.3.5	Vergleich der Technologien.....	36
4.3.6	Vergleich zwischen Multicast und Unicast.....	37
4.3.7	Paketverluste.....	38
4.3.8	IP Multicast Routing.....	41
4.3.9	Parallele TCP- und Multicast-Ströme.....	43
4.3.10	Sendeverhalten der Applikation PLANSIM-T.....	45
4.4	Zusammenfassung und Ausblick.....	46
5	Zusammenfassung.....	49
6	Berichte und Vorträge.....	51
7	Literatur.....	53
8	Anhang.....	55
8.1	LANE.....	55
8.1.1	LECS.....	55
8.1.2	LES/BUS.....	56
8.1.3	Klienten.....	57
8.2	XTI.....	58
8.3	IP Router.....	58

# 1 Einleitung

Am 1. August 1997 startete das Projekt Gigabit Testbed West. Dieses vom DFN eingeleitete und vom Bundesministerium für Bildung und Forschung (BMBF) geförderte Projekt hatte die Erprobung und Untersuchung der Gigabit-Netztechnik zum Ziel, welche im Frühjahr 2000 vom DFN in Betrieb genommen wurde. Zusätzlich sollte der exemplarische Nachweis über den Nutzen dieser Netztechnik geliefert werden, indem wissenschaftliche Anwendungen entwickelt werden, die erst durch das GTB möglich sind.

Die Motivation des Projektes „3V – Verteilte Verkehrssimulation und Visualisierung“ vor diesem Hintergrund ist, die neuen Möglichkeiten von Netzwerken mit GigaBit-Bandbreiten zu nutzen, um Formen der Kopplung zwischen Anwendungen (hier am Beispiel einer Verkehrssimulation) und ihren Visualisierungen zu untersuchen, die bislang nicht möglich waren. Dazu sollte eine Verkehrssimulation, die aufgrund ihres mikroskopischen Charakters (Molekulardynamik Simulation einer sehr großen Anzahl einzelner Fahrzeuge mit Zeitschrittweiten von 1 Sekunde in vielfacher Echtzeit) sehr große Datenraten generiert, so erweitert werden, dass die errechnete ???, über Multicast verteilt visualisieren. Darüber hinaus sollte die Visualisierung flexibel gestaltet sein. Verschiedene Nutzer sollten dem jeweiligen Interesse angepasste Visualisierungen zur Verfügung haben (für den Experten die wissenschaftliche Darstellung, für den Politiker die VR-Visualisierung, um die beiden Eckpunkte des Spektrums zu nennen). Schließlich ist es für die Nutzer von Interesse, in die laufende Simulation auch eingreifen zu können.

Dieser kurzen Beschreibung entsprechend, unterteilt sich das Projekt in drei Teilbereiche, die in diesem Abschlußbericht jeweils separat diskutiert werden:

- Verkehrssimulation und der Eingriffsmöglichkeiten in diese Simulation (Rückkanal), sowie der Parallelisierung (ZAIK/DLR),
- Netzwerkperformance und Multicast (GMD),
- VR-Visualisierung (echtzeit) und wissenschaftliche Visualisierung (ZAIK).

Dass diese Fragen nur beispielhaft und auch nicht immer in der erwünschten Perfektion bearbeitet werden konnten, kann bei einem Ansatz der zum Teil auf die erforderliche Hardware warten musste, nicht überraschen. Beispielsweise ist der Ausbau der DLR-Infrastruktur so weit hinter den ursprünglichen Planungen zurückgeblieben, dass es im Interesse des Projektes sinnvoll erschien, die Verbindung zwischen GMD und Uni Köln dort einfach nur durchzureichen. Was in diesem Fall nur bedeutet, dass ein Multicast-Partner weniger vorhanden war.

Folgende Institutionen und Firmen waren an dem Projekt beteiligt:

- Universität zu Köln (Zentrum für Paralleles Rechnen und Rechenzentrum)  
Ansprechpartner: Jürgen Gräfe  
Zentrum für Paralleles Rechnen (ZPR)  
Universität zu Köln  
Weyertal 80  
D-50931 Köln  
Tel.: 0221 / 470 6010  
Fax.: 0221 / 470 5160  
Email: [graefe@zpr.uni-koeln.de](mailto:graefe@zpr.uni-koeln.de)
- GMD – Forschungszentrum Informationstechnik GmbH  
Ansprechpartner: Lothar Zier

GMD – Forschungszentrum Informationstechnik GmbH  
Institut für Medienkommunikation  
Schloß Birlinghoven  
D-53757 Sankt Augustin  
Tel.: 02241 / 14 29 43  
Fax.: 02241 / 14 20 95  
Email: lothar.zier@gmd.de

- DLR – Deutsches Zentrum für Luft- und Raumfahrt  
Ansprechpartner: Dr. Peter Wagner  
Deutsches Zentrum für Luft- und Raumfahrt – DLR  
Abteilung FF-MS  
Linder Höhe  
D-51147 Köln  
Tel.: 02203 / 603 28 53  
Email: peter.wagner@dlr.de
- echtzeit GmbH  
Ansprechpartner: Kay-Ingo Ahlers  
echtzeit GmbH  
Kantstraße 165  
D-10623 Berlin  
Tel.: 030 / 88 45 86 167  
Fax.: 030 / 88 45 86 22  
kayingo@echtzeit.de

## **2 Universität zu Köln/Deutsches Zentrum für Luft- und Raumfahrt**

### **Ansprechpartner**

Jürgen Gräfe, E-mail: graefe@zpr.uni-koeln.de

Peter Wagner, E-mail: peter.wagner@dlr.de

### **Weitere Mitarbeiter und Alumni:**

Frank Brockners, Marcus Metzler, Felix Pütsch, Olaf Schumann, Achim Gädke, Nils Eissfeldt, Christian Rössel

### **2.1 Einleitung**

Die unterschiedlichen Maßnahmen zur Beeinflussung des Verkehrsgeschehens werden von den Verantwortlichen häufig gefällt, ohne die Möglichkeit zu haben, die Auswirkungen abschätzen zu können. Da die Auswirkungen aufgrund der komplexen Wechselwirkungen zwischen Verkehr und Netzwerk analytisch schwer abzuschätzen sind, sind Simulationen des Straßenverkehrs für solche Entscheidungen von großer Bedeutung.

Ein wesentliches Problem bei der Nutzung der aus der Verkehrssimulation gewonnenen Informationen ist die heterogene Zusammensetzung der Entscheidungsträger und der damit verbundenen unterschiedlichen Gewohnheiten, Informationen zu visualisieren und auszuwerten. Deshalb ist es notwendig, die von der Simulation gelieferten Daten den einzelnen Zielgruppen in Art und Umfang ihrer Gewohnheit entsprechend anzubieten.

Die Bandbreite des GTB ermöglicht es erstmals, die Daten einer Simulation großer Straßennetze in Echtzeit verteilt zu visualisieren und somit heterogene Gruppen zusammenarbeiten zu lassen. Einzelne Anwender können mit Hilfe eines Rückkanals in das Simulationsgeschehen eingreifen, wobei sich dies auf alle Mitglieder der Arbeitsgruppe auswirkt.

Die bisherige Version der zugrundeliegenden Simulationssoftware PLANSIM-T war vor allem für Forschungsarbeiten ausgelegt. Es fehlte die Möglichkeit die Verkehrsflüsse und die Auswirkungen straßenbaulicher Maßnahmen auf die unterschiedlichsten Arten und losgelöst von der Simulation zu visualisieren.

Zwar existierte in PLANSIM-T bereits eine 2D-Visualisierung, die allerdings fest in den Programmablauf eingebunden und tief in der Programmstruktur verwurzelt ist. Sie diente hauptsächlich dazu, den Ablauf der Simulation kontrollieren zu können und eventuelle Fehler in der Struktur des Straßennetzes, wie z.B. fehlende Verbindungen zwischen Kanten, oder in der Dynamik der Fahrzeuge, erkennen zu können. Aufgrund der Einbettung der Visualisierung war es nicht möglich, diese zwischenzeitlich ab- oder zuzuschalten. Dies bedeutete eine beträchtliche Einschränkung in der Nutzbarkeit der Visualisierung, da mit ihr die Laufzeit der Simulation um ein Vielfaches wächst.

Dies führte zu der Idee, die Visualisierung von der Simulation abzuspalten und autonom laufen zu lassen. Das GTB ermöglicht es nun zusätzlich, die Visualisierung getrennt auf einem anderen Rechner laufen zu lassen, welcher durch ein Netzwerk mit sehr hoher Bandbreite mit der Simulation verbunden ist. Zum Vergleich: Bei der Simulation eines Stadtnetzes mit 100.000 Fahrzeugen fallen ca. 400 MBit/s an Daten an. Eine herkömmliche Netzleitung kann

maximal 155 MBit/s übertragen. Das GTB hingegen ermöglicht Übertragungsraten bis zu 622 MBit/s.

Mit dieser Methode ist es damit grundsätzlich auch möglich, verschiedene Visualisierungen auf unterschiedlichen Rechnern gleichzeitig zu betreiben.

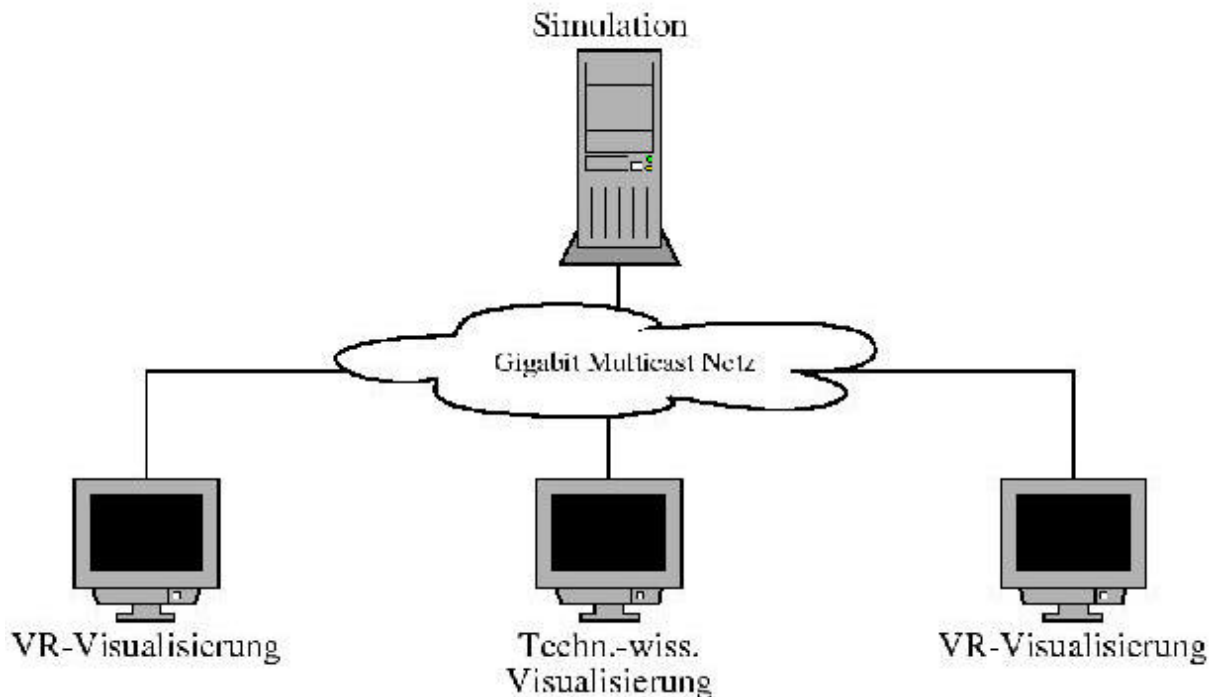


Abbildung 2-1: Verteilte Verkehrssimulation und Visualisierung

Sobald verschiedene Visualisierungen der Simulationsdaten an das Netzwerk angeschlossen werden, tritt ein weiteres Problem auf. Da jeder Empfänger mit Daten versorgt werden muß, vergrößert sich die Netzlast um einen entsprechenden Faktor.

Um dieses Problem zu lösen, wird auf Multicast-Verfahren zurückgegriffen. Hierbei muß sich jeder Klient an eine Multicast-Gruppe anmelden. Die Simulation schickt die Daten nun nicht mehr an jeden Empfänger, sondern an die Gruppe und das Netz verteilt die Pakete nach einem bestimmten Schema, so dass jedes Mitglied der Gruppe die Daten erhält, jedoch die Netzbelastung insgesamt so gering wie möglich gehalten wird.

Ein weiteres Defizit von PLANSIM-T war die fehlende Möglichkeit, mit der Simulation interaktiv zu kommunizieren. Alle Einstellungen und Optionen mußten beim Programmstart festgelegt werden und konnten während des Simulationslaufes nicht geändert werden. Gerade im Hinblick auf die Bewertung von Planungsmaßnahmen wäre es aber sinnvoll, wenn diese der Simulation übermittelt werden und die Auswirkungen im laufenden Verkehrsgeschehen untersucht werden könnten.

Um diese Funktionalität zu ermöglichen, wurde ein Rückkanal eingebaut. Dieser stellt die Kommunikationsschnittstelle zwischen Visualisierung und Simulation in umgekehrter Richtung dar.

Insgesamt wurden folgende Aufgaben im Rahmen des Projektes durchgeführt:

- Abspaltung der Visualisierung von PLANSIM-T
- Entwicklung und Implementation eines Kommunikationsprotokolles zwischen der Simulation und den Visualisierungen über eine Netzleitung
- Senden der Simulationsdaten über Multicast-Protokolle (IP-Multicast, ATM)
- Entwicklung, Implementation und Test eines Rückkanals zwischen Visualisierung und Simulation.

In den folgenden Abschnitten erfolgt eine nähere Beschreibung der durchgeführten Arbeiten.

## 2.2 Netzwerkschnittstelle

Speziell werden in der Simulation die Protokolle IP-Multicast und ATM eingesetzt. Beide Protokolle unterstützen das gleichzeitige Senden von Paketen an mehrere Teilnehmer, ohne die Netzbelastung proportional zu der Anzahl der angebundenen Klienten zu erhöhen. Zusätzlich wurde noch das TCP/IP-Protokoll implementiert, welches jedoch nicht Multicast-fähig ist. Es spielt deshalb zwar nur eine untergeordnete Rolle in der anvisierten Anwendung, da aber die Abspaltung der Visualisierung unabhängig von der Übermittlung der Daten und der dazu verwendeten Protokolle realisiert wurde, konnte TCP/IP für Testzwecke genutzt werden. Beide Multicast-Protokolle haben ihre speziellen Vorzüge. Während IP-Multicast keine spezielle Hardware voraussetzt und auf fast allen Rechnern verfügbar ist, kann ATM eine bestimmte Bandbreite für eine Netzverbindung garantieren und ist somit besonders gut geeignet, wenn keine Pakete im Netz verloren gehen dürfen.

Das gewünschte Protokoll muß beim Start von PLANSIM-T mit angegeben werden und kann während des Simulationslaufes nicht mehr geändert werden. Dazu besteht jedoch in der Regel auch keine Notwendigkeit. Aus dem gleichen Grund ist es nicht möglich, eine Simulation mit mehreren Protokollen gleichzeitig zu betreiben.

## 2.3 Simulation

### 2.3.1 Das Simulationsmodell

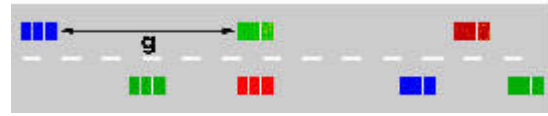
Bei PLANSIM-T handelt es sich um eine Mikrosimulation, welche auf einem Fahrzeugfolgemodell beruht. Wahlweise lassen sich hier verschiedene mikroskopische Modelle, auch innerhalb eines Netzes, einsetzen. Gebräuchlich sind das CA-Modell (Cellular Automata) von Nagel und Schreckenberg [KN] oder das SK-Modell (nach Stefan Krauß) [SK]. In einem Zeitschritt (Microstep) der Simulation, üblicherweise eine Sekunde, werden die Autos aufgrund der Regeln des Modells bewegt.

In beiden Modellen wird zunächst für jedes Auto eine sichere Geschwindigkeit  $v_{safe}$  berechnet, die Kollisionsfreiheit garantiert. Dazu fließen der Abstand zum Vordermann  $g$  und im SK-Modell auch dessen Geschwindigkeit  $v_{leader}$  und die eigene Geschwindigkeit  $v$  ein. Suboptimales Fahrverhalten wird durch eine stochastische Komponente abgebildet, indem die Fahrer mit einer gewissen Wahrscheinlichkeit trödeln, also langsamer fahren als es die Sicherheitsbedingung ermöglicht. Im folgenden werden die beiden Modelle kurz skizziert:



$v_{\text{safe}} := g$   
 $v := \min\{v + 1, v_{\text{max}}, v_{\text{safe}}\}$   
 if  $(v > 0)$  and  $(\text{rnd} < p)$   
 then  $v := v - 1$   
 $x := x + 1$

mit maximaler Geschwindigkeit  $v_{\text{max}}$  und Trödelwahrscheinlichkeit  $p$



$v_{\text{safe}} := v_{\text{leader}} + (g - v_{\text{leader}})/(r + 1)$   
 mit  $r := (v + v_{\text{leader}})/2b$   
 $v := \min\{v + a, v_{\text{max}}, v_{\text{safe}}\}$   
 $v := \max\{0, v - pa\}$   
 $x := x + v$

mit maximaler Geschwindigkeit  $v_{\text{max}}$ , Beschleunigungsvermögen  $a$ , Bremsvermögen  $b$  und Trödelparameter  $p$

Abbildung 2-1: Verkehrsmodelle

Ein Unterschied zwischen diesen beiden Modellen liegt darin, dass beim CA die Straße in einzelne Zellen unterteilt wird. Die Bewegung findet hier entlang der Zellen statt, wobei jede Zelle maximal ein Auto enthalten darf. Beim SK-Modell hingegen besteht diese Einteilung nicht, die Verkehrssimulation behält aber dennoch die Zellen bei (das hat Vorteile beim Suchen des nächsten Fahrzeuges vor dem betrachteten) und verwendet einen kontinuierlichen Offset in den Zellen. Ein Fahrzeug kann sich an einer beliebigen Position auf einer Kante befinden, solange es nicht mit anderen Fahrzeugen kollidiert. Im Gegensatz zum CA-Modell ist hier das Bremsvermögen begrenzt.

Mit diesen Modellen lassen sich viele Eigenschaften des Verkehrs wie das Spurwechselverhalten und die Dynamik von Staus realistisch abbilden. Da auf Einzelfahrzeugsbasis gearbeitet wird, benötigen die Modelle für große Straßennetze mit realistischer Verkehrsmenge eine hohe Rechenleistung. Die damit verbundenen großen Datenmengen, die zur Visualisierung notwendig sind, lassen sich jetzt erstmals aufgrund der hohen Bandbreite des GTB in Echtzeit übertragen.

### 2.3.2 Abspalten der Visualisierung

Bevor die Visualisierung von der Simulation getrennt werden konnte, mußte im ersten Schritt eine Kommunikationsschnittstelle geschaffen werden. Dies geschah mit Hilfe einer eigenen Klasse, welche zwischen Simulation und Visualisierung vermittelt. Der Vorteil dieser Lösung ist, dass man so die Simulation und Visualisierung unabhängig von der verwendeten Netzstruktur implementieren kann. Daraus ergibt sich eine höhere Flexibilität, wenn andere Netze zugrunde gelegt werden.

Die Trennung von Simulation und Visualisierung, welche aufgrund der komplexen Struktur von PLANSIM-T und der tiefen Verankerung der vorher vorhandenen Visualisierung in die Programmstruktur nur schwierig durchzuführen war, ermöglichte es nun, im zweiten Schritt die Visualisierung endgültig abzuspalten und autonom auszuführen.

Hieraus ergeben sich mehrere Vorteile. Zum Einen können nun mehrere unterschiedliche Visualisierungen auf verschiedenen Rechnern gleichzeitig von einer Simulation Daten empfangen und so unterschiedlichen Nutzerkreisen die jeweils benötigten Daten zur Verfügung gestellt werden. Die Visualisierungen können zu jedem Zeitpunkt der Simulation zu- und abge-

schaltet werden, ohne dass dies einen Einfluß auf die Simulation selbst hat. Schließlich wird die Simulation nun nicht mehr durch die Visualisierung gebremst, was insbesondere bei großen Netze für sehr lange Laufzeiten sorgte.

### 2.3.3 Die Kommunikationsschnittstelle

Um die Daten aus der Simulation zu übermitteln, wurde eine Kommunikationsschnittstelle zwischen diesen beiden Programmen geschaffen. Die hierzu erstellte Klasse kann auf Seite der Simulation sowie der Visualisierung eingebettet werden und ist in der Lage, die Daten sowohl mit ATM als auch mit IP Multicast auszutauschen.

Im Wesentlichen besteht die Kommunikationsschnittstelle aus einem Sender und einem oder mehreren Empfängern. Der Sender kommuniziert lokal mit der Simulation und die Empfänger mit den zugehörigen Visualisierungen.

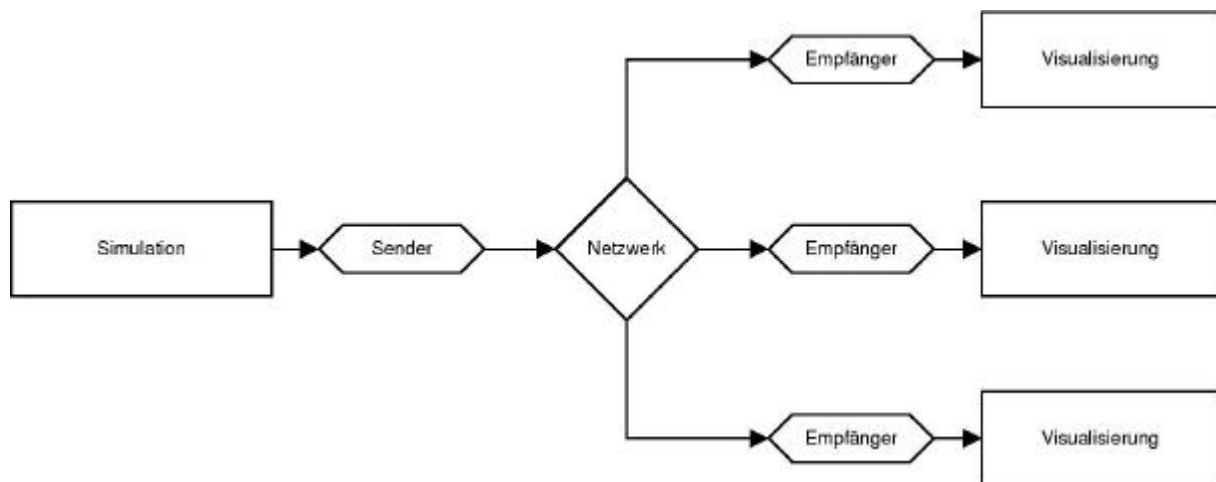


Abbildung 2-1: Verteilung der Simulationsdaten an mehrere Empfänger zur Visualisierung

Sowohl Sender als auch Empfänger besitzen einen eigenen Puffer, um Daten zwischen zu speichern, die nicht schnell genug gesendet bzw. an die Visualisierung weitergeleitet werden können. So können Störungen innerhalb der Kommunikation aufgefangen werden, wenn z.B. eine große Netzlast eines anderen Programmes das Versenden der Daten behindert oder die Visualisierung nicht schnell genug arbeitet.

Zusätzlich enthält der Empfänger einen Filter, der an die Visualisierung nur die benötigten Daten weiterreicht. Dies ist nötig, da selbst auf den zur Verfügung stehenden Computern die Darstellung der Daten am zeitaufwendigsten war. Ob dies auch bei Computern mit modernster Graphikhardware ein Problem darstellt, hängt stark von der speziellen Visualisierung ab und muß im Einzelfall untersucht werden.

Aufgrund der Flexibilität und Modularität dieser Lösung ist es einfach möglich, weitere Klienten zu entwickeln, welche die Daten auf unterschiedliche Weise darstellen. Man ist also nicht auf die im Projekt entwickelten 2D und 3D-Visualisierungen beschränkt.

### 2.3.4 Mehrprozessorsysteme

Um zusätzlich die Leistung der Simulation zu verbessern wurde PLANSIM-T erweitert, um Mehrprozessorsysteme mit verteiltem Speicher besser ausnutzen zu können [RW].

Hierbei wird das Straßennetz in verschiedene disjunkte Bereiche unterteilt die sich an möglichst wenig Knoten berühren sollen, wobei darauf geachtet wird, dass die Bereiche möglichst gleich viele Kanten haben.

Jeder dieser Bereiche wird von einem eigenen Cluster simuliert. Kommunizieren müssen die einzelnen Prozesse nur an den Schnittknoten, deren Anzahl gering gehalten wurde.

Je nach Größe der verwendeten Netze und der Leistungsfähigkeit der CPU ermöglicht dieses Verfahren bei kleiner Anzahl der CPU  $n$  einen fast linearen Geschwindigkeitszuwachs mit  $n$ . Danach fällt der Leistungsgewinn stärker ab, da die Kommunikation zwischen den einzelnen Rechnern immer mehr Rechenzeit in Anspruch nimmt.

In der folgenden Grafik sieht man die Leistungssteigerung welche man für ein großes Netz mit über 50.000 Kanten und 65.000 Knoten erhält.

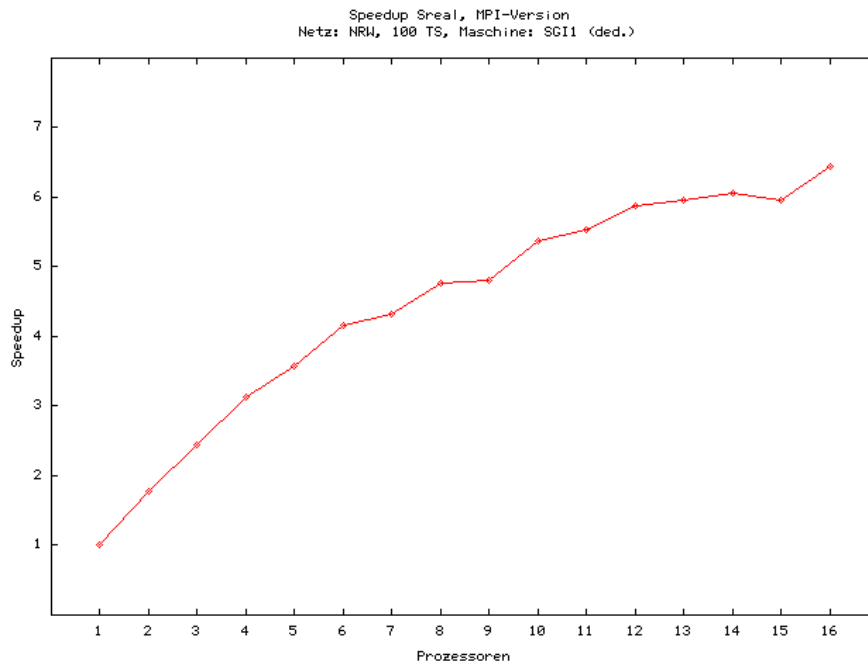


Abbildung 2-1: Leistungssteigerung durch Parallelisierung

Die größte Schwierigkeit bei der Verwendung von Mehrprozessorsystemen ist die Aufteilung des Straßennetzes. Das verwendete Verfahren erreicht dieses Clustering jedoch in einer, gemessen an der Laufzeit der Simulation, geringen Zeit wodurch dieser Schritt nicht weiter ins Gewicht fällt.

## 2.4 Visualisierung

Die bisherige Visualisierung von PLANSIM-T war fest in die Simulation eingebettet. Dies brachte die oben beschriebenen Probleme mit sich und machte als Lösung die Abspaltung der Visualisierung notwendig.

Konkret wurden in diesem Projekt zwei Visualisierungen entwickelt. Zum Einen wurde eine 2D-Visualisierung erstellt, welche vom Layout weitgehend identisch ist zu der in PLANSIM-T fest eingebauten Darstellungsmethode ist. Zum Anderen wurde von *echtzeit GmbH* eine 3D-Visualisierung entwickelt und an die Simulation, mit Hilfe der entwickelten Kommunikationsschnittstelle, angekoppelt. Das ist in Abschnitt 3 genauer beschrieben.

### 2.4.1 2D-Visualisierung

Die 2D-Visualisierung wurde, auf der Basis der in PLANSIM-T eingebauten Visualisierung, unter Beibehaltung des Layout neu entwickelt. Hierbei wurde auf GTK zurückgegriffen, eine Bibliothek zum Erstellen von grafischen Benutzeroberflächen, da diese Bibliothek für alle UNIX-Plattformen kostenlos erhältlich ist und somit inzwischen auf vielen Rechnern installiert wurde.

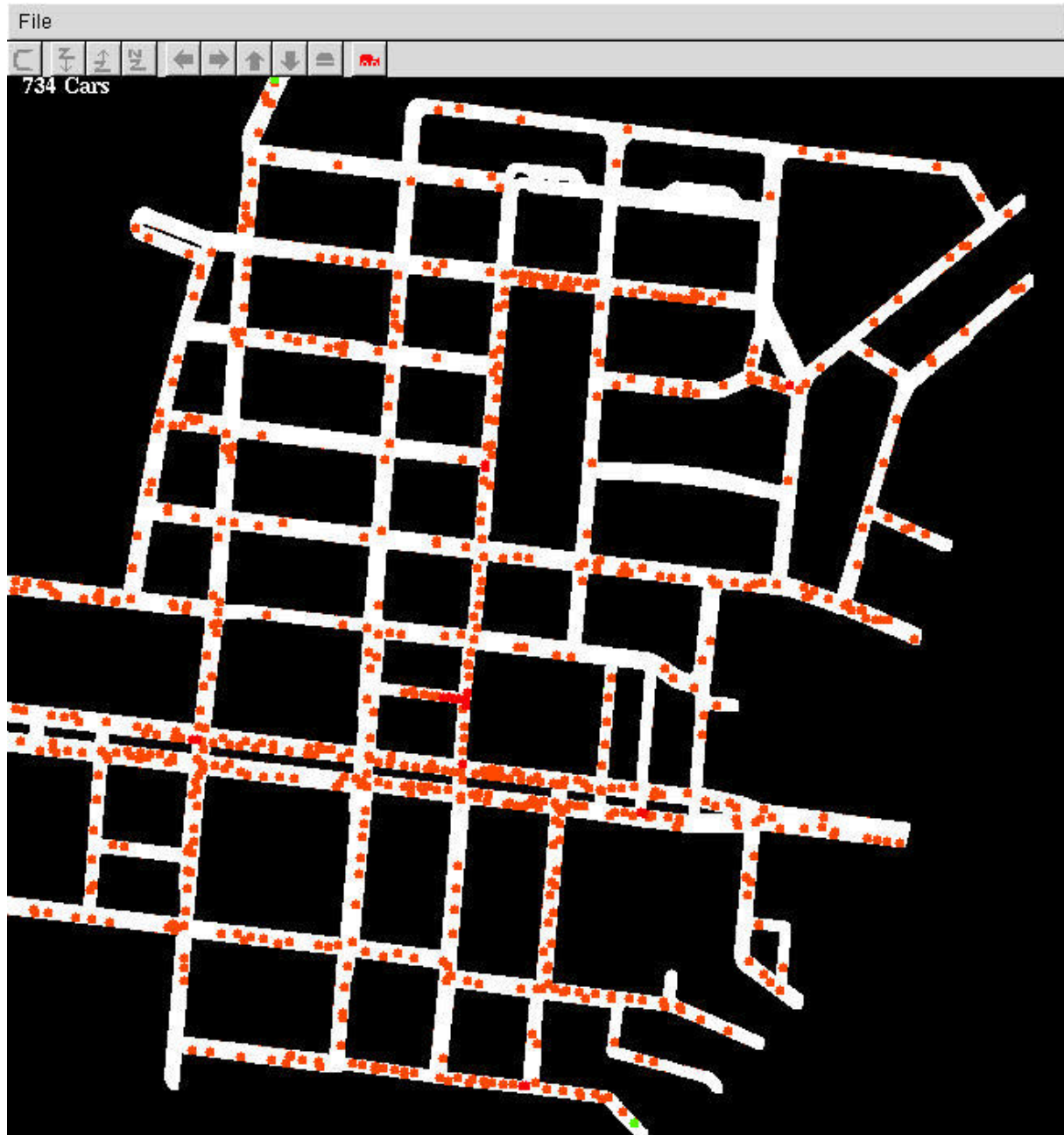


Abbildung 2-1: 2D-Visualisierung eines Berliner Stadtteils

Gleichzeitig wurden verschiedene Verbesserungen implementiert wie z.B. die Kopplung mit der Kommunikationsschnittstelle, um flexibler auf zukünftige Entwicklungen von Netzprotokollen reagieren zu können. Auch mußte hier der oben erwähnte Filter beim Empfänger berücksichtigt werden, um einen minimalen Datenstrom zwischen Empfänger und Visualisierung zu erhalten.

Abbildung 2-1 zeigt eine 2D-Visualisierung des zugrunde liegenden Straßennetzes, ein kleiner Ausschnitt aus dem Straßennetz von Berlin. Auf diesem Straßennetz werden die einzelnen Autos als Punkte dargestellt. Es ist also direkt zu erkennen wo z.B. der Verkehr lang läuft, wo sich Staus bilden und welche Straßen wenig oder gar nicht benutzt werden.

Innerhalb des Stadtplans kann man Bereiche auswählen, die sich dann vergrößert darstellen lassen. Somit ist es möglich, bestimmte Verkehrsschwerpunkte auszuwählen und näher zu betrachten.

Zusätzlich wäre es denkbar, dass statistische Daten über den gesamten Verkehrsfluß oder über einzelne Straßen eingeblendet werden. Auch könnte man einzelne Straßen farblich darstellen, um so die Dichte oder andere Informationen über eine Straße ablesen zu können. Mit Hilfe der zugrunde liegenden Struktur von Sender und Empfänger kann die Visualisierung leicht auf diese Darstellungsmöglichkeiten erweitert werden.

Neben der reinen Darstellung bietet die Visualisierung noch die Möglichkeit über einen Rückkanal mit der Simulation zu kommunizieren. Mittels verschiedener Buttons kann man die Simulation dazu veranlassen, keine Daten mehr zu senden, die Netzdaten neu zu schicken oder vorgegebene Straßen zu sperren. Auch eine Selektionsmöglichkeit für bestimmte Straßen ist leicht zu realisieren. Diese könnten dann der Simulation gemeldet werden mit der Aufforderung, eine bestimmte Aktion durchzuführen, wie z.B. das Setzen einer Höchstgeschwindigkeit auf diesen Straßen oder eine Sperrung.

## **2.5 Der Rückkanal**

Um die Kommunikation zwischen der Visualisierung und der Simulation zu ermöglichen, wurde ein Rückkanal implementiert. Diese Kommunikation ist z.B. erforderlich, damit die Visualisierung die von ihr benötigten Daten anfordern kann. So wird das Straßennetz getrennt von den Autodaten angefordert, da dieses sich während des Simulationslaufes nicht mehr ändert und somit ständige Übertragungen nur unnötige Netzlast erzeugen würden.

Zusätzlich ermöglicht es der Rückkanal dem Benutzer mit Hilfe der Visualisierung in das Simulationsgeschehen einzugreifen. So gibt es in der Simulation die Möglichkeit, Szenarios zu erstellen, in der einzelne Situationen wie die Sperrung von Straße und das Setzen von Höchstgeschwindigkeiten in einzelnen Bereichen festgelegt werden. Mit Hilfe der Visualisierung veranlaßt der Benutzer die Simulation, die Szenarios schrittweise durchzuspielen. Das Ergebnis lässt sich direkt am Bildschirm verfolgen.

Diese Möglichkeit ist besonders wichtig, wenn eine Verkehrs-System-Management-Zentrale (VSM) schnell auf aktuelles Verkehrsgeschehen reagieren will. So kann festgestellt werden, wie sich z.B. nach einem Unfall die Sperrung einer bestimmten Straße auswirken würde und welche Maßnahmen zusätzlich nötig werden. Auch lassen sich so sehr einfach und schnell die Auswirkungen von straßenbaulichen Maßnahmen wie z.B. Tempo-30 Zonen, die Einführung von Busspuren usw. untersuchen.

### **2.5.1 Umsetzung**

Der Rückkanal konnte nicht mit dem normalen Datenkanal zusammengelegt werden, da auf einer Multicast-Verbindung nur Kommunikation in eine Richtung möglich ist. Deshalb wird eine eigene TCP/IP-Verbindung zwischen Visualisierung und Simulation aufgebaut, welche diese Richtung der Kommunikation übernimmt.

Es werden zur Zeit nur die wichtigsten Kommandos unterstützt, um die Funktionalität des entwickelten Rückkanals zu demonstrieren. Der Befehlssatz ist jedoch auf einfache Weise erweiterbar und es lassen sich weitere, wesentlich komplexere Anweisungen definieren, welche sich durch den Rückkanal realisieren lassen. Diese Möglichkeiten werden nur insofern begrenzt, als das diese nicht den Ablauf des zugrunde liegenden Modells verändern dürfen.

Der wichtigste Befehl veranlasst die Visualisierung, überhaupt Daten übers Netz zu senden. Beim Starten der Simulation ist dies ausgeschaltet, da es keinen Sinn macht die Daten zu versenden, wenn kein Klient angeschlossen ist, der die Daten überhaupt empfangen kann. Es würde nur eine unnötige Netzlast erzeugt. Mit dem gleichen Befehl lässt sich auch das Senden wieder abschalten, um somit einen Zeitraum zu überbrücken, in der keine Visualisierung gewünscht ist. Damit erhält man die Möglichkeit, zwischendurch den Verlauf der Simulation zu kontrollieren, ohne in der Zwischenzeit das Netz durch eine überflüssige Visualisierung zu belasten.

Ein weiterer Befehl fordert von der Simulation die Daten über das Straßennetz an. Normalerweise werden nur die Fahrzeugdaten verschickt, um die Netzbelastung zu minimieren. Wenn nun eine Visualisierung gestartet wird, fordert sie die Straßendaten solange an, bis ein komplettes Netz vorliegt. Da sich dieses nicht mehr ändert, wird bis zum Ende der Visualisierung mit diesem Netz gearbeitet.

Weitere Befehle führen innerhalb der Simulation vorgefertigte Aktionen aus. So kann man z.B. ausgewählte Kanten sperren lassen oder den Zufluss des Verkehrs begrenzen und erweitern. Dies wird in folgendem Beispiel demonstriert.

### **2.5.2 Ein Beispiel: Sperrung einer Straße**

Um die Möglichkeiten des Rückkanals zu demonstrieren wurde ein Beispielszenario erstellt. Die Basis bildet ein Teilausschnitt Berlins. Die Verkehrsströme zwischen den Quellen und Senken wurden in Form einer Start/Zielmatrix vorgegeben. Während des Simulationslaufes wird eine festgelegte Straße für den Verkehr gesperrt. Am einfachsten lässt sich die Wirkung bei Betrachtung eines einzelnen Start / Zielpaars betrachten.

Auf den folgenden zwei Bildern sieht man den Verkehrsfluss solch eines Start / Zielpaars, dessen kürzester Weg entlang der markierten, zu sperrenden Kante verläuft. Man erkennt den Verkehrsfluß entlang des kürzesten Weges, einige Autos nehmen einen individuellen Weg.

Auf dem rechten Bild wurde die Straße gesperrt. Alle Autos die noch nicht an dieser Straße angekommen sind, müssen sich von dem gegenwärtigen Standpunkt aus einen neuen Weg suchen. Nach einer kurzen Einschwingphase, in der sich alle Fahrzeuge die zur Zeit der Sperrung schon unterwegs waren einen individuellen Weg suchen, kommt es zu zwei Hauptverkehrsströmen welche nur Streckenweise mit dem alten Fluß übereinstimmt. Ein Strom fließt oberhalb der gesperrten Straße entlang, während der andere sich seinen Weg unterhalb gesucht hat.

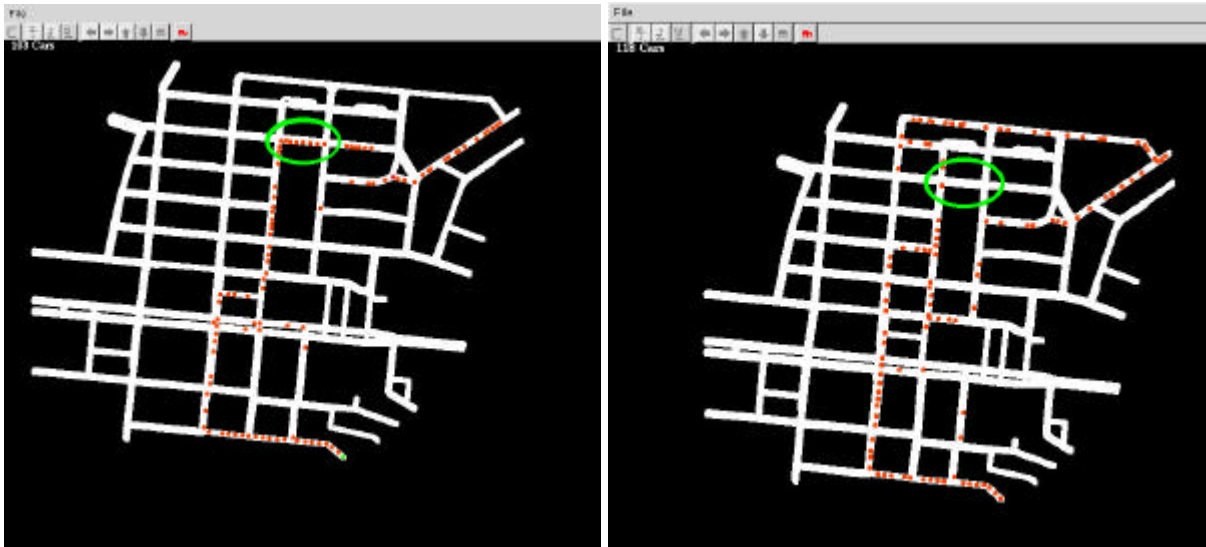


Abbildung 2-1: Verkehrsfluss eines Start / Zielpaares

Hebt man nun nach einer gewissen Zeit diese Sperrung wieder auf, so folgt wiederum eine kurze Einschwingphase. Danach erhält man wieder den anfängliche Verkehrsfluß, wie er vor der Straßensperrung war.

Nun wird die Simulation erneut gestartet, diesmal jedoch lässt man den Verkehr über das ganze System fließen. Jetzt sind die Auswirkungen der Straßensperrung nicht mehr ganz so einfach zu erkennen.

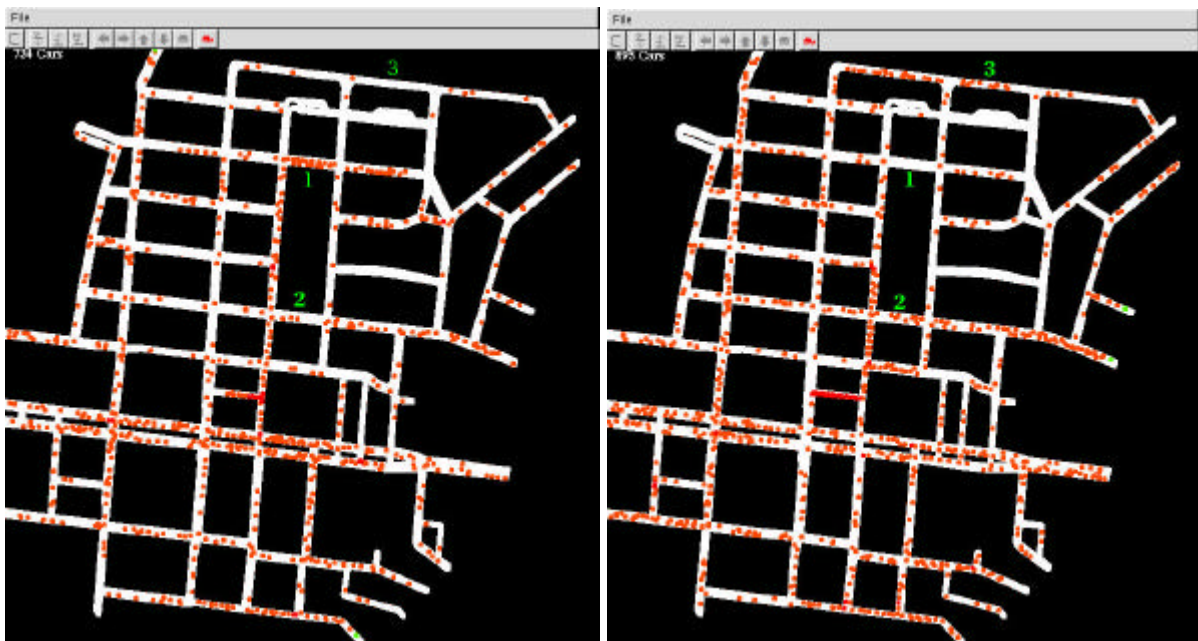


Abbildung 2-2: Verkehrsfluss über das gesamte System

Man kann jedoch den Fluß, welcher an den mit 1, 2 und 3 beschrifteten Kanten vorliegt, in einem Diagramm betrachten. In diesem Beispiel wird die Kante 1 alle 500 Zeitschritte für den Durchgangsverkehr gesperrt und nach weiteren 500 Zeitschritten wieder freigegeben.

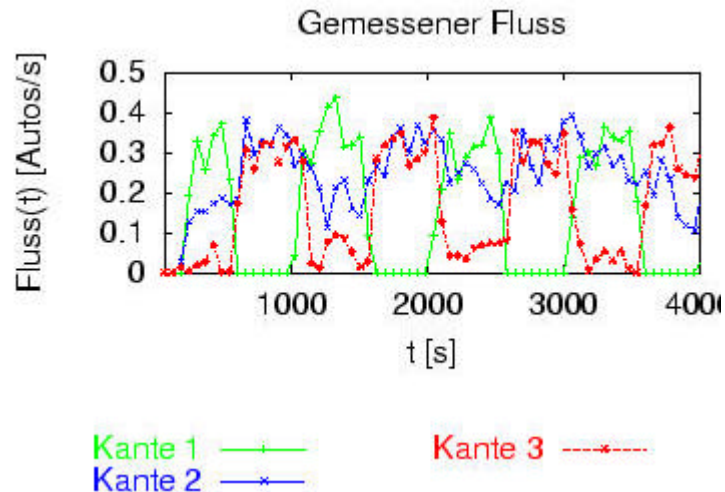


Abbildung 2-3: Gemessener Fluss auf 3 Kanten des Systems

Anhand der grünen Linie ist gut zu erkennen, dass der Fluß zu den gesperrten Zeiten fast sofort bis auf Null abfällt. Gleichzeitig steigt die Verkehrsbelastung auf der Ausweichstrecken über die Kanten 2 (blau) und 3 (rot) deutlich an.

## 2.6 Zusammenfassung und Ausblick

Dieser Beitrag demonstriert anhand eines einfachen Verkehrsszenarios die Eignung des verwendeten Ansatzes zur bidirektionalen Kommunikation zwischen Verkehrssimulation und Visualisierung, die auf verschiedenen Rechnern laufen können. Auch wurde die Simulation um eine Komponente für verteiltes Rechnen erweitert, die allerdings aufgrund der rasch wachsenden CPU-Leistungen nur für sehr große Netz interessant wird. Für die Zwecke dieses Projektes war die SMP-Variante von PLANSIM-T ausreichend. Im nächsten Abschnitt wird eine sehr viel komplexere Visualisierung an die Simulation angekoppelt, während in Abschnitt 4 den Fragen nach den erreichten Bandbreiten nachgegangen wird.



## 3 Echtzeit

### Ansprechpartner

Kay-Ingo Ahlers

E-mail: kayingo@echtzeit.de

### 3.1 Einleitung

Das Arbeitsfeld der echtzeit AG ist in zwei Teile zu gliedern. Der eine Teil umfasst die Implementierung eines Modellierungswerkzeuges zum Erstellen eines Verkehrsnetzes. Der andere enthält die Implementierung eines VR-Visualisierungsclients, der die simulierte Verkehrsbewegung in einer dreidimensionalen Umgebung darstellt.

### 3.2 Modellierungswerkzeug für Verkehrsnetze

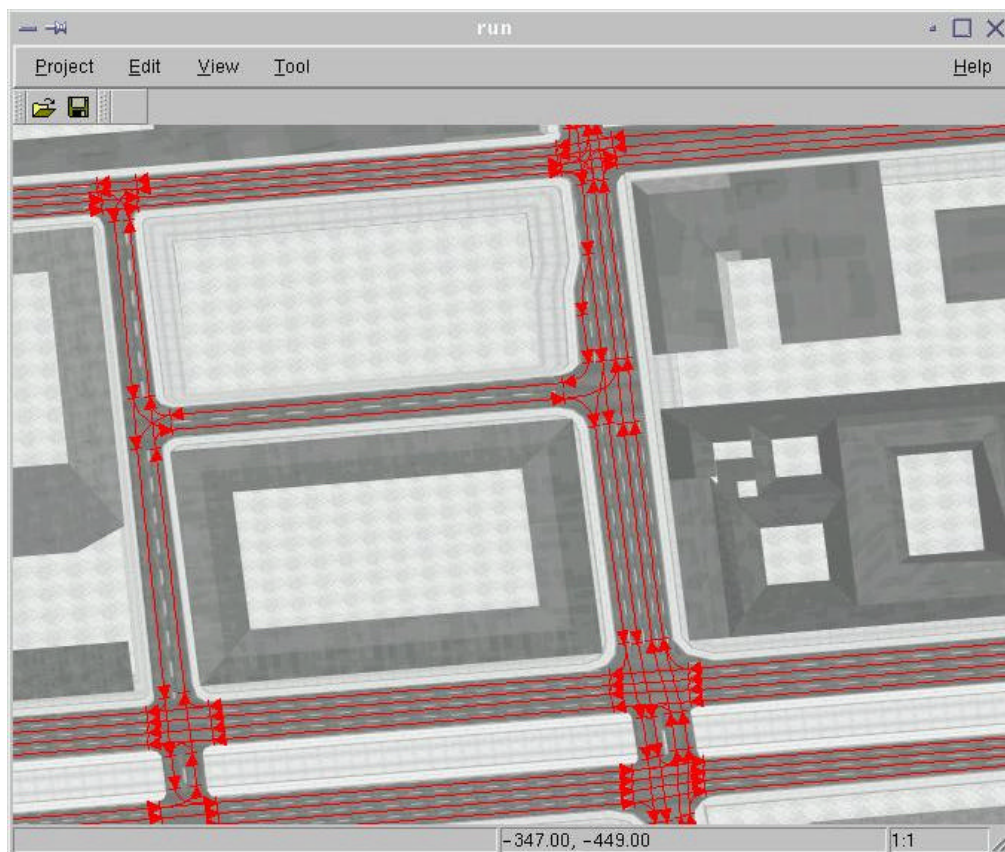


Abbildung 3-1: Modellierungswerkzeug

Für die Erstellung und Bearbeitung der Netzdaten wurde ein UNIX/C/Motif basierter Editor entwickelt. Der Editor ermöglicht die Definition und Verknüpfung von Linien- und (linear

interpolierten) Kurvensegmenten zu einem Graphen. Die einzelnen Segmente können mit Attributen wie Verkehrszeichen und Ampeln versehen werden. Im Hintergrund wird dazu als Vorlage eine digitalisierte Karte eingeblendet.

Mit Hilfe des Editors wurden die Netzdaten für ein bestehendes Modell der Berliner Innenstadt erarbeitet.

Nach Konvertierung der Daten aus dem Editor in das vereinbarte Dateiformat können die Verkehrsnetzdaten in den von der Uni Köln entwickelten Simulationsserver eingelesen werden.

### 3.3 VR-Visualisierungsclient

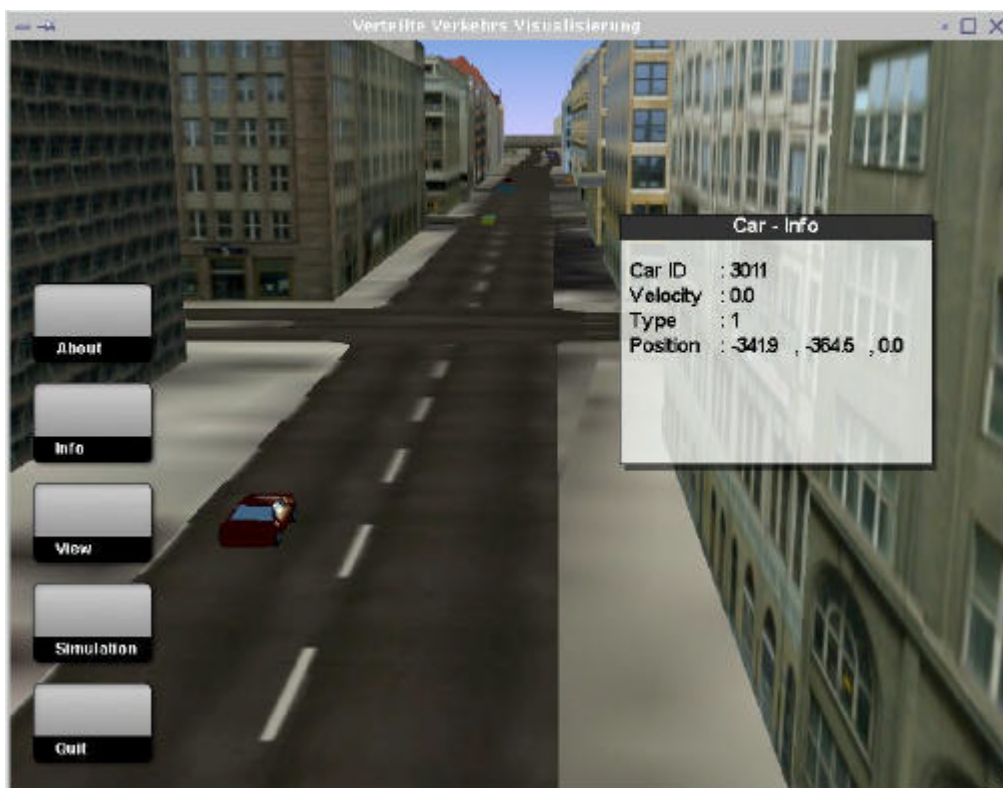


Abbildung 3-1: 3D-Visualisierung einer Berliner Straße

#### 3.3.1 Laufzeitumgebung

Das Zielbetriebssystem, auf dem der VR-Visualisierungsclient ausgeführt werden soll, ist IRIX bzw. Linux. Dies ist in der Verwendung der Grafikschnittstelle „Performer“ begründet. Performer ist eine von SGI entwickelte IRIX-Grafikbibliothek, die speziell auf die Darstellung komplexer Szenen zugeschnitten ist. Einen großen Vorteil brachte die Freigabe der Performer-Version für das PC-Betriebssystem Linux. Damit ist es uns nun möglich, die VR-Visualisierung der Simulation auf normalen PCs darzustellen. Eine schnelle Netzwerkanbindung von Linux Maschinen kann mit Gigabit Ethernet Karten realisiert werden.

### **3.3.2 Datenaustausch**

Der Datenaustausch zwischen dem Netzwerk und der Darstellungseingine erfolgt über einen Datenaustauschpuffer. Dieser ist notwendig, weil die Darstellungsschritte mit den Berechnungsschritten des Simulationsserver zeitlich nicht übereinstimmen. Die Applikation wurde so implementiert, dass ein eigener Netzwerk-Thread alle Daten aus dem Netz liest und sie in den Datenaustauschpuffer schreibt. Der Main-Thread, der auch für die Visualisierung zuständig ist, liest die Daten aus diesem Puffer und arbeitet sie für die Visualisierung auf.

## **3.4 Visualisierung**

Die Visualisierung erfolgt zum einen durch die Darstellung der statischen Szenenelemente wie Gebäude, Straßenzüge, Bäume und Verkehrszeichen. Über eine Konfigurationsdatei kann eingestellt werden, welche Geometrien geladen werden sollen.

In dieser Datei werden zum anderen die Geometrien für die nicht statischen Elemente, in unserem Fall die Fahrzeuge, festgelegt. Nach dem Laden aller in der Szene enthaltenen Geometrien werden Referenzen auf die Fahrzeuggeometrien erzeugt, die vom Typ der über das Netz übertragenen Fahrzeugdaten abhängig sind. Jede Referenz wird noch mit zusätzlichen Informationen versehen, wie z.B. ID, Geschwindigkeit oder Status. Diese Referenzen auf die Fahrzeuggeometrien werden anhand der Positions- und Orientierungsdaten aus dem Simulator in der VR-Szene platziert.

### **3.4.1 Steuerung**

Es wurde eine Menüführung entwickelt, um Informationen über die Simulation und die Navigation in der virtuellen Welt auszugeben. Mit Hilfe der Menüs können Details über das zur Zeit betrachtete Fahrzeug oder Informationen über die aktuelle Verbindung zum Simulationsserver angezeigt werden. Die Menüs sind über Konfigurationsdateien frei gestaltbar. Diese Konfigurierbarkeit hält es offen, die Ausgabe zu erweitern und frei zu gestalten.

Die Navigation durch die VR-Szene ist an die Steuerung bekannter VR-Szenenbetrachter angelehnt, so dass eine intuitive Navigation möglich ist. Sie kann mit der Tastatur, vorrangig jedoch mit der Maus erfolgen. Das Selektieren eines gewünschten Fahrzeuges erfolgt ebenfalls durch Anklicken mit der Maus. Nach der Selektion eines Fahrzeuges folgt die virtuelle Betrachterkamera diesem Fahrzeug. Der Betrachter hat die Möglichkeit, mittels Tastatureingabe die Blickrichtung zu beeinflussen.

### **3.4.2 Stadtmodell**

Das Stadtmodell ist die Nachbildung eines Stadtteils der Berliner Innenstadt. Dieses Szenario bietet eine gute Kulisse für ein stark befahrenes Verkehrsnetz. Die Komplexität der Szene ist auf die Darstellung in der Echtzeitumgebung ausgelegt, so dass eine angemessene Detaillierung gewählt wurde. Eine Nachbildung des real existierenden Verkehrsnetzes lässt schon in der Entwicklungsphase Probleme und Besonderheiten besser erkennen als ein fiktiv erzeugtes Netz.

### **3.5 Resümee**

Die Entwicklung des VR–Betrachters ergab das Resultat, dass eine wissenschaftliche Simulation für eine VR–Visualisierung speziell angepasst werden muss. So zeigte sich, dass die gegenwärtig vorliegende Konfiguration zwar wichtige Informationen für die Verkehrsanalyse liefert, dabei aber leider nicht die Darstellungsqualität erreicht wird, die für Virtual Reality erforderlich ist. Grund dafür ist die große Schrittweite, in der die Verkehrsdaten auf dem Simulationsserver berechnet werden.

### **3.6 Ausblick**

Um eine flüssige Darstellung in der virtuellen Umgebung zu erreichen, sollen noch verschiedene Interpolationsverfahren getestet werden. Die Synchronisationsprobleme können durch einen verbesserten Informationsaustausch im Rückkanal vermindert werden. Mit Hilfe dieser beiden Maßnahmen eine wesentlich realistischere Darstellung zu erwarten.

Eine mögliche Weiterentwicklung des Projektes ist es, die Interaktionsmöglichkeiten für den Betrachter auszubauen, um dem Betrachter größere Einflussmöglichkeiten auf das Verkehrsgeschehen zu gestatten. Des weiteren soll die Simulation um das korrekte Verhalten der Fahrzeuge an Kreuzungen, Kurven und Einmündungen erweitert werden. Dadurch würde die VR–Darstellung harmonisiert werden.

## 4 GMD – Forschungszentrum Informationstechnologie

### **Ansprechpartner**

Lothar Zier

E-Mail: Lothar.Zier@gmd.de

Telefon: 02241 14 2943

Telefax: 02241 14 2071

Adresse: GMD, IMK.HSN, 53754 Sankt Augustin

### **Weitere Mitarbeiterinnen im Projekt:**

Gundula Dörries, E-mail: Gundula.Doerries@gmd.de

Eva Pless, E-mail: Eva.Pless@gmd.de

Ursula Eisenblätter, E-mail: Ursula.Eisenblaetter@gmd.de

Martha Merzbach, E-mail: Martha.Merzbach@gmd.de

### **4.1 Übersicht**

Im Rahmen des 3V-Projektes ([3V]) baute die GMD eine Multicast-Infrastruktur innerhalb des Gigabit Testbed (GTB) West ([GIGAW]) auf. Zielsetzung war, die Verteilung von Daten eines Senders an eine kleine Gruppe von Empfängern mit hohen Bandbreiten zu ermöglichen, Leistungstests durchzuführen und Grundlagen für die Konfiguration von Gigabit-Multicastnetzen zu ermitteln.

Multicast ist die Verteilung von Daten an mehrere Empfänger, wobei ein Sender die Daten nur einmal verschicken muss. Im Rahmen des 3V-Projektes werden folgende Annahmen getroffen:

- es gibt einen Sender und n Empfänger, wobei n klein ( $< 20$ ) ist
- die Verteilung erfolgt auf Basis von ATM Punkt-zu-Mehrpunkt Verbindungen oder IP-Multicast
- die Datenverteilung erfolgt ungesichert
- ein einzelner Datenstrom erfordert eine hohe Bandbreite ( $>100\text{Mbit/s}$ ), so dass in der Regel der Sender mit der Duplizierung der Daten überlastet wäre

Die beiden wesentlichen Ziele sind das Bereitstellen von Netztechniken für hohe Multicast-Übertragungsraten und die Sicherstellung eines störungsfreien Betriebs der Produktionsnetze. Multicast ist aufgrund seiner Struktur (kein Regelmechanismus für Congestion Management und Verteilung an viele Empfänger) ein äußerst aggressiver Datenverkehr. Dies kann in Kombination mit den im Projekt geforderten hohen Bandbreiten zu massiven Störungen in den bestehenden Netzen führen.

Kapitel 4.2 beschreibt die realisierten Verfahren für die Multicast-Verteilung, Kapitel 4.3 die durchgeführten Tests zur Leistungsbewertung. Kapitel 4.4 zieht ein Fazit. Im Anhang des Abschlussberichtes sind die verwendeten Konfigurationen beschrieben.

## 4.2 Methoden für Multicast-Verteilung

In diesem Kapitel werden Methoden für die Verteilung von Multicast-Daten mit hohen Übertragungsraten sowie gegebenenfalls ihre Realisierung im GTB West beschrieben. Die Methoden sind im Einzelnen:

- Direkte Nutzung von ATM Punkt-zu-Mehrpunkt Verbindungen über die XTI-Programmierschnittstelle.
- Nutzung von LAN Emulation für die Verteilung von IP-Multicast-Paketen innerhalb eines IP-Netzes.
- Routing von IP-Multicast-Paketen.
- Sonstige Verfahren: MARS, proprietäre Verfahren von Cisco und FORE.

Als Verteilungsmodell wird ein Sender mit n Empfängern (1-n Verteilung) angenommen.

Die im GTB West im Rahmen von Multicast eingesetzte Hardware beruht auf folgenden Geräten:

- Workstations der Firmen Sun und SGI
- ATM Adapter der Firmen FORE/Marconi<sup>1</sup> und Sun mit 622 Mbit/s
- Gigabit Ethernet Adapter der Firmen Sun, SysKonnnect, Packet Engines (jetzt Alcatel)
- ATM Switche der Firmen FORE (ASX-1000/4000) und Cisco (LS1010, C8540MSR<sup>2</sup>)
- Ethernet Switche der Firma Cisco (C5505 und C6509) mit 622 Mbit/s ATM Interface
- IP Router der Firma Cisco (C8540MSR) mit 622 Mbit/s ATM und mit Gigabit Ethernet Schnittstellen

### 4.2.1 Native ATM

ATM erlaubt die 1-n Verteilung von Daten mittels Punkt-zu-Mehrpunkt Verbindungen. Diese werden von allen durch die vom ATM Forum definierten Schnittstellen (UNI und IISP 3.0/3.1/4.0, PNNI 1.0) unterstützt, wobei zur Zeit Punkt-zu-Mehrpunkt Verbindungen nur vom Sender aus initiiert werden können. UNI 4.0 sieht Empfänger initiierte Punkt-zu-Mehrpunkt Verbindungen vor; dies ist aber in der Praxis noch nicht umgesetzt. Da Punkt-zu-Mehrpunkt Verbindungen wichtige Voraussetzung für LAN Emulation sind, werden sie von allen Herstellern im LAN-Bereich unterstützt. Gab es mit älteren Geräten noch teilweise Probleme, so kann bei neueren Geräten (wie bei den im Gigabit Testbed West eingesetzten FORE ASX-1000/4000 mit D-Modulen und Cisco C8540MSR) davon ausgegangen werden, dass der Aufbau von Punkt-zu-Mehrpunkt Verbindungen und die Verteilung der Daten mit Leitungsgeschwindigkeiten problemlos möglich ist.

Die XTI-Schnittstelle ([XTI]) ist eine von der X-Open Gruppe standardisierte Programmierschnittstelle für UNIX basierte Systeme. Sie enthält eine Erweiterung, die ATM-Verbindungen von Anwendungen aus direkt nutzbar macht. Die Funktionalität von XTI mit ATM ähnelt der IP-UDP-Nutzung; der Hauptunterschied liegt darin, dass der Sender alle Empfänger kennen und jeden einzeln zum Verteilungsbaum hinzufügen muss. Eine Umstel-

---

<sup>1</sup> Die Firma FORE wurde 1999 von einer englischen Firma übernommen und ist nun Teil der Firma Marconi über die die erwähnten Produkte erhältlich sind. Im Rahmen dieses Berichtes wird der alte Firmennamen verwendet.

<sup>2</sup> Der C8540MSR kann gleichzeitig als ATM Switch, Ethernet Switch und IP-Router arbeiten. Im Rahmen des Projektes wurde er als ATM Switch und IP Router eingesetzt.

lung bestehender IP-Multicast Programme ist in der Regel mit vertretbarem Aufwand möglich.

Die Firma FORE hat in ihrer Adapter Software für UNIX basierte Systeme die XTI-Schnittstelle realisiert. ATM Punkt-zu-Mehrpunkt Verbindungen können somit über die XTI-Schnittstelle von allen Sun und SGI Maschinen mit FORE Interface genutzt werden. Auf ATM Adaptionern der Firma Sun wird die Schnittstelle nicht unterstützt.

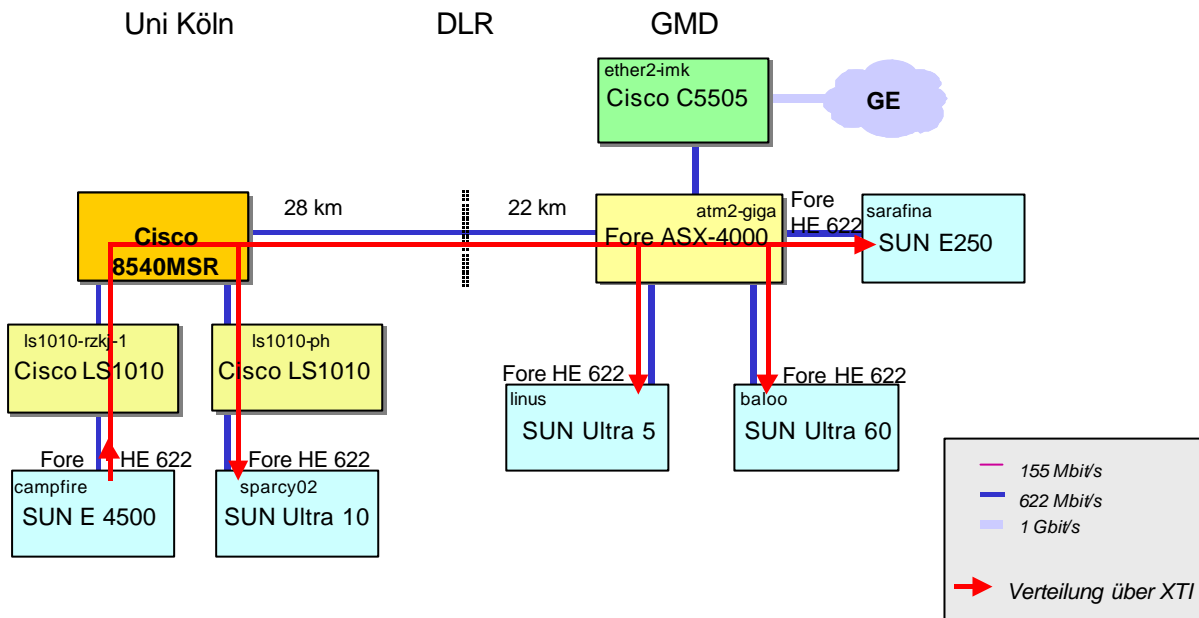


Abbildung 4-1: XTI Punkt-zu-Mehrpunkt Verteilung

Aus Sicht der 3V-Anwendung hat die Nutzung der XTI-Schnittstelle für die Multicast-Übertragung folgende Vorteile:

- Es wird eine optimale Verteilung der Multicast-Daten erreicht. Jeder Link des Netzes wird höchstens einmal belastet; es wird der kürzeste Weg zum Empfänger genommen; die Daten werden nur an die Endgeräte geschickt, wo auch die entsprechende Anwendung aktiv ist.
- Es können beliebige Pakete bis zu einer Größe von 64KB ohne Fragmentierung verschickt werden. Dies ermöglicht hohe Übertragungsraten.

Diesen Vorteilen stehen folgende Nachteile/Probleme gegenüber:

- Die Schnittstelle verlangt eine vollständige ATM-Infrastruktur zwischen allen beteiligten Systemen. Insbesondere ist keine Nutzung von Gigabit Ethernet möglich.
- Die Schnittstelle wird nur von einem Hersteller (FORE) unterstützt.
- Eine Weiterentwicklung findet nicht mehr statt. Eine Socket-Schnittstelle ist zwar im Standard definiert, aber von keinem Hersteller implementiert. In der Praxis auftretende Probleme, wie z.B. die Verwaltung der Empfänger, haben kaum Aussichten, behoben zu werden.
- Für die korrekte Benutzung der Programmierschnittstelle und die Behandlung von Fehler-situationen ist teilweise sehr viel Wissen über ATM und insbesondere die Signalisierung erforderlich.

Insgesamt gesehen, ist die Multicast-Verteilung mittels der XTI-Schnittstelle und ATM Punkt-zu-Mehrpunkt Verbindungen im Rahmen des 3V-Projektes eine interessante Option.

Sie wurde im Rahmen des 3V-Projektes sowohl für Test-/Lastmessprogramme (mgen) als auch für die Kommunikation von Simulation und Visualisierung realisiert.

#### 4.2.2 Gemeinsames IP-Netz / LANE

LAN Emulation ([LANE1], [LANE2], [LNNI2]) ist ein vom ATM Forum definierter Standard, der die Emulation eines Ethernet oder Token Ring auf ATM-Netzen erlaubt. Für die Verteilung von Multicast-Daten wird ein zentraler Server namens Broadcast and Unknown Server (BUS) benutzt. Der Multicast-Sender schickt die Daten zum BUS, welcher sie über eine Punkt-zu-Mehrpunkt Verbindung an alle Teilnehmer des emulierten LAN (ELAN) verschickt. Der Standard erlaubt die Nutzung von unterschiedlichen MTUs, sowohl bei Ethernet als auch bei Token Ring.

LAN Emulation wird von vielen Herstellern unterstützt und ist weitverbreitet. Auf einem ATM Adapter können in der Regel mehrere LANE-Klienten und/oder LANE-Server gleichzeitig laufen. Die meisten Hersteller (wie z.B. Sun) unterstützen nur die Ethernet Emulation mit 1500 Bytes MTU. Einige (wie Cisco) zusätzlich Token Ring Emulation mit 4490 Bytes MTU. FORE bietet ein breites Spektrum und unterstützt z.B. auf den Sun Workstations Ethernet Emulation mit 1500, 4528 und 9218 Bytes MTU und Token Ring Emulation mit 4490, 9180 und 18174 Bytes MTU. Ein gemeinsames IP-/LAN Emulation Netz mit allen bei den Projektpartnern vorhandenen Geräten ist aber nur mit Ethernet und 1500 Bytes MTU durchführbar.

Im 3V-Projekt wurden unterschiedliche emulierte LANs realisiert:

- 3v/192.168.2.x - Ethernet Emulation mit 1500 Bytes MTU. BUS ist ein Cisco Catalyst 5505 Ethernet Switch mit 622 Mbit/s ATM Interface. Dieser ist auch gleichzeitig LANE-Klient und realisiert den Übergang zu Gigabit Ethernet. Alle an 3V beteiligte Maschinen sind Mitglieder dieses Netzes.
- 3vv/192.168.4.x - Ethernet Emulation mit 9218 Bytes MTU. BUS ist eine Sun Ultra 60 mit FORE HE622 ATM Adapter. Maschinen mit FORE ATM Interface sind Mitglieder dieses Netzes.
- 3v-e1516, 3v-e4544, 3v-e9234, 3v-tr4544, 3v-tr9234, 3v-tr18190 – Testnetze, um alle Varianten zu überprüfen, Probleme festzustellen und gezielte Lastmessungen durchzuführen. Als BUS werden der Cisco C5505 und eine Sun mit FORE HE622 Adapter verwendet. Klienten sind nur die für die Tests/Messungen benötigten Maschinen.

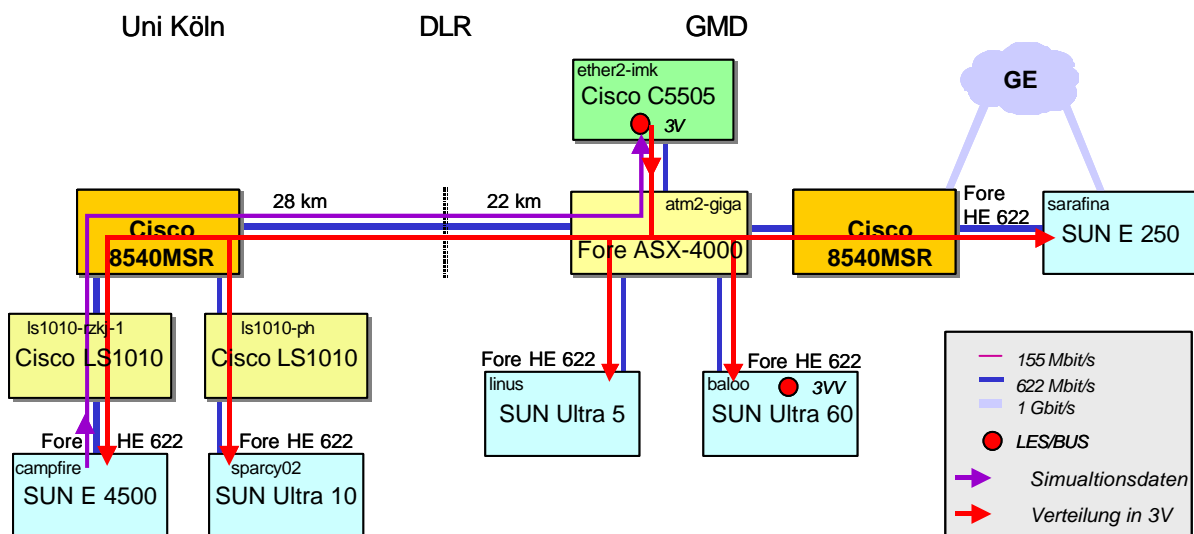


Abbildung 4-1: LANE-Netz im 3V-Projekt



Die wesentlichen Nachteile sind:

- Der Verteilbaum der Daten ist nicht optimal. Zum einen werden Strecken unnötig in zwei Richtungen belastet (vom BUS zum Sender). Zum anderen bekommen alle Teilnehmer des Netzes die Daten zugestellt, ob eine Anwendung läuft oder nicht.
- Bei ATM-basierten Endgeräten geschieht das Herausfiltern der nicht benötigten IP-Multicast-Pakete in Software und belastet damit die Maschine.
- Die Paketgröße ist in der Praxis meist auf 1500 Bytes beschränkt.
- Sender und alle Empfänger der Multicast-Daten sind Mitglieder eines gemeinsamen IP-Netzes.

Wegen der einfachen Einrichtung der Netze auf den Endgeräten wurden im Projekt zwei emulierte Netze (3v, 3vv mit Ethernet Emulation und MTUs von 1500 bzw. 9280 Bytes) dauerhaft realisiert und bilden die Grundlage für die Nutzung von IP Multicast.

### 4.2.3 IP Multicast-Routing

IP Multicast Router erlauben den Aufbau größerer IP Multicast Netze. Insbesondere im WAN beim Verbinden mehrerer Organisationen und unterschiedlicher Link-Layer-Technologien ist häufig der Einsatz von IP Routern erforderlich. Während früher das Forwarding von IP-Multicast-Paketen durch den Prozessor des Routers erfolgte, ist bei modernen Routern das Multicast Forwarding in Hardware/ASICs realisiert. Solche Router unterstützen Multicast Forwarding bis Leitungsbandbreiten. Im 3V-Projekt fanden Cisco C8540MSR Router Verwendung, da sie zum einen Multicast Forwarding in Hardware durchführen und zum anderen einen guten Übergang zwischen ATM und Gigabit-Ethernet-Netzen ermöglichen.

Die Verteilung der Multicast Pakete wird durch IP Multicast Routingprotokolle gesteuert. Auf Cisco Routern sind drei verschiedene Multicast Routing Protokolle implementiert ([Wil00]): Distance Vector Multicast Routing Protocol (DVMRP), Protocol Independent Multicast (PIM) dense-mode (PIM-DM) und PIM sparse-mode (PIM-SM).

DVMRP ist nur auf Tunnel-Schnittstellen implementiert. IP-Pakete über ein Tunnel Interface werden vom Hauptprozessor des Routers geforwarded und sind daher für hohe Bandbreiten nicht geeignet. DVMRP kann mit Cisco Routern nicht für Hochgeschwindigkeits-Multicast eingesetzt werden.<sup>3</sup>

PIM dense-mode (PIM-DM) gehört zur Klasse der flood&prune-Protokolle. In regelmäßigen Zeitabständen (3 Minuten) werden die Pakete aller Multicast-Gruppen im gesamten Netz geflutet. Router, die keine Empfänger für die Gruppen haben, schicken Prune-Nachrichten. Auf diese Weise wird ein optimaler, Senderbasierter Verteilbaum (Source Based Tree, SBT) aufgebaut. Nachteil ist das erhöhte Verkehrsaufkommen durch das regelmäßige Fluten der Pakete, das in größeren Multicast-Netzen zu erheblichen Problemen führt – insbesondere, wenn Backbone Leitungen mit geringerer Kapazität vorhanden sind. Generell ist von der Verwendung von flood&prune-Protokollen für Hochgeschwindigkeits-Multicast in größeren Netzen abzuraten. Sie sind allenfalls einsetzbar, wenn das Multicast Routing für die Hochgeschwindigkeits-Anwendung streng vom generellen Multicast-Routing getrennt wird.

PIM sparse-mode (PIM-SM) benutzt einen expliziten Anmelde Mechanismus für einzelne Multicast-Gruppen. Empfänger melden sich bei einem Rendezvous Point (RP) an. Multicast-Pakete werden nur an Empfänger verteilt, die sich vorher bei dem RP für die spezielle Multi-

---

<sup>3</sup> Dies gilt nicht für Router anderer Hersteller! DVMRP wird von den meisten Herstellern auf normalen Interfaces (Ethernet/ATM) unterstützt. DVMRP ist dann einsetzbar; allerdings mit den gleichen Einschränkungen und Problemen wie PIM dense-mode.

cast-Gruppe angemeldet haben. Die Multicast Verteilung erfolgt zunächst vom Sender zum RP und von dort zu den Empfängern (Core Based Tree mit RP als Core). Dies kann gegenüber Sender basierten Verteilbäumen zu einer erhöhten Netzbelastung führen. Aus diesem Grunde schaltet PIM-SM auf eine Sender basierte Verteilung um. Dies geschieht normalerweise bereits nach dem ersten gesendeten Multicast-Paket. Ist der RP zwischen Sender und allen Empfängern positioniert, so liegt von Anfang an ein Sender basierter Verteilbaum vor. Aufgrund dieses Verhaltens ist PIM-SM für Hochgeschwindigkeits-Multicast bevorzugt zu verwenden.

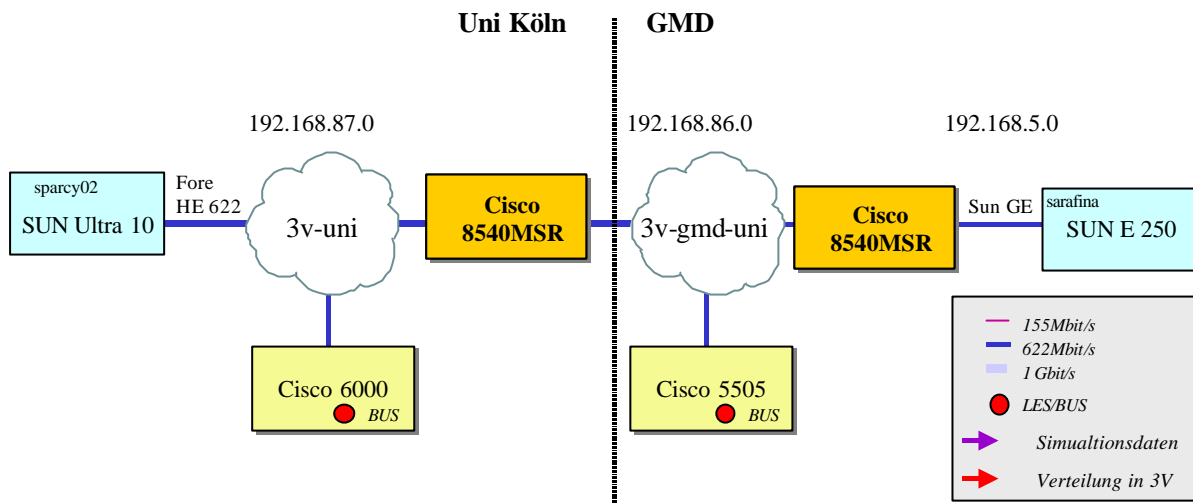


Abbildung 4-1: IP Multicast-Routing Aufbau

Im Rahmen von 3V wurde die Kopplung von lokalen Multicast-Netzen mittels eines LANE-Netzes realisiert (Abbildung 4-1). Die lokalen Multicast-Netze basierten dabei auf LAN Emulation und Gigabit Ethernet. Es wurden die Multicast-Routing-Protokolle PIM-DM und PIM-SM eingesetzt. Um Wechselwirkungen mit Multicast-Produktionsnetzen (GMD, Universität Köln, aber auch dem B-WiN) zu vermeiden, wurde das Routing von dem normalen Multicast Routing der Organisationen strikt getrennt. Untersucht wurden mit dieser Konfiguration sowohl PIM dense-mode als auch PIM sparse-mode mit einem Rendezvous Point.

Bei einem Netzaufbau wie im 3V-Projekt wäre es sinnvoll, die beiden Standorte mittels eines PVC zwischen den Routern zu koppeln, da die Kopplung von zwei Geräten mittels LAN Emulation einen großen Overhead darstellt. PVCs wurden von dem für den C8540MSR zur Verfügung stehenden IOS aber noch nicht unterstützt. Dies ist für das nächste Release angekündigt.

Vorteile der Router-Lösung sind:

- Wie bei der LANE kann IP Multicast genutzt werden und lokale Netze können auf beliebigen Technologien wie ATM oder Gigabit Ethernet beruhen.
- Ähnlich wie bei der native ATM-Lösung mit XTI kann ein optimaler Verteilbaum (Source Based Tree, SBT) aufgebaut werden. Multicast Pakete werden nur zugestellt, wenn auch Empfänger vorhanden sind.

Nachteile sind:

- Bei Fehlkonfigurationen und/oder Verwendung von flood&prune-Protokollen kommt es zu erheblichen Netzbelastungen. Die Gefahr, die Multicast-Produktionsnetze zu fluten, ist relativ hoch. Dies führt in der Regel zu erheblichen Netzstörungen.

- Hochleistungs-Router unterstützen Multicast Forwarding meist in Hardware Konfigurationen, die dazu führen, dass dieser schnelle Weg nicht genutzt wird, sind zu meiden. Forwarded die Haupt-CPU des Routers, so sinkt die Leistung beträchtlich und es kommt zu Störungen der anderen Router-Prozesse.

Insgesamt gesehen ist die Verwendung von IP Multicast Routing für Hochgeschwindigkeits-Multicast-Verteilung heute mit der geeigneten Hardware wie z.B. Cisco C8540MSR oder 12000 möglich. Gerade in größeren Netzen ist aber dringend die Verwendung von PIM sparse-mode mit einem Rendezvous Point (RP) anzuraten und es ist darauf zu achten, dass das Forwarding auch tatsächlich Hardware basiert erfolgt.

#### 4.2.4 Sonstige Verfahren

Für die Verteilung von IP Multicast über ATM existieren eine Reihe weiterer Verfahren. Zum einen hat die IETF einen Standard für IP Multicast über ATM entwickelt, zum anderen gibt es Hersteller proprietäre Verfahren von Cisco und FORE.

Die IETF hat schon 1996 eine Methode, IP Multicast über ATM Netze zu nutzen, standardisiert (Support for Multicast over UNI 3.0/3.1 based ATM Networks, [RFC2022]). Der ursprüngliche Ansatz wurde in einer Reihe von weiteren RFCs verfeinert ([RFC2121], [RFC2149], [RFC2443]). Der IETF-Ansatz sieht die Einführung eines Multicast Address Resolution Server (MARS) vor, welcher analog dem Classical IP ATMARP Server Multicast-Adressen nach ATM-Adressen auflöst. Für die Verteilung der Multicast-Pakete können sowohl Multicast Server (bei vielen Sendern mit niedrigen Bandbreiten; Minimierung der Anzahl der Verbindungen), als auch direkte Punkt-zu-Mehrpunkt Verbindungen (bei wenigen Sendern und hohen Bandbreiten; keine zentralen Engpässe) zwischen Sender und Empfängern benutzt werden. Kritik hat sich insbesondere gegen die Komplexität dieses Ansatzes erhoben. Keiner der großen Hersteller von ATM-Geräten (z.B. FORE, Cisco) unterstützt zur Zeit diesen Ansatz und eine Realisierung ist auch in Zukunft nicht zu erwarten. Ein Test im Rahmen des Projektes war daher nicht möglich.

FORE hat schon Mitte der 90iger Jahre im Rahmen des proprietären FORE-IP IP-Multicast über ATM unterstützt. FORE-IP benutzt nicht die (damals noch nicht definierte bzw. implementierte) ATM Forum Signalisierung, sondern basiert auf einer FORE proprietären Signalisierung. Der Austausch der IP-Multicast-Pakete erfolgt über einzelne Punkt-zu-Mehrpunkt Verbindungen für jede Multicast-Gruppe. Da FORE die proprietäre Lösung für FORE-IP und Signalisierung nicht mehr weiterentwickelt und die proprietäre Signalisierung aufgrund der heterogenen Infrastruktur (FORE, Cisco) nicht zwischen den Projektpartnern realisierbar war, machte der Einsatz und die Untersuchung dieses Ansatzes keinen Sinn.

Cisco hat eine proprietäre Lösung für das Versenden von IP-Multicast-Paketen über ATM für seine Router entwickelt. Die Router verwenden Punkt-zu-Mehrpunkt Verbindungen zum Austausch einzelner Multicast Gruppen. Das Verfahren beruht auf der ATM Form 3.0/3.1 Signalisierung. Auf dem zur Zeit neuesten IOS für den C8540MSR Router (IOS 12.0(7)) ist der Ansatz allerdings nicht implementiert und konnte daher nicht getestet werden.

#### 4.2.5 Fazit

Im 3V-Projekt wurden drei Ansätze für die Verteilung von Multicast mit hohen Datenraten realisiert: die direkte Nutzung von ATM Punkt-zu-Mehrpunkt Verbindungen über die XTI Programmierschnittstelle, die Verteilung von IP-Multicast-Paketen mittels LAN Emulation und die Verteilung mittels IP Multicast Routing. Der Vorteil der direkten ATM-Nutzung liegt darin, dass die Daten optimal im Netz verteilt werden; sowohl die Netzbelastung als auch die Belastung von Endgeräten ist minimal. Bei LANE kommt es zu einer wesentlich höheren

Belastung. Diese wirkt sich insbesondere beim Sender aus und im Fall, dass BUS und Klient auf einem Gerät laufen. Großer Vorteil von LANE ist, dass IP Multicast genutzt werden kann und Maschinen auch über Gigabit Ethernet angeschlossen werden können. Die Verwendung von IP Multicast Routing erlaubt eine bessere Aufteilung des IP-Multicast-Verkehrs als bei LANE. Die Konfiguration des Multicast Routings hat allerdings mit großer Sorgfalt zu erfolgen, um Probleme in Produktionsnetzen zu vermeiden.

Bei den übrigen Ansätzen macht eine nähere Untersuchung wenig Sinn: MARS ist zur Zeit von den großen Herstellern nicht realisiert und FORE IP Multicast beruht auf einem FORE spezifischem Signalisierungsprotokoll zwischen ATM Switchen, welches im 3V-Umfeld nicht genutzt werden kann.

## **4.3 Messungen**

Im Rahmen des Projektes wurden zahlreiche Messungen durchgeführt, um die Leistungsfähigkeit der Übertragungsmethoden zu untersuchen und einen Einblick in die wichtigen Parameter zu gewinnen.

Zunächst wird mgen vorgestellt - die Software, die wir für die Generierung der Multicast Daten benutzen - und es erfolgt eine Aufstellung der theoretischen Übertragungsraten der einzelnen Technologien. Anschließend werden die einzelnen Messungen vorgestellt.

### **4.3.1 Testwerkzeug**

Als Tool zur Messung der Performance von Multicast über UDP wurde das Programm MGEN verwendet. Dieses steht im Sourcecode zur Verfügung ([MGEN]), so dass die nötigen Modifikationen selbst durchgeführt werden konnten. MGEN besteht aus zwei Einzelprogrammen. Ein Sender (mgen) schickt über ein bestimmtes Interface UDP-Multicast-Pakete mit definierter Paketgröße und Senderate. Ein Empfänger (drec) empfängt Multicast Pakete einer IP Multicast-Gruppe über ein vorgegebenes Interface.

Für die Messungen in 3V wurde MGEN modifiziert, so dass auch größere Pakete, z.B. 9190 Bytes wie sie bei LAN Emulation möglich sind, verschickt werden können.

Zusätzlich wurde auf der Grundlage des Source-Codes von MGEN ein Tool entwickelt, das die Versendung von Multicast über native ATM mittels der FORE-XTI-Schnittstelle ermöglicht. Dazu wird vom Sender aus eine ATM Punkt-zu-Mehrpunkt Verbindung zu den Empfängern aufgebaut. Dem Tool werden in einem Skript die gewünschten Empfängersysteme, die ATM-Dienstkategorie (UBR, CBR, VBR) zusammen mit den jeweiligen Verkehrsparametern, die Senderate und die Paketgröße vorgegeben. Auf den Empfängern läuft ein Programm, das den Verbindungsaufbau entgegennimmt und als Datensenke fungiert.

Bei beiden Varianten von MGEN (IP Multicast/UPD und XTI) wird beim Versenden bzw. Empfangen eine Statistik über die Pakete geführt, die als Grundlage für die Leistungsmessungen dient.

### **4.3.2 Theoretische Übertragungsraten**

Die in 3V verwendeten Netztechnologien ATM und Gigabit Ethernet weisen Übertragungsraten von 622 Mbit/s bzw. 1 Gbit/s auf. Aufgrund des Overhead der einzelnen Protokolle steht einer Anwendung jedoch weniger Bandbreite zur Verfügung. Tabelle 4-1 gibt einen Überblick über die maximal erreichbaren Übertragungsraten bei Verwendung von UDP/IP Multicast über LANE/ATM und GE sowie von XTI über ATM.

Protokoll	Paketgröße/Byte	Übertragungsrate/(Mb/s)
XTI	1472	536,7
	8192	541,5
XTI ohne Padding	1432	539,51268
	8152	541,99483
UDP/LANE	1472	519,9
	9190	538,2
UDP/LANE ohne Padding	1436	523,56745
	9212	539,48145
UDP/GE	1472	964,61337

Tabelle 4-1: Übertragungsraten unter Berücksichtigung des Protokoll Overhead

Bei Verwendung von 622.08 Mbit/s ATM beträgt der Sonet/SDH Overhead 23,04 MBit/s. Der ATM Schicht stehen 599.04 Mbit/s zur Verfügung. Nach Abzug des Overhead für den ATM-Zell-Header stehen für AAL 5 maximal 542,53 MBit/s bzw. 1.412.830 Zellen bereit.

Bei AAL5 erfolgt ein Padding auf ein Vielfaches des ATM-Zell-Nutzinhaltes von 48 Bytes. Der AAL5 Overhead ist daher für bestimmte Paketgrößen (Vielfaches von 48 Bytes) am geringsten. Maximale Übertragungsraten bei LANE und XTI werden bei Werten erreicht, die ein Vielfaches von 48 Bytes sind.

Durch die jeweiligen Header ergeben sich die folgenden Overheads:

UDP/LANE: 52 Bytes: 8 Bytes UDP + 20 Bytes IP + 16 Bytes LANE + (8 + Padding) Bytes AAL5

XTI: (8 + Padding) Bytes AAL5

UDP/GE: 54 Bytes: 8 Bytes UDP + 20 Bytes IP + 26 Bytes GE

Für kleinere Pakete ergibt sich eine weitere Reduzierung der maximal möglichen Bandbreite.

### 4.3.3 Einfluss der Rechnerarchitektur

Die maximal mögliche Sendeleistung muß in realen Systemen nicht der theoretisch möglichen aus Abschnitt 4.3.2 entsprechen. Unter anderem kann sie begrenzt sein durch die Hardwareausstattung des Rechners. Die wichtigsten Komponenten sind dabei die Geschwindigkeiten von Prozessor und Cachegröße, die Anzahl der CPU's, sowie die Bus-Frequenz und Bus-Breite.

Uns standen für die Messungen die folgenden Rechner zur Verfügung:

- Sun Ultra 5 ("ULTRA-5"), Solaris 7, 1 CPU mit 330 MHz,  
FORE HE622 ATM Adapter im 32Bit/33MHz PCI-BUS
- Sun Ultra 60 ("ULTRA-60"), Solaris 7, 2x 300 MHz CPU  
FORE HE622 ATM Adapter im 64Bit/66MHz PCI-BUS
- Sun E250 ("E250"), Solaris 2.6, 2x400 MHz CPU  
FORE HE622 ATM Adapter im 64Bit/66MHz PCI-BUS  
Sun GE Adapter in 64Bit/33MHz PCI-BUS
- Sun E250 ("E250-33"), Solaris 2.6/7, 2x400 MHz CPU  
FORE HE622 ATM Adapter im 64Bit/33MHz PCI-BUS  
SysKonnnect GE Adapter in 64Bit/33MHz PCI-BUS

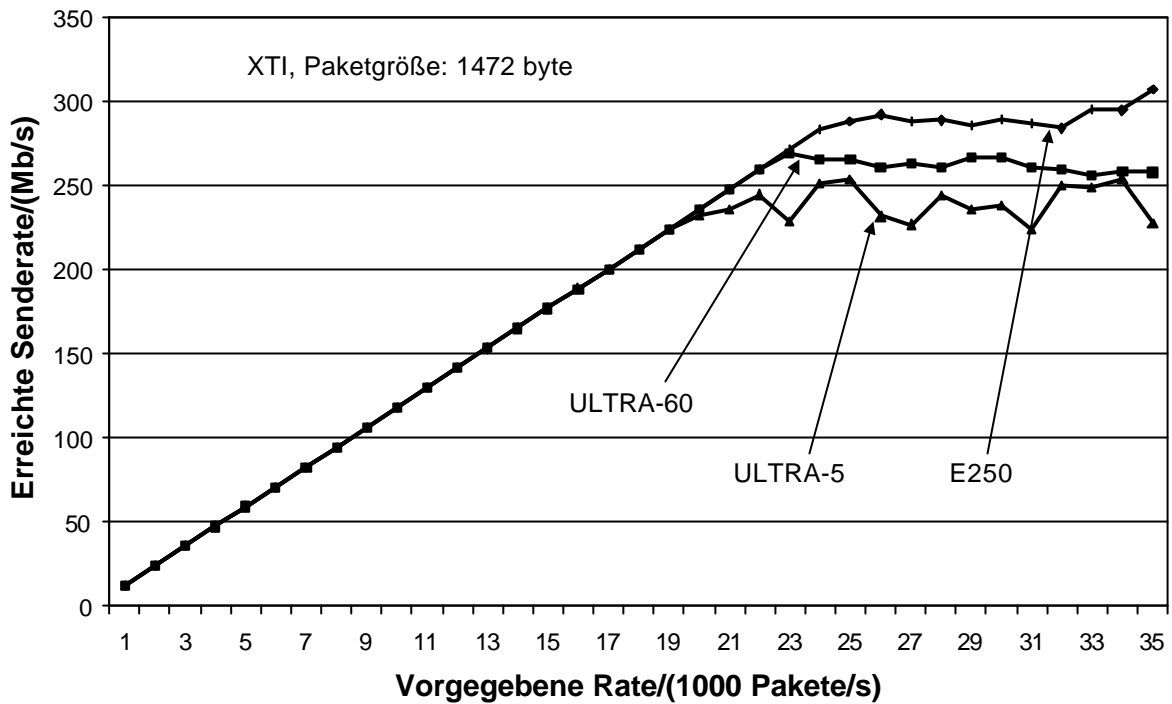


Abbildung 4-1: Multicast-Sendeleistung verschiedener Systeme über XTI

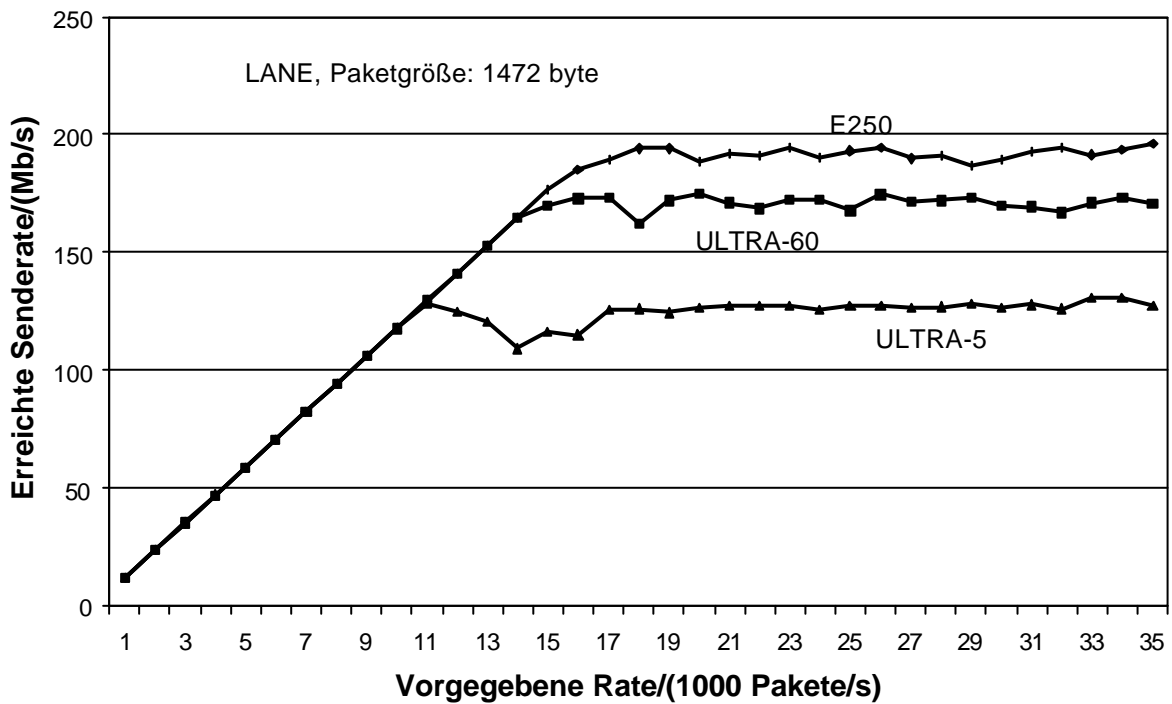


Abbildung 4-2: Multicast-Sendeleistung verschiedener Systeme über LANE

Für XTI und LANE haben wir auf den drei Systemen ULTRA-5, ULTRA-60 und E250 im Tool mgen die vorgegebene Senderate schrittweise erhöht und parallel dazu die tatsächlich erreichte Sendeleistung gemessen. Die Paketgröße betrug 1472 Bytes. Man erkennt in Abbildung 4-1 und Abbildung 4-2 Bereiche linearen Wachstums, in denen das System die vorgegebene Senderate ohne Probleme erreicht. Danach gehen die Kurven über in eine hori-

zontale Linie, was anzeigt, dass die unter den gegebenen Randbedingungen mögliche maximale Senderate auf dem Adapter erreicht wurde. Man erkennt, dass ein leistungsfähigeres System, z.B. mit einer schnelleren CPU, auch höherer Sendeleistungen erreicht. Allerdings liegen auch die Maxima der E250, ca. 290 MBit/s bei XTI und 190 MBit/s bei LANE, nicht annähernd bei dem theoretisch möglichen Wert von 519,9 MBit/s (LANE) bzw. 536,7 MBit/s (XTI).

Ein Grund dafür kann bei der hier verwendeten Paketgröße von 1472 Bytes liegen, für die das Sendeverhalten der vorhandenen Hardware möglicherweise nicht optimal ist. Dies wird im folgenden Abschnitt genauer untersucht.

#### 4.3.4 Einfluss der Paketgröße

Die folgenden Messungen wurden auf der E250 durchgeführt. Wir haben nach dem gleichen Verfahren wie im vorhergehenden Abschnitt schrittweise die Senderate in mgen erhöht und diesmal zusätzlich noch die Größe der Pakete im Bereich von ca. 100 Bytes – 9000 Bytes variiert.

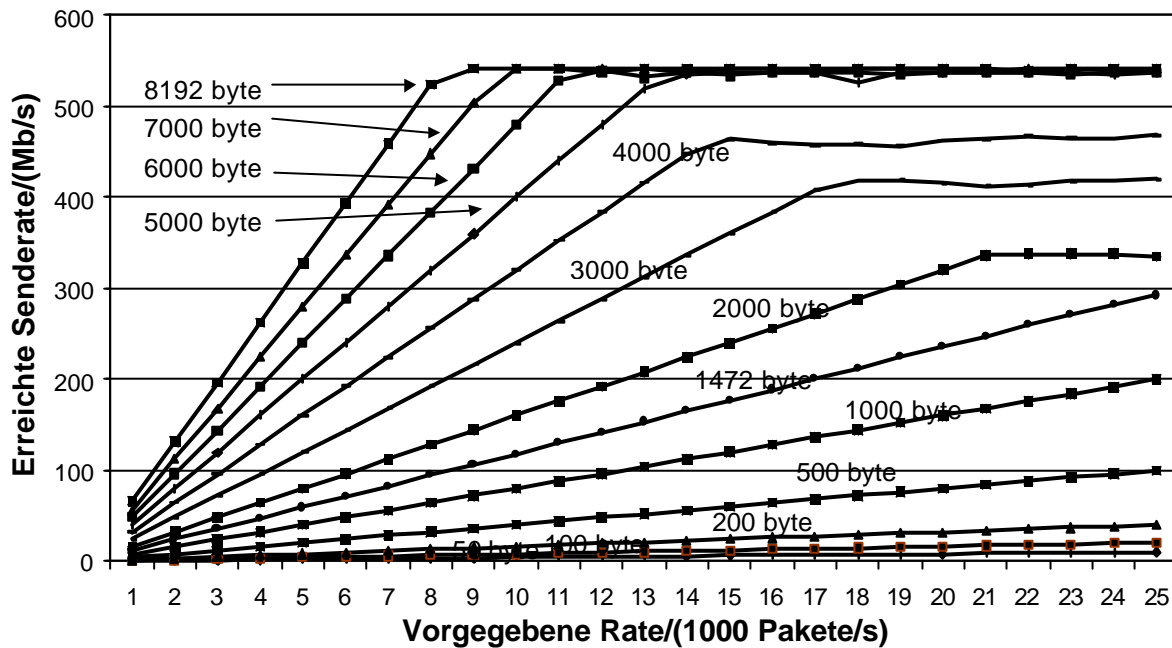


Abbildung 4-1: Multicast-Sendeleistung einer E.250 über XTI

Für XTI liegt das Maximum in etwa bei dem theoretischen Wert von 540 MBit/s (Abbildung 4-1). Dieser Wert kann mit Paketgrößen oberhalb 5000 Bytes erreicht werden. Bei kleineren Paketgrößen wird das theoretische Maximum offenbar nicht mehr erreicht. Als Grund vermuten wir die Begrenzung der Anzahl möglicher Interrupts auf der Hardware.

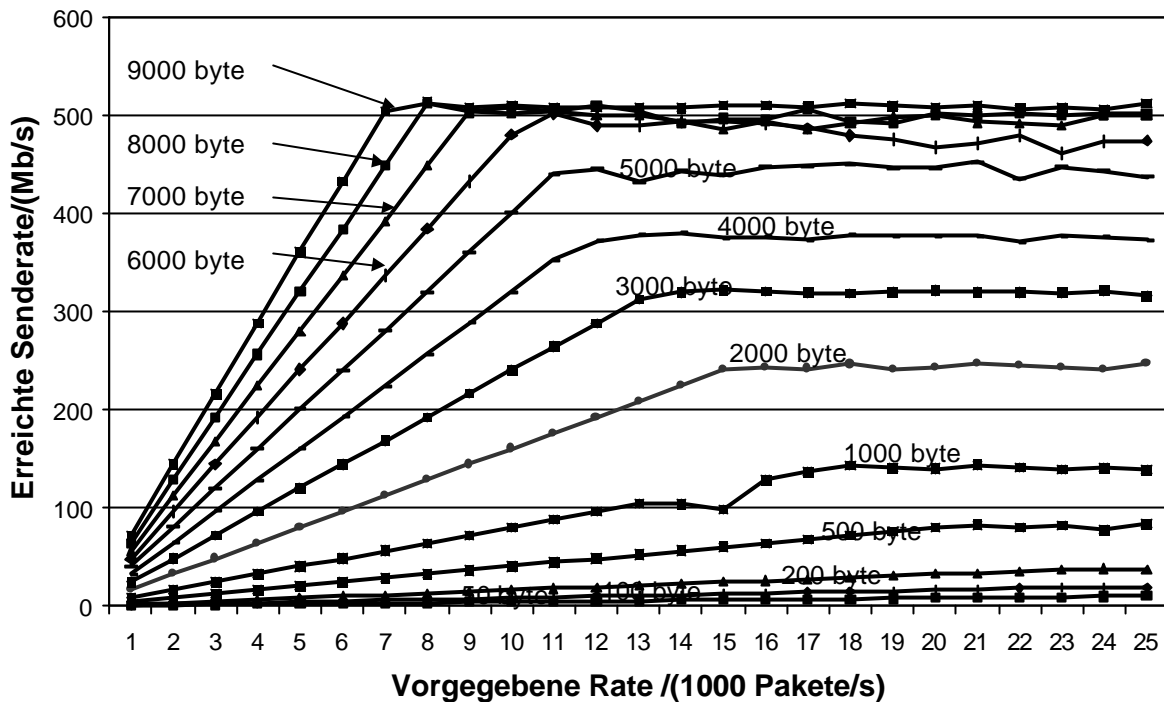


Abbildung 4-2: Multicast-Sendeleistung einer E.250 über LANE

Für UDP/LANE ist das Verhalten ähnlich (Abbildung 4-2). Das absolute Maximum liegt etwas niedriger. Dies kann aber durch den zusätzlichen IP-Overhead bedingt sein. Daneben hat Multicast über LANE den Nebeneffekt, dass der Sender die Pakete die raus gehenden Pakete über den BUS wieder zugestellt bekommt und sie verwerfen muss. Hierdurch entsteht eine zusätzliche Belastung des Prozessors. Dieser Effekt wird in Abschnitt 4.3.6 genauer untersucht werden.

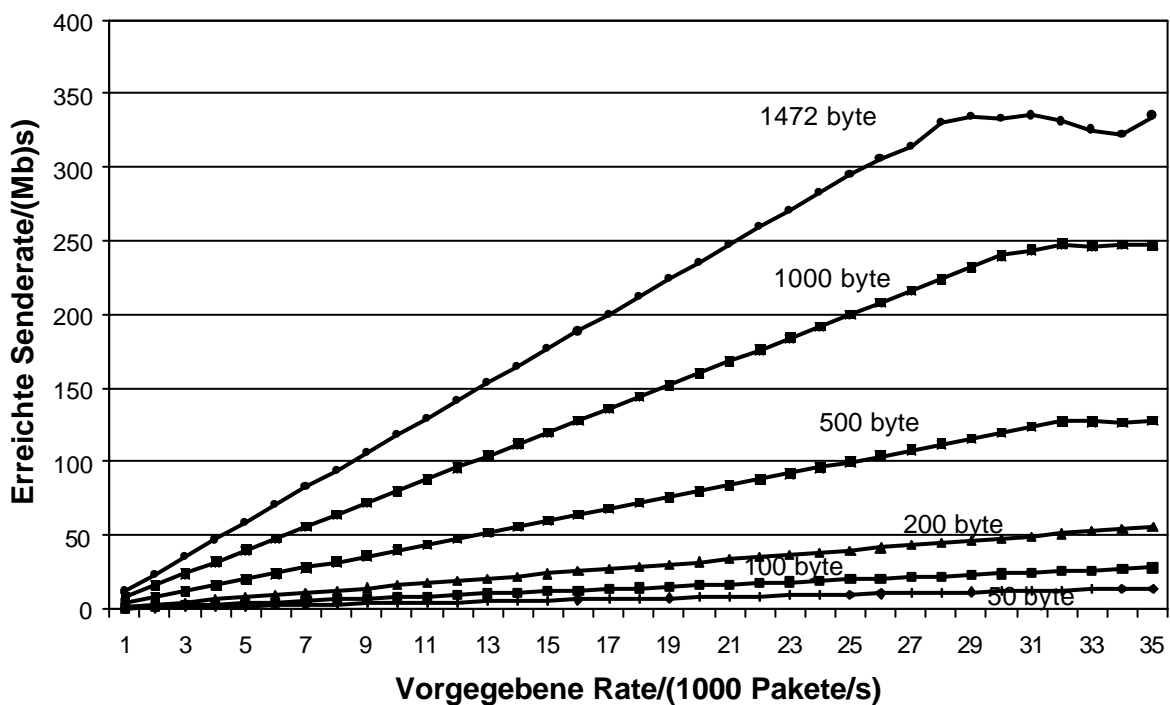


Abbildung 4-3: Multicast-Sendeleistung einer E.250 über GE

Abbildung 4-3 zeigt das Verhalten bei Gigabit Ethernet. Die Sendeleistung verhält sich bis zu der maximalen Paketgröße von 1472 Bytes analog zu XTI und LANE. Allerdings wird bei GE die theoretisch mögliche Senderate nicht annähernd erreicht.

Insgesamt gesehen, hat die Paketgröße einen wesentlichen Einfluss auf die maximale Übertragungsleistung. Die Anwendung sollte daher versuchen, die Simulationsdaten in möglichst großen Paketen zu schicken.

#### 4.3.5 Vergleich der Technologien

Zielsetzung dieser Messungen war es, festzustellen, welche Unterschiede in der Leistungsfähigkeit zwischen den im Projekt untersuchten Multicast-Technologien besteht. Auf einer Sun E250 mit FORE ATM Adapter und Sun GE Adapter wurde die Sendeleistung für XTI, LANE und GE ermittelt. Hierbei wurde eine einheitliche Daten-Größe von 1472 Bytes – die jede Technologie ohne Fragmentierung unterstützt - und die jeweils maximale Daten-Größe der einzelnen Technologien untersucht.

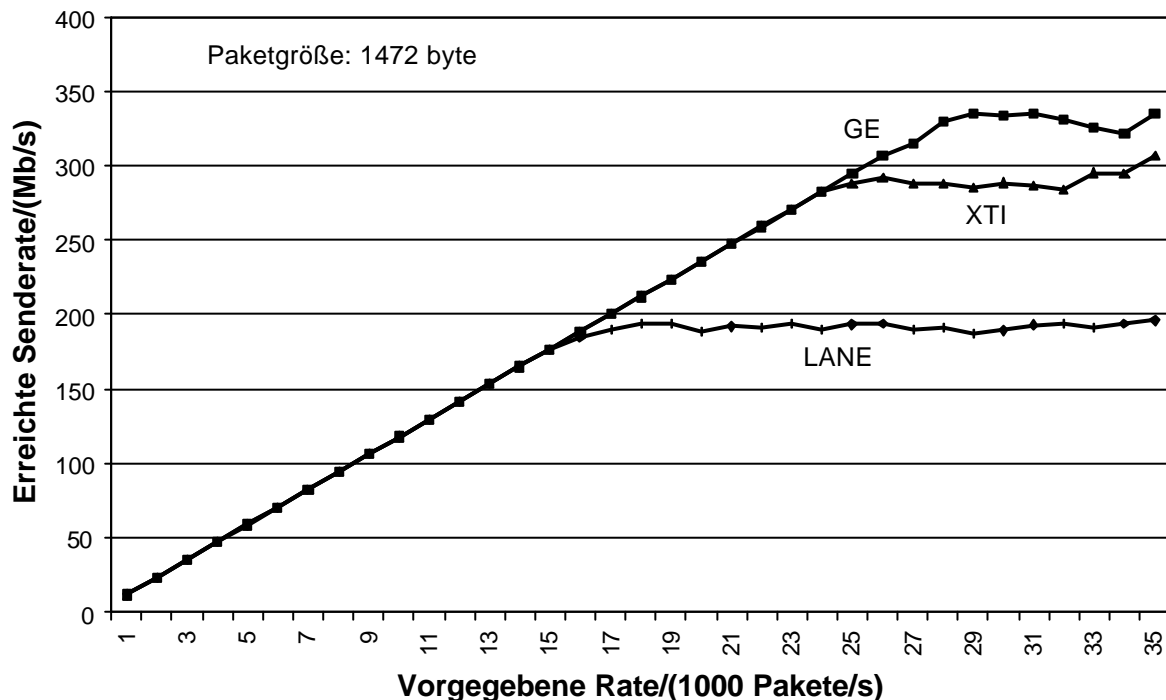


Abbildung 4-1: Multicast-Sendeleistung über verschiedene Technologien, 1472 Bytes

Abbildung 4-1 zeigt den direkten Vergleich zwischen XTI, LANE und GE bei 1472 Bytes großen Daten-Paketen. Bei dieser Größe ist GE mit maximalen Übertragungsraten von ca. 340 MBit/s die leistungsfähigste Netztechnologie, dicht gefolgt von XTI mit ca. 300 Mbit/s. LANE dagegen erreicht keine 200 Mbit/s und ist damit deutlich langsamer. Ursache für das schlechte Abschneiden von LANE ist wohl das bereits erwähnte Versenden der Multicast-Pakete vom BUS zum Sender.

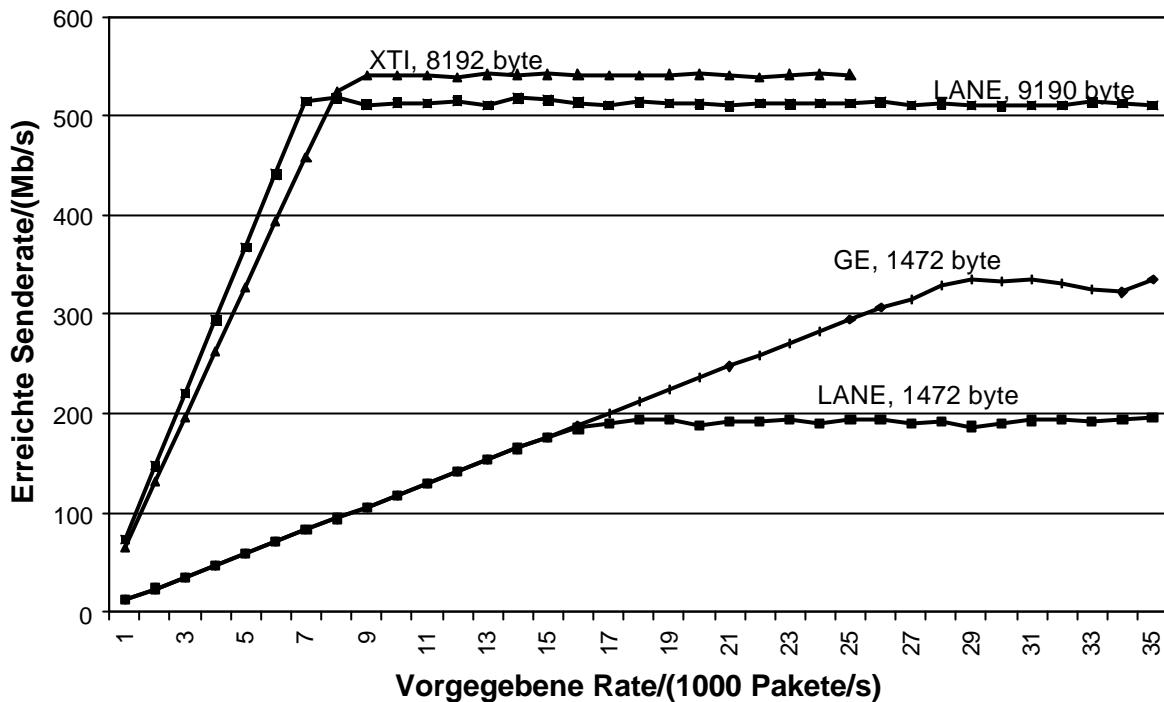


Abbildung 4-2: Multicast-Sendeleistung mit jeweils maximaler Paketgröße

Untersucht man das Sendeverhalten der Technologien bei der jeweils maximalen Datengröße, die ohne Fragmentierung übertragen werden kann, so ändern sich die Verhältnisse grundlegend (Abbildung 4-2). XTI und LANE können große Pakete<sup>4</sup> übertragen werden und beide Technologien erreichen dann in etwa ihr theoretisches Maximum von über 500 Mbit/s.

mit werden bei großen Paketen die höchsten Senderaten erzeugt und die theoretischen Grenzen der

Eine Entscheidung zwischen GE, XTI und LANE muß also immer unter Einbeziehung der Anwendung und speziell der von ihr verwendeten Paketgröße erfolgen.

#### 4.3.6 Vergleich zwischen Multicast und Unicast

Wie bereits erwähnt besteht ein Nachteil von LANE gegenüber XTI darin, dass der Sender die Multicast-Pakete zweimal verarbeiten muss, als raus gehende wie auch als herein kommende Last. Die rein kommenden Pakete müssen vom System extra verworfen werden. Dies stellt eine zusätzliche Belastung dar. Wie diese sich auswirkt, erkennt man in Abbildung Abbildung 4-1, wo wir für LANE und GE die jeweiligen Sendeleistungen vergleichen haben. Der Unterschied bei GE ist überraschend. Vermutlich liegt der Grund in einer unterschiedlich effektiven Implementierung von Multicast und Unicast. Die Differenz zwischen Unicast und Multicast bei LANE ist jedoch deutlich größer als der bei GE.

<sup>4</sup> XTI unterstützt auch noch größere Daten-Pakete bis 64 Kbyte. Eine Vergrößerung über 8 Kbyte hinaus bringt bei 622 Mbit/s keine wesentliche Leistungssteigerung. Ähnliches gilt für LANE mit Token Ring Emulation und 18174 Byte MTU.

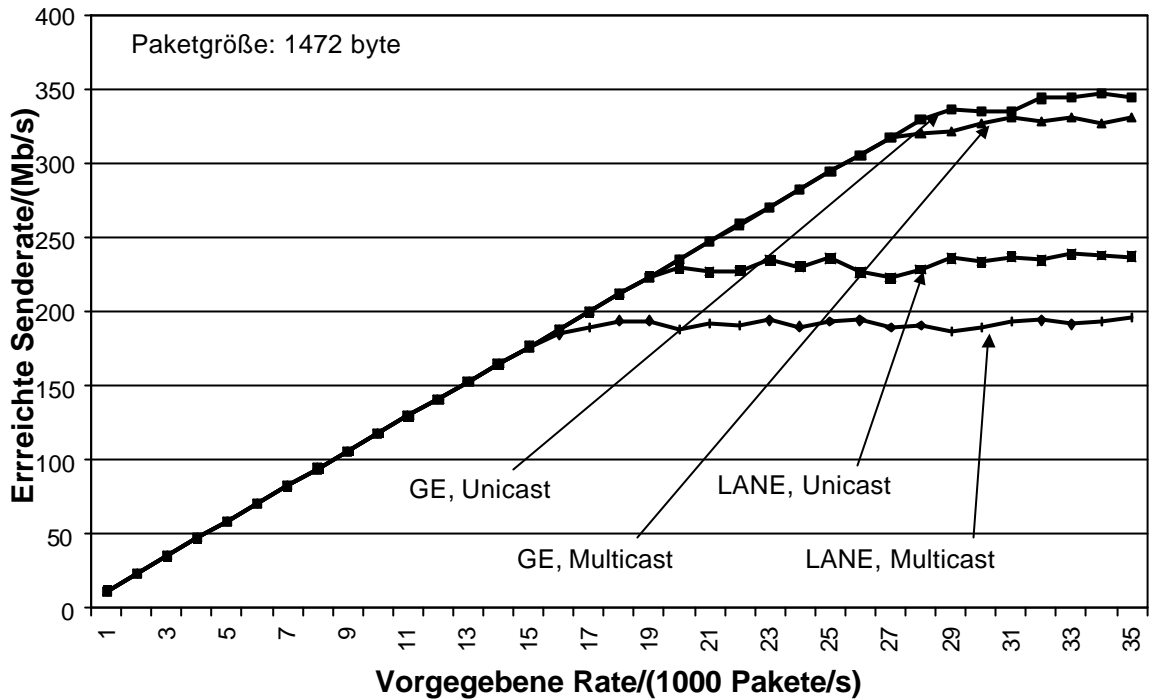


Abbildung 4-1: Multicast versus Unicast

### 4.3.7 Paketverluste

Bei den Messungen fiel uns teilweise eine sehr hohe Verlustrate bei den Empfängern auf. Dies trat besonders in Fällen auf, in denen der Sender leistungsfähiger war als der Empfänger - eine für das Multicast-Szenario von 3V nicht untypische Situation.

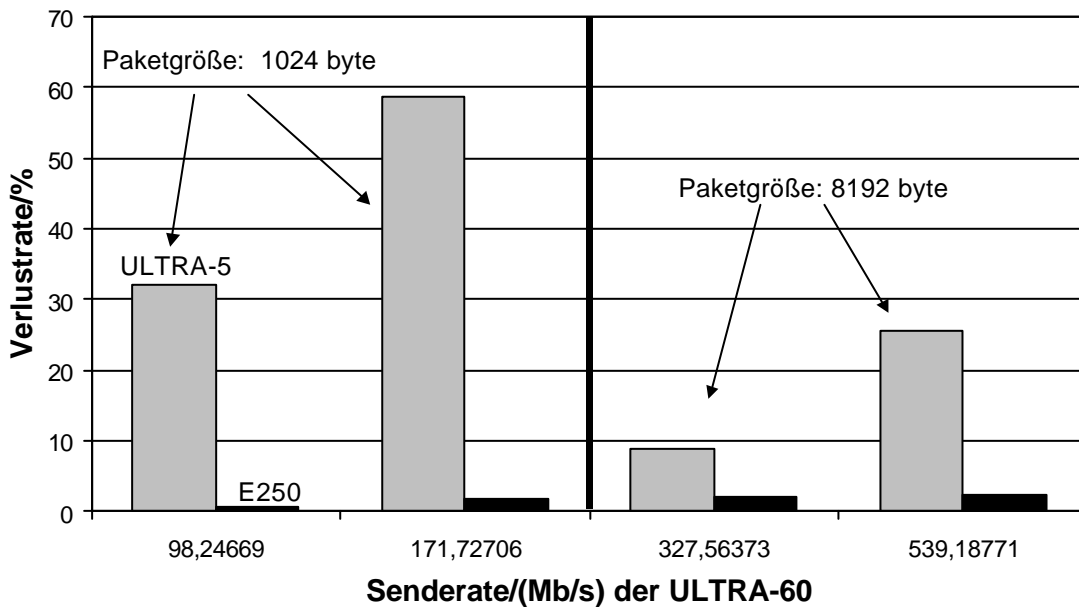


Abbildung 4-1: Verluste bei XTI, Interrupt Coalescing OFF

In der Abbildung 4-1 sieht man die Verlustraten, die sich beim Versenden über die XTI-Schnittstelle ergeben haben. Bei diesen Messungen schickt die ULTRA-60 ATM Multicast an die ULTRA-5 und an die E250. Wir haben diese Messungen für zwei Paketgrößen (1024 Bytes und 8192 Bytes) und für zwei Senderaten (maximal mögliche Rate auf der ULTRA-60 und die Hälfte davon) durchgeführt.

Offenbar hat die E250 keine Probleme bei der Verarbeitung der Pakete. Entsprechend niedrig sind die Verlustraten bei allen vier Messungen. Die ULTRA-5 dagegen verliert bedingt durch CPU-Überlastung unter Umständen bis zu 60 % der eintreffenden Daten. Dieser Effekt ist besonders ausgeprägt bei kleinen Paketen, da hier mehr Interrupts ausgelöst werden.

Bei Solaris gibt es einen Kernel-Parameter, der für die ATM-Karte festlegt, ob beim Empfang jedes Paketes direkt ein Interrupt ausgelöst wird. "Interrupt Coalescing" fasst mehrere Pakete zusammen und löst dann einen Interrupt aus, so dass mehrere Pakete innerhalb eines Interrupts bearbeitet werden. Dies senkt die Belastung der CPU. Wir haben die gleiche Messung wie oben noch einmal durchgeführt und dabei Interrupt Coalescing auf den Empfängersystemen aktiviert. Offenbar kann die Verlustrate insbesondere des schwächeren Systems dadurch erheblich gesenkt werden (Abbildung 4-2).

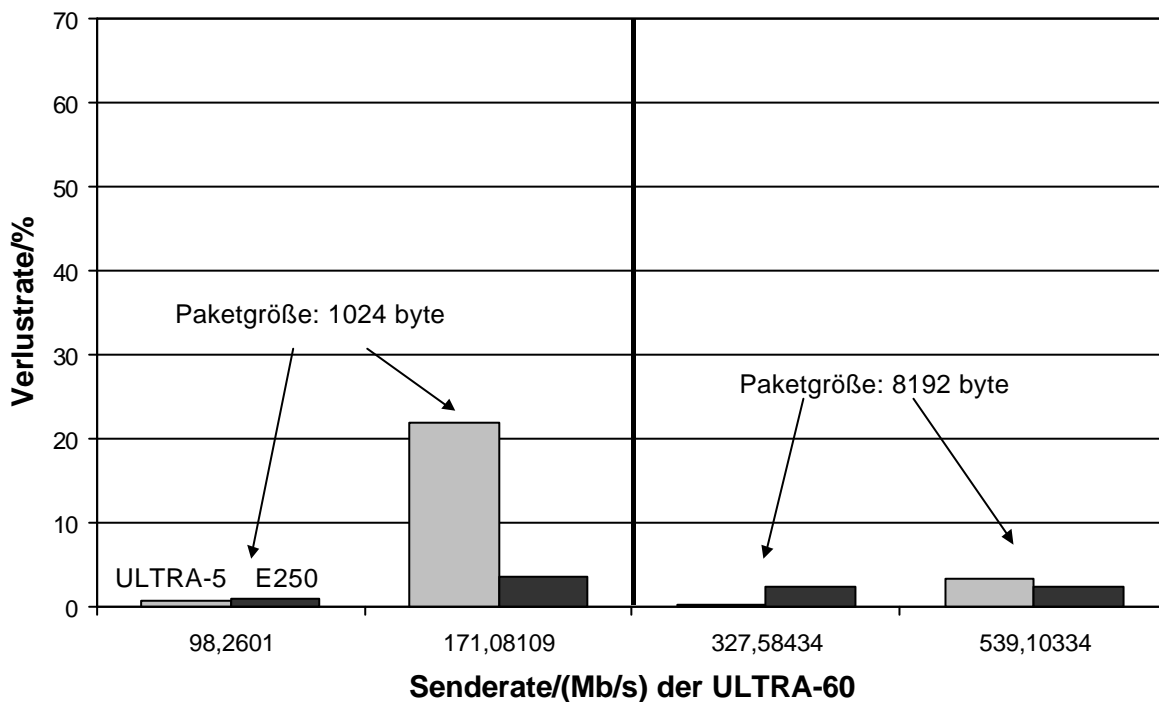


Abbildung 4-2: Verluste bei XTI, Interrupt Coalescing ON

Auch bei Multicast über LANE/UDP kommt es in der empfangenden Applikation unter Umständen zu erheblichen Verlusten. Bei den Messungen in Abbildung 4-3 schickte dieses Mal die E250-33 UDP Multicast (Paketgrößen: 1024 Bytes, 9190 Bytes, jeweils zwei verschiedene Senderaten) an die E250 und die ULTRA-5. Wieder sind die Verluste bei der ULTRA-5 besonders ausgeprägt, jedoch verliert jetzt auch die E250 einen großen Teil der Daten (etwa 30-50%), vermutlich bedingt durch den gegenüber der XTI Messung leistungsfähigeren Sender.

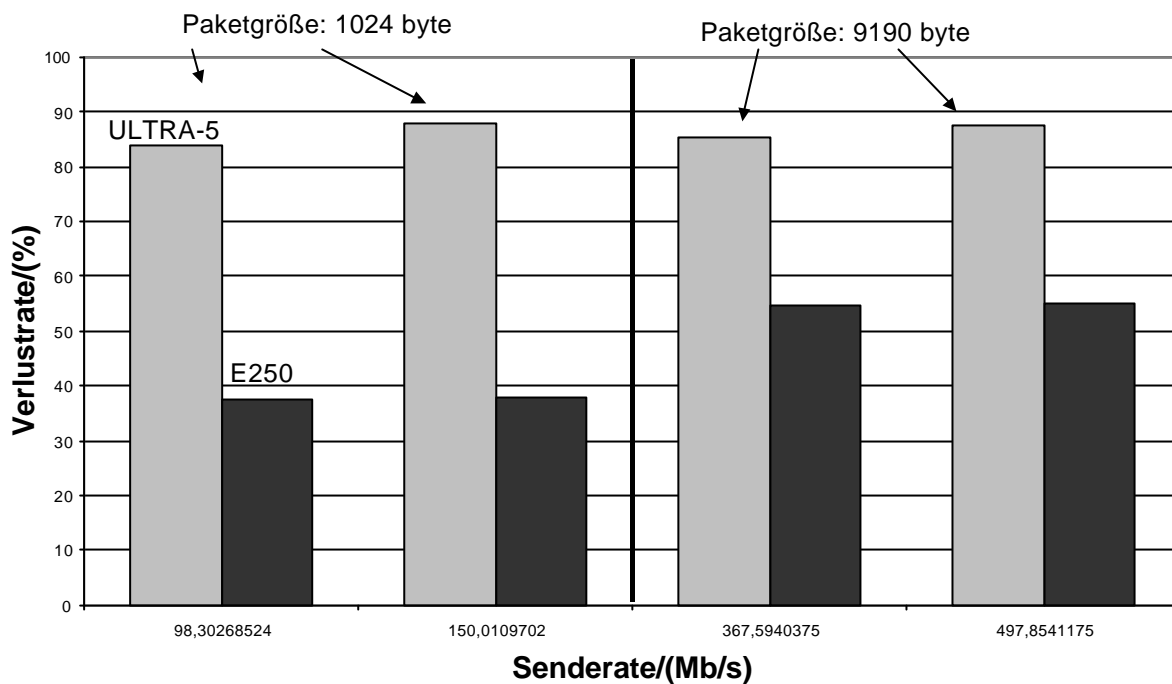


Abbildung 4-3: Verluste bei UDP/LANE, udp\_recv\_hiwat = 8192

Um diese hohen Verluste zu senken, haben wir über den Kernel Parameter `udp_recv_hiwat` auf den Empfängern vom Default-Wert 8192 Bytes auf 65535 Bytes erhöht (Abbildung 4-4). Bei der E250 ist der Effekt gut zu sehen: ein größerer Empfänger-Puffer senkt die UDP Verluste. Die ULTRA-5 jedoch ist insgesamt so überlastet, dass sich auch ein größerer Buffer kaum auswirkt. Es gehen der Applikation auch hier bis zu 80% der Daten verloren.

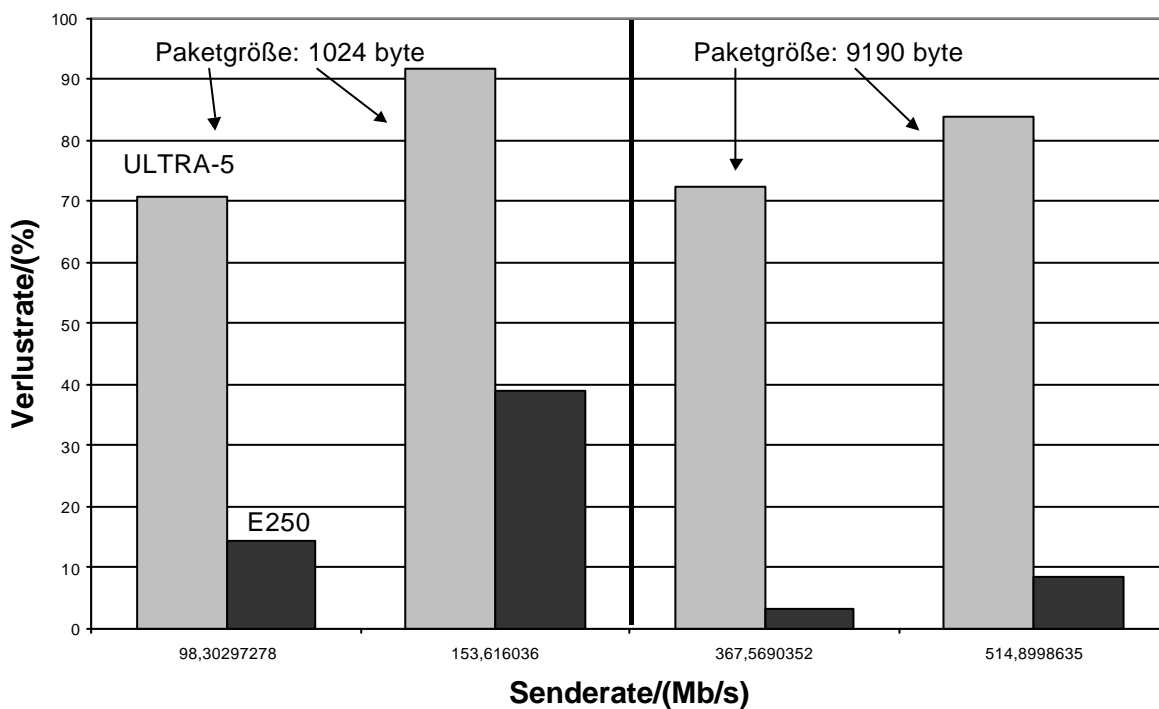


Abbildung 4-4: Verluste bei UDP/LANE, udp\_recv\_hiwat = 65535

Die Untersuchung, wie sich Interrupt Coalescing auf das UDP/LANE Verlustverhalten auswirkt, ließ sich nicht mehr durchführen, da in der Zwischenzeit eine neue ATM-Treiberversion (FT 5.0.2.5) auf den Systemen installiert worden ist, bei der das Setzen des Kernel-Parameters „he\_intr\_coalesce“ offenbar keine Wirkung mehr hat.

Die Untersuchungen haben gezeigt, dass Interrupt Coalescing eine sinnvolle Technik ist, um bei kleinen Paketen die Last auf dem Empfänger zu reduzieren. UDP-Puffer sollten bei den hohen Geschwindigkeiten unbedingt erhöht werden – sie gleichen letztendlich aber nur kurze Burst im Netz und Lastspitzen im Rechner aus.

### 4.3.8 IP Multicast Routing

Bei den Untersuchungen zum IP Multicast Routing wurden zwei Cisco C8540MSRs als Multicast Router benutzt, die LANE und Gigabit Ethernet basierte Netze verbanden (Abbildung 4-1). Die Sun Enterprise 250 mit Sun GE Interface bei der GMD diente als Sender; teilweise wurde die E250-33 mit Syskonnect GE Interface als zusätzlicher Sender benutzt. Die E250 erzeugte mit mgen eine Datenlast von ca. 300 Mbit/s. Dies entspricht über ATM ca. 800000 Zellen. Über die Cisco C8540MSR Router wurde der Multicast-Verkehr zur Uni Köln auf eine Sun Ultra 10 geleitet. Mit einem GNN ATM Analyser wurden jeweils die Zellen gemessen, die über den BUS bei der GMD gingen.

Der C8540MSR forwarded IP-Multicast-Pakete normalerweise Hardware/ASICs basiert und ist in der Lage, die Last von 300 Mbit/s ohne Belastung des Hauptprozessors zu verarbeiten. Beim Auftreten sehr vieler Multicast-Gruppen ist eventuell mit Problemen zu rechnen, da die ASICs nur eine beschränkte Anzahl von gleichzeitig aktiven Gruppen unterstützen. Wie im folgenden Abschnitt beschrieben ist, traten bei dem Zusammentreffen mehrerer Multicastströme Paketverluste auf.

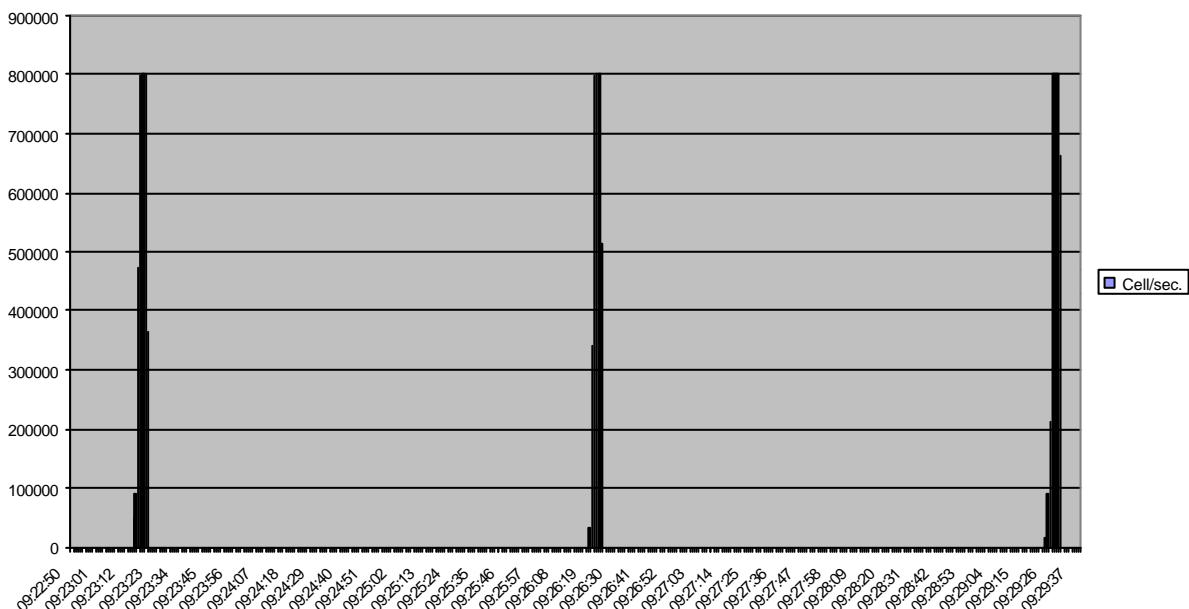


Abbildung 4-1: Multicast-Last bei PIM dense-mode

Abbildung 4-1 zeigt das typische flood&prune-Verhalten bei PIM-DM, wenn keine Empfänger aktiv sind. Alle 3 Minuten entstehen Lastspitzen von ca. 3-4 Sekunden Dauer. Die Lastspitzen entstehen durch das Fluten der Multicast Pakete bei PIM-DM, wobei während dieser

Zeit alle Multicast-Pakete der Gruppe weitergegeben werden. Besitzt ein down-stream Router keine Empfänger, so schickt er nach Ablauf eines Prune-delay Timers (3 Sekunden) eine Prune-Nachricht an seinen upstream Router. Dieser stoppt daraufhin die Multicast-Weiterleitung ([Wil00]). Als Konsequenz dieser Protokoll-Eigenschaft kommt es im gesamten PIM-DM Multicast-Netz alle 3 Minuten zu massiven Lastspitzen. Dies macht in den meisten Fällen die Verwendung von PIM-DM unmöglich.

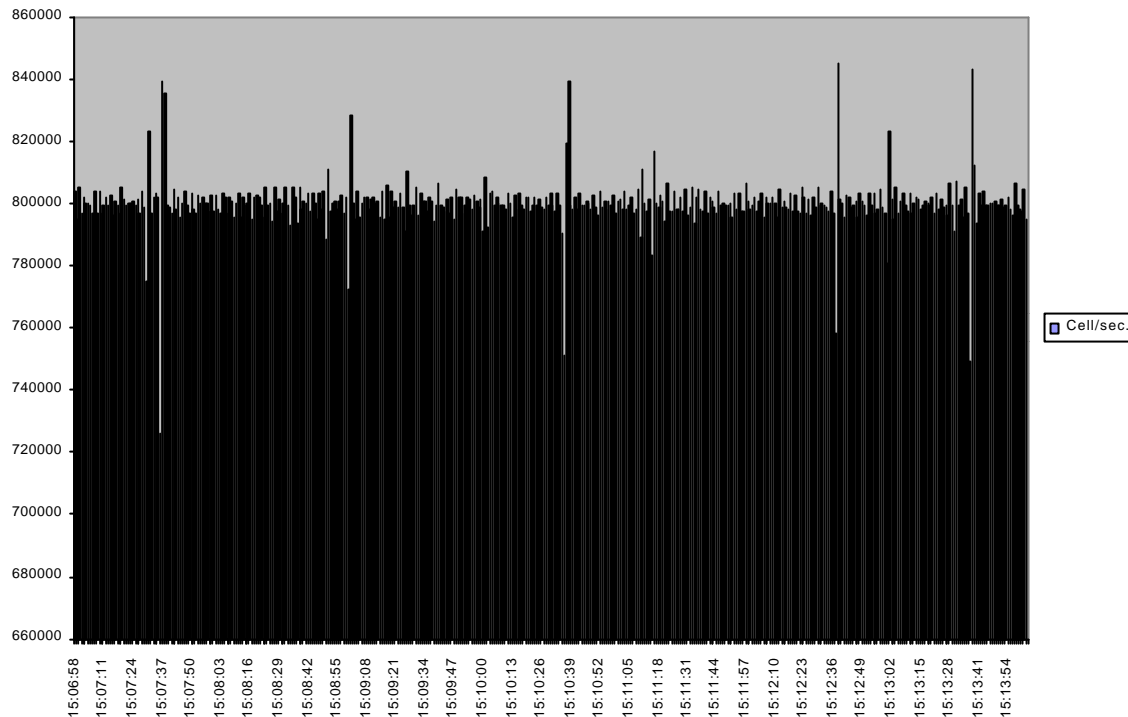


Abbildung 4-2: Multicast-Last bei PIM dense-mode mit fehlerhaftem Pruning

Das Reverse Path Forwarding (RPF) Interface eines IP Multicast Router ist das Interface, auf das die IP Route zum Sender der Multicast-Pakete zeigt. Um Paket-Duplizierungen auszuschließen, akzeptiert ein Router nur Multicast-Pakete, welcher er über das RPF Interface hereinbekommt. Pakete von einem Nicht-RPF-Interface werden gelöscht und normalerweise sendet der Router eine Pruning-Nachricht, um die überflüssigen Pakete zu stoppen. Abbildung 4-2 zeigt das Verhalten des C8540MSR in einem solchen Fall – die Route zum Sender war nicht definiert. Der C8540MSR schickt offensichtlich keine Pruning Nachrichten, so dass die Multicast Pakete ständig an ihn gesendet werden. Der C8540MSR löscht die Pakete, aber nicht in Hardware, sondern mit dem Hauptprozessor auf dem Supervisor Board. Dies führt zu einer völligen Überlastung der Maschine und zu erheblichen Betriebsproblemen. Beide Fehler wurden an Cisco gemeldet und dürften in zukünftigen Softwareversionen behoben werden. Pruning-Probleme treten relativ häufig in Multicast-Netzen auf.

Sinnvollerweise sollten daher Multicast-Routing-Protokolle verwendet werden, bei denen sich die Empfänger explizit für Multicast-Gruppen registrieren (explicit join) und die Zugehörigkeit zu einer Gruppe regelmäßig erneuert werden muss (soft state). Abbildung 19 zeigt die Last zwischen den Multicast Routern bei Verwendung von PIM-SM und keinem registriertem Empfänger. Über den gesamten Messbereich trat so gut wie kein Verkehr am BUS auf. Der upstream Router bei der GMD sendet also keine Multicast-Pakete zum downstream Router bei der Universität Köln. Alle Teile des Multicast-Netzes, in denen keine Empfänger vorhanden sind, werden nicht belastet.

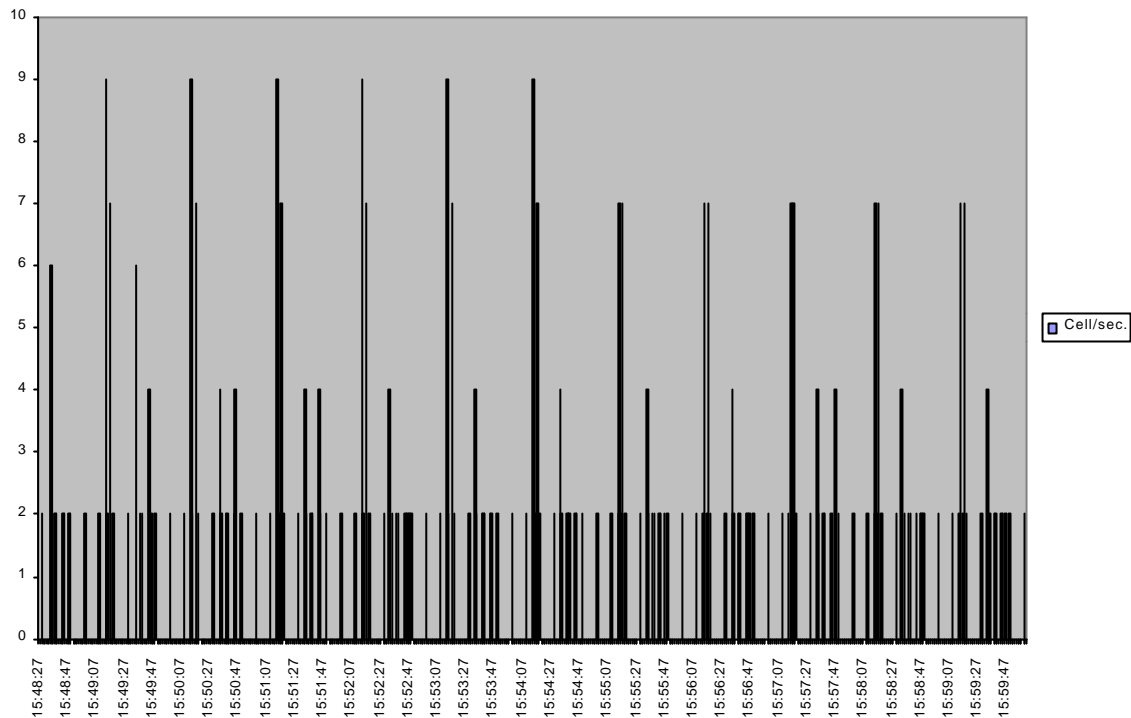


Abbildung 4-3: Multicast-Last bei PIM sparse-mode mit RP

#### 4.3.9 Parallele TCP- und Multicast-Ströme

Ein zentrales Problem bei Multicast mit hohen Geschwindigkeiten ist, dass Multicast-Verkehr massive Engpässe (Congestion) im Netz verursachen kann. Anders als bei TCP gibt es keinen Regelmechanismus, der Engpässe erkennt und die Sendeleistung entsprechend reduziert. Das Problem ist aufgrund der vielen Empfänger in der Regel noch um einiges größer als bei UDP-Unicast-Verkehr.

An Techniken zur Lösung dieses Problems wird seit längerem gearbeitet. Zum einen werden Regelmechanismen für Multicast Transportprotokolle entwickelt (z.B. [Bro99]), damit der Sender die Multicast-Senderate im Falle von Überlast reduziert. Zum anderen kann mit geeigneten Queueing-Mechanismen eine faire Aufteilung der vorhandenen Bandbreite zwischen den einzelnen Datenströmen im Netz erzielt werden. Ein gängiger Ansatz ist dabei das Generalized Processor Sharing (GPS) bzw. das Weighted Fair Queueing (WFQ) ([Kes98]). Hierbei wird die vorhandene Bandbreite auf die Verkehrsströme gleichmäßig aufgeteilt. Solche Queueing-Mechanismen finden sowohl bei ATM Switchen als auch bei IP Routern Verwendung. Bei modernen ATM Switchen ist per-VC Queueing mit round-robin in Hardware implementiert. IP Router unterstützen Weighted Fair Queueing, aber häufig nur in Software mit geringen Bandbreiten bis 2 Mbit/s.

Von den im Projekt eingesetzten ATM Switchen unterstützen Cisco C8540MSR und FORE ASX-1000/D-Modul per-VC Queueing mit round-robin. Der FORE ASX-4000 und die Cisco LS-1010 haben diese Unterstützung nicht.<sup>5</sup> In neueren Ausführungen beider Geräte ist dieses Queueing allerdings auch implementiert.

Die Wirksamkeit von per-VC Queueing mit round-robin wurde am C8540MSR untersucht. Eine Sun E250 sendet über Gigabit Ethernet IP Multicast zum C8540MSR. Dieser schickt die

<sup>5</sup> Entsprechende Tests mit dem ASX-4000 wurden im Projekt GIGAnet durchgeführt ([GTBW00], S. 13f).

Pakete über ATM zu einem BUS. Eine Sun Ultra 60 sendet über eine HE622-Adapterkarte parallelen TCP-Verkehr über eine direkte ATM-Verbindung. Der Engpass entsteht im ATM Teil des C8540MSR, in dem beide Datenströme auf einer 622 Mbit/s Leitung zusammentreffen.

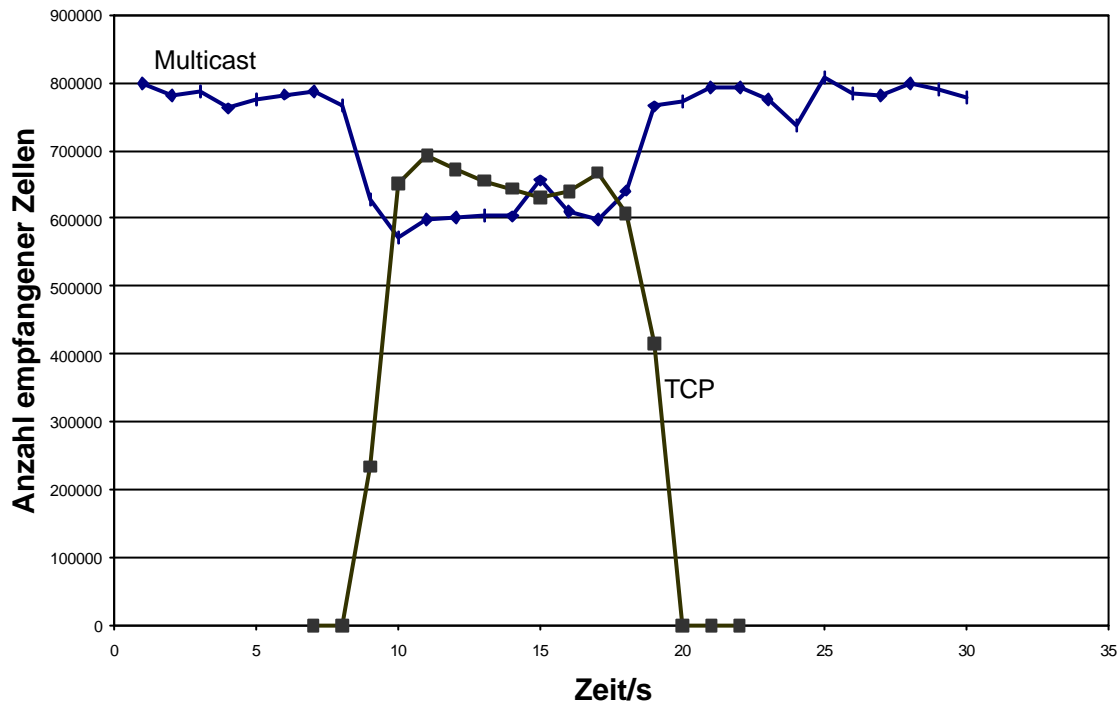


Abbildung 4-1: Parallele TCP-Verbindung, WFQ

Abbildung 4-1 zeigt die Übertragungsraten des IP Multicast und der TCP-Verbindung. Der IP-Multicast-Strom von ca. 300 Mbit/s bzw. 800000 Zellen wird durch den einsetzenden TCP-Datenstrom um ca. 200000 Zellen reduziert, die im C8540MSR verworfen werden. Beide Datenströme bekommen in etwa die gleiche Bandbreite zugeteilt. Mit Hilfe dieser Queueing-Methode lassen sich also negative Auswirkungen von UDP basiertem Multicast begrenzen.

Bei der Untersuchung des Queueing-Verhaltens des C8540MSR auf IP-Ebene stellte sich heraus, dass das Gerät IP Pakete auch in Situationen verliert, in denen keine Überlast auftritt. Um das Problem genauer zu studieren, wurde das Zusammenführen zweier IP-Multicast-Ströme genauer untersucht. Die beiden Suns E250 (mit Sun bzw. Syskonnect Gigabit Ethernet Adapter) senden über Gigabit Ethernet IP Multicast zum C8540MSR, welcher die Pakete über LAN Emulation weiter sendet. Die Anzahl der weiter geschickten Pakete bzw. der ATM Zellen wurde mit Hilfe eines GNN Analyser ermittelt.

Die Senderaten der Suns betragen 2000, 3000, 4000, 5000, 10000, 15000, 20000, 25000 Pakete zu je 1500 Bytes (= 32 ATM Zellen). Für jede einzelne Rate sendete zunächst die E250 mit Sun GE Adapter für einige Sekunden alleine; dann setzte die E250 mit Syskonnect Interface zusätzlich ein; anschließend stoppte die E250 mit Sun GE, so dass nur noch die E250 mit Syskonnect Interface sendete. Auf diese Weise wurde das Forwarding-Verhalten sowohl für die einzelnen Multicast-Ströme als auch für den gemeinsamen Multicast-Strom ermittelt.

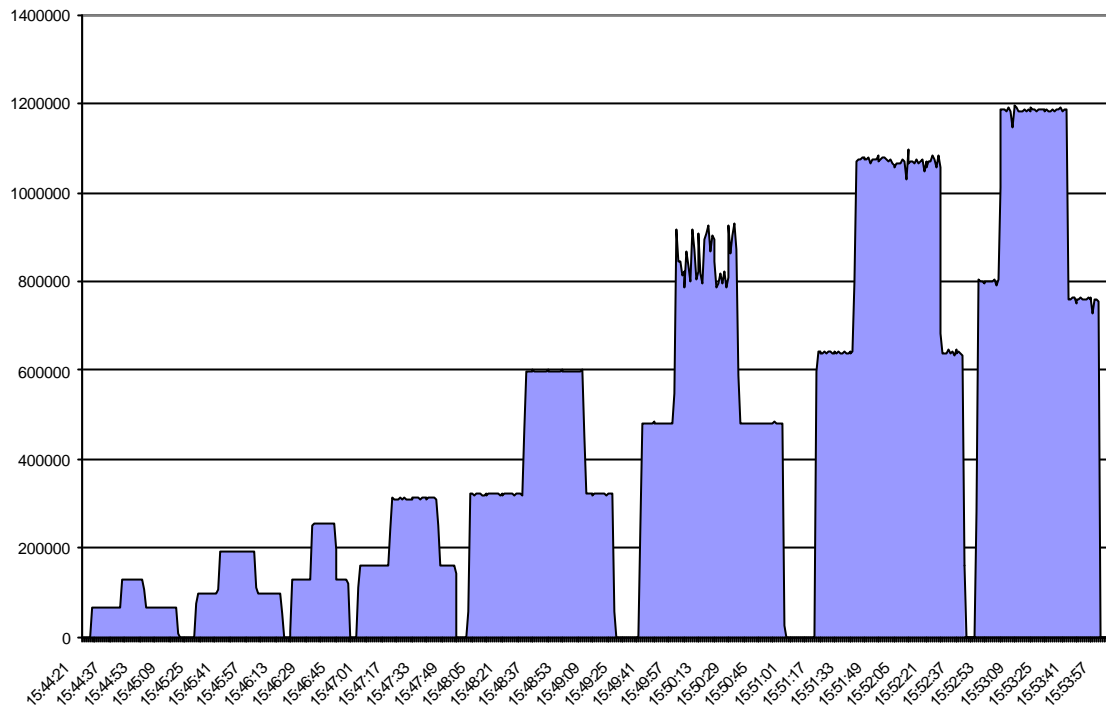


Abbildung 4-2: Zwei parallele Multicast-Ströme im C8540MSR

Abbildung 4-2 zeigt die Ergebnisse. Die Rate für die einzelnen Multicast-Datenströme entspricht - bis auf eine Ausnahme - der mit mgen eingestellten Senderate. Damit ist nachgewiesen, dass sowohl die Suns mit dieser Rate senden, als auch der C8540MSR mit dieser Rate Einzelströme forwarden kann. Die Ausnahme ergibt sich bei einer Senderate von 25000 Paketen/800000 ATM Zellen. In diesem Fall ist die Sun mit Sysconnect GE Interface nicht mehr in der Lage, die Senderate zu generieren. Sie schafft lediglich ca. 23750 Pakete bzw. 760000 Zellen.

Beim Forwarden der beiden Multicast-Ströme müsste der Zellstrom exakt der Summe der zwei Einzelströme entsprechen. Während dies für kleine Senderate (bis 4000 Pakete) der Fall ist, zeigen sich ab 5000 Paketen Abweichungen. Der C8540MSR forwarded nicht mehr alle Pakete bzw. verwirft Pakete. Der Anteil der verworfenen Pakete wird um so höher, je höher die Senderaten werden. Hierbei handelt es sich um einen Fehler des C8540MSR – zumindest bis 20000 Pakete/640000 Zellen passt die Summe von 1280000 Zellen problemlos in eine 622 Mbit/s Leitung. Der Fehler tritt bei Datenraten auf, die der C8540MSR bei Einzelströmen problemlos forwarded. Während bei den Einzelströmen die gemessenen Raten sehr konstant sind, zeigen sich bei den zusammengesetzten Datenströmen bei hohen Bandbreiten zudem einige Schwankungen.

#### 4.3.10 Sendeverhalten der Applikation PLANSIM-T

Die im Projekt weiter entwickelte Simulationssoftware PLANSIM-T wurde so ausgelegt, dass die Verkehrsdaten für die Visualisierung sowohl über ATM Multicast/XTI wie auch über UDP Multicast (LANE) verschickt werden können. Für beide Ansätze wurde die erreichte Senderate in Abhängigkeit von der in der Simulation verwendeten Paketgröße gemessen. Auf dem Sender wurde mit dem Tool „adstat“, das zusammen mit dem FORE Treiber installiert wird, die Anzahl der hinaus gehenden ATM Zellen gemessen.

Die von PLANSIM-T erzeugte Datenmenge und deren zeitlicher Verlauf (Burstiness) ist wesentlich von dem zugrunde liegenden Verkehrsnetz und vom Grad der Parallelisierung abhän-

gig. Große Simulations-Szenarien z.B von Städten wie Berlin führen in der Regel zu sehr ungleichmäßigen Netzlasten und damit Messergebnissen. Wir wählten daher für die Messungen das kleine Verkehrsnetz „Autobahnkreuz“, da es eine gleichmäßige und hohe Datenlast produziert. Dies führt zu reproduzierbaren Ergebnissen. Auf der anderen Seite zeigten unsere Messungen, dass in einem kleinen Netz wie dem Autobahnkreuz kein Leistungsgewinn durch Parallelisierung erreichbar ist.

PLANSIM-T erreicht in der Ein-Prozessor Version Übertragungsraten von etwa 120 – 145 MBit/s (Abbildung 4-1). Wie bei den Leistungs-Untersuchungen mit mgen werden bei der Verwendung von größeren Paketen höhere Senderaten erzielt. Die von PLANSIM-T erzeugten Bandbreiten liegen noch weit unterhalb der theoretischen Grenzen (Abschnitt 4.3.2) und der mit Testprogrammen ermittelten praktischen Grenzen (Abschnitt 4.3.4) wurden, so dass bei einer höheren Simulations-Leistung noch deutlich höhere Übertragungsraten zu erzielen sind. Für genauere Feststellungen müsste eine detaillierte, Prozess bezogene Leistungsanalyse durchgeführt werden.

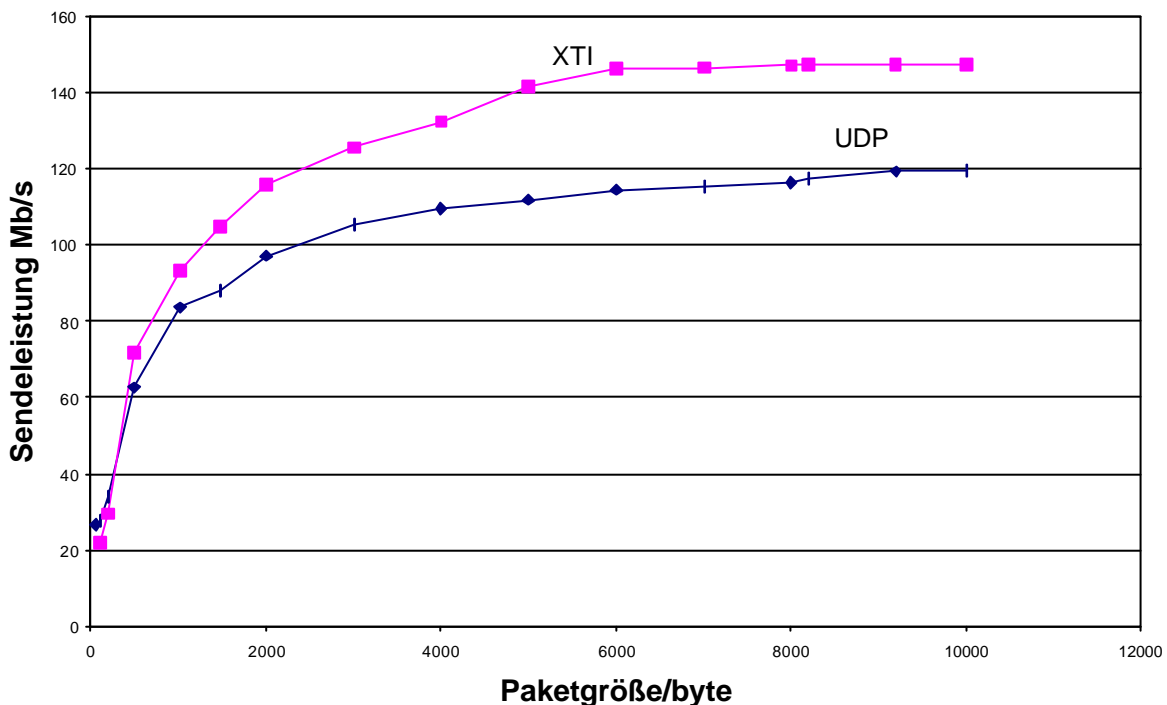


Abbildung 4-1: Sendeleistung von PLANSIM-T auf der E250 über LANE und XTI

Ein kritischer Punkt ist die schnelle Visualisierung der Daten auf den Empfängern. Wie in Kapitel 4.3.7 beschrieben, kommt es schon bei Testprogrammen - die Daten nur entgegennehmen, ohne sie weiter zu bearbeiten - zu erheblichen Datenverlusten. Das Problem verschärft sich, wenn die empfangenen Daten für 2D- und 3D-Visualisierungen noch aufwendig bearbeitet werden müssen.

#### 4.4 Zusammenfassung und Ausblick

Im Gigabit Testbed West hat die GMD im Rahmen des 3V-Projektes Methoden untersucht, um Multicast Daten mit hohen Bandbreiten über ATM und Gigabit Ethernet basierte Netze zu verteilen. Drei Ansätze wurden intensiver betrachtet und praktisch umgesetzt, so dass sie von der Anwendung (Verteilte Verkehrssimulation und Visualisierung) genutzt werden konnten.

Für die Anwendung wurden vom Projektpartner Universität Köln IP-Multicast und ATM-XTI basierte Schnittstellen realisiert.

Die drei verschiedenen Ansätze sind:

- direkte Nutzung von ATM Punkt-zu-Mehrpunkt Verbindungen mittels der XTI-Schnittstelle
- Verteilung von IP-Multicast in einem gemeinsamen IP Subnetz über LAN Emulation
- Verteilung von IP-Multicast über IP Multicast Router

Jeder dieser Ansätze erfüllt die im Projekt gestellte Zielsetzung, Multicast mit hohen Bandbreiten an kleine Gruppen von Empfängern zu verteilen. Die direkte ATM-Nutzung mittels XTI bewirkt die beste Verteilung der Daten, arbeitet aber nur in reinen ATM-Netzen und wird nur von einem Anbieter unterstützt. LAN Emulation ist einfach zu realisieren und zu nutzen. Die Datenverteilung hat aber einige Nachteile, die zu erhöhten Belastungen des Netzes und der beteiligten Endgeräte führen können. Die IP Router basierte Multicast-Verteilung ist wie bei LANE einfach zu nutzen, gestattet aber eine bessere Verteilung der Daten. Die Auswahl der Multicast-Routing-Protokolle hat mit Sorgfalt zu erfolgen und teilweise haben die Geräte noch Probleme bei hohen Bandbreiten.

Durch Tuning kann die Übertragungsleistung zum Teil beträchtlich erhöht werden. Hierzu gehören z.B. die Verwendung von Interrupt Coalescing, großen UDP-Puffern und großen Paketen. Insgesamt ist jedoch eine klare Abstufung festzustellen:

- Die Technologien ermöglichen theoretische Übertragungen von ca. 540 Mbit/s (XTI, LANE) bzw. 965 Mbit/s (GE).
- Mit der verfügbaren Hardware in Endgeräten sind mit Testprogrammen und gängigen Paketgrößen von 1500 Bytes nur Bandbreiten zwischen 200-300 Mbit/s erreichbar. Lediglich mit sehr großen Paketen sind die theoretischen Grenzen bei den ATM basierten Verfahren erreichbar. Bei den Empfängern gibt es im Grenzbereich zum Teil erhebliche Datenverluste.
- Die Anwendung erreicht im Sendeverhalten in der nicht-parallelisierten Version maximal ca. 120 Mbit/s. Bei dieser Datenrate sind die Empfänger in der Regel durch die Visualisierung bereits so überlastet, dass es zu erheblichen Datenverlusten kommt.



## 5 Zusammenfassung

Ein großes Problem, mit dem sich moderne Verkehrssimulationen auseinander setzen müssen, ist die Art und Weise wie die gewonnenen Daten dargestellt werden können. Hierbei ist es häufig wünschenswert, daß eine Simulation die Daten für mehrere Visualisierungen liefert, die von verschiedenen Personen mit unterschiedlichen Zielen gleichzeitig genutzt werden können. Verantwortliche wie Verkehrsplaner, Politiker, Wissenschaftler und nicht zuletzt einzelne Verkehrsteilnehmer können so individuell Nutzen aus der Simulation ziehen.

Mit neuen Netztechnologien zur Verteilung von Multicast mit hohen Bandbreiten, wie sie in den Gigabit Testbeds des DFN Vereins realisiert wurden, ist es nun erstmals möglich geworden, eine Verkehrssimulation mit mehreren verteilten Visualisierungen zu realisieren.

Ziel des Projektes 3V - Verteile Verkehrssimulation und Visualisierung - war es, die bestehende Verkehrssimulation PLANSIM-T zu erweitern, um eine oder mehrere Visualisierungen über ein Netzwerk mit Gigabit-Technik anzuschließen. Es existierten die folgenden Arbeitsschwerpunkte:

### 1. **Aufbau einer Multicast-Infrastruktur und Analyse von Multicast-Technologien**

Es wurden mehrere Verfahren zur Multicast-Verteilung realisiert (XTI, LANE und IP Router basiert) und Parameter für das Tuning der Übertragung bestimmt. Mit geeigneten Verfahren und der richtigen Wahl der Parameter wurden Übertragungsraten bis über 500 Mbit/s erreicht. Bei hohen Datenraten kommt es auf Empfängerseite jedoch häufig zur Überlastung mit resultierenden Datenverlusten bei der Übergabe der Daten an die Anwendung.

### 2. **Trennung der eingebauten Visualisierung von der Simulation und Entwicklung von Kommunikationsschnittstellen**

Die bei Projektbeginn bestehende enge Verzahnung zwischen Simulation und Visualisierung wurde aufgelöst, so dass es nun möglich ist, an eine Simulation mehrere unterschiedliche Visualisierungen anzuschließen. Hierfür wurden Kommunikations-Schnittstellen für die Nutzung sowohl von ATM-Multicast als auch von IP-Multicast Verteilungen entwickelt.

### 3. **Erstellen einer 2D- und einer 3D-Visualisierung mit Rückkanal**

Es wurde eine 2D-Version als Beispiel für eine technisch-wissenschaftliche Visualisierung und eine 3D-Version als Beispiel für eine Virtual-Reality Visualisierung entwickelt. Über einen Rückkanal ist ein interaktiver Eingriff, wie z.B. das Sperren einer Straße, von der Visualisierung auf die Simulation möglich. Die 3D-Visualisierung wurde in einem Video demonstriert.

Probleme bestehen noch im Zusammenspiel der Komponenten, insbesondere bei der Synchronisation zwischen Simulation und Visualisierungen. So erwartet z.B. der Betrachter einer Virtual-Reality-Visualisierung den Ablauf der Szene in Echtzeit, so dass die Simulation ein Realzeit-Verhalten aufweisen muss.

Insgesamt gesehen, ist es mit der GTB Netztechnik nun möglich, eine Verkehrssimulation auf einem zentralen Rechner ablaufen zu lassen und die Daten über ein Hochgeschwindigkeits-Netz an mehrere verschiedene Arbeitsplätze für die Visualisierung zu verteilen.



## 6 Berichte und Vorträge

Die folgenden Berichte und Vorträge wurden im Rahmen des Projektes gehalten.

- [3V] 3V – Verteilte Verkehrssimulation und Visualisierung, DFN-Projekt, <http://www.3V.Uni-Koeln.DE>.
- [MM00] Marcus Metzler: 3V – Verteilte Verkehrssimulation und Visualisierung, Gigabit West Abschlußkolloquium, Jülich, 1. Februar 2000.
- [ZD99] Lothar Zier, Gundula Dörries: Multicast im GTB-West, Technischer Bericht, November 1999. (Die aktualisierte Fassung ist Teil dieses Abschlussberichtes)
- [LZ00] Lothar Zier: High Speed Multicast, DFN-Symposium, 15. Februar 2000.
- [MZ99] Marcus Metzler, Lothar Zier: 3V – Verteilte Verkehrssimulation und Visualisierung, Gigabit West Treffen, Sankt Augustin, 22. Oktober 1999.



## 7 Literatur

- [Bro99] Frank Brockners: Flußsteuerung für verlässliche Muticastkommunikation im Internet, Dissertation, Universität zu Köln, Shaker Verlag, Aachen, 1999.
- [GIGAW]Gigabit Testbed West, DFN-Projekt, <http://www.dfn.de/projekte/gigabit/gtb-west.html>.
- [GTBW00] Gigabit Testbed West, Abschlußbericht, März 2000.
- [Kes98] S. Keshav: An Engineering Approach to Computer Networking, Addison-Wesley, 1998.
- [LANE1] ATM Forum: LAN Emulation over ATM, Version 1.0, Januar 1995.
- [LANE2] ATM Forum: LAN Emulation over ATM Version 2 – LUNI Specification, Juli 1997.
- [LNNI2] ATM Forum: LAN Emulation over ATM Version 2 – LNNI Specification, Februar 1999.
- [MGEN] Naval Research Laboratory (NRL) "Multi-Generator" (MGEN) Toolset, Version 3.1, <ftp://manimac.itd.nrl.navy.mil/Pub/MGEN/index.html>.
- [RFC2022] IETF: Support for Multicast over UNI 3.0/3.1 based ATM Networks, RFC 2022, November 1996.
- [RFC2121] IETF: Issues affecting MARS Cluster Size, G. Armitage, RFC2121, March 1997.
- [RFC2149] IETF: Multicast Server Architectures for MARS-based ATM multicasting. R. Talpade, M. Ammar, RFC 2149, May 1997.
- [RFC2337] Farinacci, et al.: Intra-LIS IP multicast among routers over ATM using Sparse Mode PIM, RFC 2337, April 1998.
- [RFC2362] IETF: Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification, RFC 2362, June 1998.
- [RFC2443] IETF: A Distributed MARS Service Using SCSP. J. Luciani, A. Gallo, RFC 2443, November 1998.
- [UNI31] ATM Forum: User-Network Interface Specification – Version 3.1, September 1994.
- [Wil00] Beau Williamson: Developing IP Multicast Networks – Volume I, Cisco Press, 2000.
- [XTI] Open Group: Networking Services, Issue 5, February 1997.
- [KN] Kai Nagel, High-speed micro simulations of traffic flow, Dissertation, Universität zu Köln, 1994.
- [SK] Stefan Krauß, Microscopic Modelling of Traffic Flow: Investigation of Collision-Free Vehicle Dynamics, Dissertation, Universität zu Köln, 1998.
- [RW] Ralf Wimmershoff, Untersuchungen zur parallelen, objektorientierten Verkehrssimulation, Dissertaion, Universität zu Köln, 2000.



## 8 Anhang

Im Anhang werden die konkreten Konfigurationen für die im 3V-Projekt eingesetzten Geräte der Firmen Cisco und FORE aufgeführt. Im Einzelnen werden Konfigurationen für LANE, XTI, und IP Router beschrieben.

### 8.1 LANE

Auf Basis von LANE sind mehrere IP-Netze mit unterschiedlichen Funktionalitäten realisiert. Ziel ist es, unterschiedliche Methoden der Verkapselung (Ethernet und Token Ring) sowie unterschiedliche MTUs (1500, 4490/4528, 9180/9218, 18136) für Testzwecke zur Verfügung zu stellen. Hauptsächlich werden die Netze 3v (Ethernet Verkapselung mit 1500 Bytes MTU, IP 192.168.3.x) und 3vv (Ethernet Verkapselung mit 9218 Bytes MTU, IP 192.168.4.x) verwendet.

Ein LECS auf einem ATM-Switch in der GMD erlaubt den Projektteilnehmern eine einfache Konfiguration ihrer Klienten. LES/BUS-Systeme sind auf FORE und Cisco Geräten bei der GMD eingerichtet und verteilen von dort aus den Multicast-Verkehr zu allen Klienten. LANE Klienten sind Sun und SGI Maschinen mit FORE HE622 ATM-Adaptern und Gigabit Ethernet Karten diverser Hersteller. Der Übergang von ATM nach Gigabit Ethernet erfolgt über einen Cisco Catalyst 5505 Ethernet Switch. Der C8540MSR hat Probleme beim Bridgen vom ATM/LANE auf Gigabit Ethernet und konnte wurde daher nur für das IP Routing benutzt.

#### 8.1.1 LECS

Der LECS läuft auf einem FORE ASX-4000 ATM Switch mit ForeThought\_6.1.0 Software. Dieser verwendet eine ANYCAST Adresse, so dass er beliebig ausgetauscht werden kann. Die Adresse lautet:

```
0x39.276f.31.000119.0000.0002.0000.111100000002.31
```

In der Konfigurationsdatei lecs.gig werden 3 ELANs definiert:

```
Match.Ordering: 3v, 3vv, 3v-tr4544
# Ethernet, 1500 Bytes MTU, 192.168.3.x
3v.Address: 39.276F.31.000119.0000.0002.0000.200000000010.10
3v.Accept: XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
3v.ELAN_Identifier: 3
# Ethernet, 9280 Bytes MTU, 192.168.4.x
3vv.Address: 39.276F.31.000119.0000.0002.0000.200000000012.12
3vv.Accept: XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
3vv.ELAN_Identifier: 4
3vv.Maximum_Frame_Size: 9234
# Token Ring, 4490 Bytes MTU, 192.168.83.x
3v-tr4544.Address: 39.276f.31.000119.0000.0002.0000.200000000004.04
3v-tr4544.Accept: XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
3v-tr4544.ELAN_Identifier: 83
3v-tr4544.LAN_Type: IEEE 802.5
3v-tr4544.Maximum_Frame_Size: 4544
```

Die Datei wird entsprechend den Anforderungen modifiziert, um z.B. weitere Testnetze hinzuzufügen.

Der LECS wird auf dem ATM Switch unter ForeThought\_6.1.0 folgendermaßen aktiviert:

```
services lane lecs new -atmaddress 31 -adminstatus up -database lecs.gig -lesflag disabled -
wkaflag other -wka 0x39.276f.31.000119.0000.0002.0000.111100000002.31
```

## 8.1.2 LES/BUS

Im 3V-Projekt finden LES/BUS Server von Cisco und Sun/FORE Verwendung. LES und BUS sind immer zusammen realisiert. Es werden ATM ANYCAST-Adressen verwendet, um die Server unabhängig von konkreten Geräten und von den ATM Switchen zu halten, an die sie angeschlossen sind. LES/BUS wurden für Ethernet und Token Ring ELANs aufgesetzt.

### Cisco

Der Cisco Catalyst C5505 Ethernet Switch ist mit einem 622 ATM-Interface (WS-X5161) ausgestattet und dient als LES/BUS für einen Teil der Netze. Als Betriebssystem wird auf dem ATM Interface IOS 11.3 eingesetzt.

Cisco unterstützt beim WS-X5161 ATM Interface unter IOS 11.3 nur die beiden folgenden Verkapselungen/MTUs:

- Ethernet mit 1500 Bytes
- Token Ring mit 4490 Bytes

Andere MTU-Größen lassen sich zwar konfigurieren (mit dem mtu Interface-Kommando), aber das Gerät arbeitet nicht ordnungsgemäß, sondern verwirft Pakete über einer bestimmten Größe.

Ethernet LES/BUS Konfiguration mit 1500 Bytes MTU:

```
interface ATM0.10 multipoint
  description LANE Server 3V
  mtu 1500
  lane server-atm-address 39.276F31000119000000020000.200000000010.10
  lane server-bus ethernet 3v
```

Token Ring LES/BUS Konfiguration mit 4490 Bytes MTU:

```
interface ATM0.4 multipoint
  description LANE Server 3v-tr4544
  mtu 4490
  lane server-atm-address 39.276F31000119000000020000.200000000004.04
  lane server-bus tokenring 3v-tr4544
```

Der Token Ring LES/BUS läuft aufgrund eines Cisco Fehlers zur Zeit nicht mit FORE Klienten.

### FORE

Eine Sun Workstation (Solaris 7) mit FORE HE622 ATM Adapter (ForeThought\_5.2) wird ebenfalls als LES/BUS eingesetzt. Vorteil dieser Lösung ist, dass FORE viele unterschiedliche MTUs unterstützt (Ethernet: 1500, 4528, 9218; Token Ring 4490, 9180, 18136). Die Klienten können so auch größere Pakete mittels Multicast austauschen und höhere Übertragungsraten nutzen.

Die Definition eines LES/BUS erfolgt in der Datei /etc/fore/fore\_lanem.conf (in der auch die Klienten definiert werden):

```
# LES/BUS #0
LES_BUS_OCT0="0x80 0x80"
LE_NAME0='3vv'
LE_ELANID0=4
LE_TYPE0=ethernet
LE_SIZE0=9234
```

```
LE_UNIT0=0
LE_ANYCAST0='39.276F.31.000119.0000.0002.0000.200000000012.12'
LE_SECURITY0='Well-Known'
```

Der Server wird anhand dieser Konfiguration wie folgt gestartet:

```
lnni -name 3vv -id 4 -type ethernet -size 9234 -unit 0 -secure -ibus 0x80 0x80
```

Verteilte Server – wie sie z.B. mit FORE Adaptern realisierbar sind – wurden bisher nicht eingerichtet.

### 8.1.3 Klienten

LANE-Klienten werden von FORE (Sun und SGI Maschinen mit FORE HE622 Adaptern; ForeThought\_5.2) und Cisco (C5505 für den Übergang zu Gigabit Ethernet, IOS 11.3) eingesetzt.

Die Definition eines LANE-Klienten (LEC) bei FORE erfolgt in der Datei /etc/fore/fore\_lanem.conf:

```
# LEC #0
INTF_ELAN0=e10
UNIT_ELAN0=0
TYPE_ELAN0=ethernet
MPOA_ELAN0='-nompoa'
HA_ELAN0=''
LECS_NSAP_ELAN0=--wellknown
ELAN0='3v'
# LEC #1
INTF_ELAN1=e11
UNIT_ELAN1=0
TYPE_ELAN1=ethernet
MPOA_ELAN1='-nompoa'
HA_ELAN1=''
LECS_NSAP_ELAN1=--wellknown
ELAN1='3vv'
# LEC #2
INTF_ELAN2=e12
UNIT_ELAN2=0
TYPE_ELAN2=ethernet
MPOA_ELAN2='-nompoa'
HA_ELAN2=''
LECS_NSAP_ELAN2=--wellknown
ELAN2='3v-tr4544'
```

Soll beim LECS eine andere als die ATM Forum wellknown Adresse verwendet werden, kann bei LECS\_NSAP\_ELAN# diese ATM-Adresse eingetragen werden.

Es gibt zwei Möglichkeiten, Multicast bei Sun Maschinen mit FORE Adapter zu tunen:

- **UDP Sende- und Empfangspuffer** sind normalerweise 8192 Bytes groß; sie können z.B. mit ndd auf größere Werte (bis maximal udp\_max\_buf = 262144) konfiguriert werden:  
/usr/sbin/ndd -set /dev/udp udp\_xmit\_hiwat <value>  
/usr/sbin/ndd -set /dev/udp udp\_recv\_hiwat <value>
- **Interrupt Coalescing** ist das Zusammenfassen von mehreren IP-Paketen auf dem Netzwerk-Adapter, bevor ein Interrupt ausgelöst wird. Die Aktivierung erfolgt in der Datei /etc/system:  
set fore\_he:he\_intr\_coalesce=1

Bei Aktivierung von Interrupt Coalescing sollten die UDP-Puffer auf den Maximalwert von 65535 gesetzt werden.

Beim Cisco C5505 ist das Management von Ethernet und ATM getrennt und erfolgt auf zwei verschiedenen Modulen (Supervisor und ATM Interface Modul):

```
interface ATM0.3 multipoint
  description LANE Client 3v
  lane config-atm-address 39.276F31000119000000020000.111100000002.37
  lane client ethernet 3 3v
```

Wird ein LECS mit wellknown Adresse benutzt, kann die „lane config-atm-address“ Zeile entfallen. Bei „lane client“ bezeichnet die Zahl 3 hinter ethernet die Nummer des VLAN, das auf der Ethernetseite verwendet wird. Die Konfiguration im Ethernet-Teil lautet:

```
set vlan 3 name 3v type ethernet mtu 1500 said 100003 state active
set vlan 3 1/1-2
```

## 8.2 XTI

Die ATM-Erweiterung der XTI-Schnittstelle ist für FORE Adapter realisiert. Sie kann von Sun und SGI Maschinen aus genutzt werden. In der Praxis existieren Unterschiede zwischen den Versionen der FORE Software (z.B. bei den unterstützten Verkehrsklassen). Die Versionen der eingesetzten Adapter Software sollten gleich sein, um problemlos Programme übernehmen zu können.

Beispielprogramme – auch für die Nutzung von Punkt-zu-Mehrpunkt Verbindungen - stehen in /opt/FOREatm/examples/xti zur Verfügung. Sinnvoll ist die Nutzung des ATM-Namensdienstes, der die Verwendung von aussagekräftigen Namen anstatt von ATM-Adressen erlaubt. Hierfür muss die Datei /etc/xti\_hosts konfiguriert werden:

```
/etc/xti_hosts
#
# /etc/xti_hosts auf baloo.gmd.de
#
# Own XTI address
39.276f.31.000119.0000.0002.0050.00204800005f.99 baloo-xti
39.276f.31.000119.0000.0002.0050.00204800005f.99 localhost
# 3V XTI adresses
39.276f.31.000119.0000.0002.0032.002048000047.99 linus-xti
39.276f.31.000119.0000.0002.0033.002048000074.99 sarafina-xti
39.276f.31.000119.0000.0004.0012.0020482d02e3.00 sparcy02-xti
```

Werden Maschinen auf andere ATM Switches umgehängt, so ändert sich auch deren ATM-Adresse und die /etc/xti\_hosts Dateien aller beteiligten Rechner müssen geändert werden.

## 8.3 IP Router

Als IP Multicast Router fanden im 3V-Projekt Cisco C8540MSR Geräte mit dem Betriebssystem IOS 12.0(7)W5(15c) Verwendung. Bei dem C8540MSR handelt es sich um ein multifunktionales Gerät, welches mit entsprechenden Interfaces sowohl als ATM Switch als auch als Ethernet Switch und IP Router arbeitet. Die im 3V-Projekt verwendeten Geräte bei der GMD und der Universität Köln sind mit ATM 622 Mbit/s und Gigabit Ethernet Schnittstellen ausgestattet.

Die Verbindung der Zell-basierten ATM-Schnittstellen mit den Paket-basierten Schnittstellen wie Gigabit Ethernet erfolgt über ein spezielles Hardware-Modul namens ATM Router Mo-

dule (ARM). Der ARM erlaubt die Definition von LANE-Klienten und realisiert einen Übergang mit maximal 2 GBit/s zu den Paket-basierten Schnittstellen.<sup>6</sup> Dieses Konzept ermöglicht einen extrem flexiblen und leistungsfähigen Einsatz des C8540MSR:

- Betrieb als ATM Switch an den ATM Interfaces: Hierüber wurden sowohl leistungsfähige Endgeräte, andere Ethernet Switche (als BUS), sowie andere ATM Switche angebunden.
- Betrieb als Router: Einzelne LANE-Kleinten auf dem ARM und Gigabit Ethernet Schnittstellen können direkt mit einer IP Adresse versehen werden. Das Routing erfolgt in der Regel Hardware-basiert, so dass hohe Übertragungsraten möglich sind.
- Betrieb als Ethernet Switch: Die Kombination mehrerer Gigabit Ethernet und LANE Klienten erfolgt in sogenannten Bridge Groups mit BVI. Aufgrund von Software-Problemen ist kein Multicast-Betrieb auf BVI Schnittstellen möglich – in älteren Softwareversionen traten Fehler auf; in 12.0(7)W5(15c) lässt sich Multicast für BVI Schnittstellen nicht konfigurieren.

Cisco unterstützt zwei Multicast-Routing-Protokolle, die im 3V-Projekt sinnvoll einsetzbar sind: PIM dense-mode und PIM sparse-mode.

Bei Verwendung von PIM muss eine Unicast Route zum Sender auf dem Interface existieren, in dem die Pakete hereinkommen (RPF Interface), da der Router die Pakete sonst verwirft. Beim C8540MSR erfolgt das Verwerfen nicht in Hardware, sondern auf der CPU des Supervisor-Moduls, was zu einer extremen Belastung und zu Betriebsstörungen führen kann.

Die Konfiguration von PIM ist einfach. Global muss Multicast Routing aktiviert werden und auf jedem Interface muss „ip pim sparse-dense-mode“ konfiguriert werden:

```
ip multicast-routing
!
interface ATM9/0/0.4 multipoint
 ip address 192.168.86.90 255.255.255.0
 no ip directed-broadcast
 ip pim sparse-dense-mode
 lane config-atm-address 39.276F31000119000000020000.111100000002.31
 lane client ethernet 3v-gmd-uni
!
interface GigabitEthernet10/0/0
 ip address 192.168.90.90 255.255.255.0
 ip pim sparse-dense-mode
 no ip directed-broadcast
!
ip route 192.168.5.0 255.255.255.0 192.168.86.250 100
```

Auf einem Router muss zudem für sparse-mode ein Rendezvous Point (RP) eingerichtet werden:

```
ip pim send-rp-announce ATM9/0/0.4 scope 1 group-list 10
ip pim send-rp-discovery scope 2
access-list 10 permit 239.255.255.0 0.0.0.255
```

---

<sup>6</sup> Aufgrund von Hardware-Beschränkungen (SAR Chips mit maximal 622Mbit/s; Anbindung des ARMs an die Paket-basierten Interfaces mit 2 mal Gigabit Ethernet) steht für einzelne Verbindungen weniger Bandbreite zur Verfügung – unter Berücksichtigung des Overhead ca. 515 Mbit/s für Multicast Daten. Trotz dieser Einschränkung ist der C8540MSR neben dem Cisco 12000 der mit Abstand leistungsfähigste Multicast Router mit ATM Schnittstelle.

Die Verwendung des RP wird mit group-list und scope eingeschränkt. Der RP ist nur für die Multicast-Gruppen in der Access-Liste 10 aktiv und wird nur von PIM Multicast Routern benutzt, die maximal 2 Hops entfernt sind.