

Lehrstuhl für Kommunikationsnetze

Technische Universität München

Reaktive verteilte
Verkehrslenkungsverfahren für
MPLS-Weitverkehrsnetze

Stefan Butenweg

Lehrstuhl für Kommunikationsnetze

**Reaktive verteilte
Verkehrslenkungsverfahren für
MPLS-Weitverkehrsnetze**

Stefan Butenweg

Vollständiger Abdruck der von der Fakultät für
Elektrotechnik und Informationstechnik der Technischen Universität München
zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr.-Ing. H.-G. Herzog

Prüfer der Dissertation: 1. Univ.-Prof. Dr.-Ing. J. Eberspächer
2. Univ.-Prof. Dr.-Ing. G. Färber

Die Dissertation wurde am 01.04.2004 bei der Technischen Universität München
eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik
am 23.08.2004 angenommen.

Vorwort

Diese Arbeit ist im Rahmen meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Lehrstuhl für Kommunikationsnetze der Technischen Universität München entstanden.

Mein besonderer Dank gilt meinem Doktorvater Prof. Dr.-Ing. Jörg Eberspächer, der durch zahlreiche Diskussionen, konstruktive Anmerkungen und seine freundliche Unterstützung wesentlich zum Gelingen der Arbeit beigetragen hat.

Herrn Prof. Dr.-Ing. Georg Färber möchte ich meinen herzlichen Dank für die Übernahme des Zweitgutachtens aussprechen.

Des weiteren möchte ich allen Mitarbeitern danken, die zur guten Stimmung und zur kreativen Atmosphäre am Lehrstuhl beigetragen haben. Insbesondere gilt der Dank meinen Kollegen Andrea Bör, Josef Glasmann und Anton Riedl. Gemeinsam sind wir häufig als das NA-Team (Network Architecture Team) den ungelösten Problemen der Telekommunikation erfolgreich gegenüber getreten.

Für die engagierte Mitarbeit danke ich auch allen Diplomanden, welche einen Beitrag zur vorliegenden Arbeit geleistet haben.

Meiner Freundin Barbara, meinen beiden Kindern Johanna und Lukas sowie meinen Eltern Maria Butenweg und Paul Butenweg möchte ich für ihre moralische Unterstützung sowie ihre Geduld danken.

München, im April 2004

Stefan Butenweg

Zusammenfassung

Um einen wirtschaftlichen Betrieb von IP-Netzen bei der zunehmenden Verkehrsdynamik sowie dem anhaltenden Wachstum des Datenvolumens zu ermöglichen, spielen lastabhängige Verkehrslenkungsverfahren in IP-Netzen eine immer wichtigere Rolle. Lastabhängige Verkehrslenkungsverfahren ermöglichen eine zeitlich flexible Anpassung der Verkehrslenkung im Netz zum Ausgleich von temporären lokalen Überlastsituationen.

In den letzten Jahren lag der Fokus der Untersuchungen auf proaktiven Verkehrslenkungsverfahren, welche basierend auf der Kenntnis des Eingangsverkehrs die optimale Verkehrslenkung im Netz festlegen. Der große Nachteil dieser Verfahren ist jedoch die Tatsache, dass keine Reaktion auf unvorgesehene Änderungen der Eingangslast erfolgt. Da bedingt durch die steigende Mobilität der Internet-Teilnehmer sowie die Realisierung neuer hochbitratiger Dienste zukünftig von einer Zunahme solcher Lastschwankungen auszugehen ist, beschäftigt sich diese Arbeit mit reaktiven Verkehrslenkungsverfahren. Diese basieren auf der Messung der Netzauslastung über einen Zeitraum von Sekunden bis Minuten und führen bei Erkennung von Überlastsituationen eine Anpassung der Lastverteilung durch. Da bekannte reaktive Verfahren Schwächen bezüglich zu hoher Komplexität bzw. zu geringer Konvergenzgeschwindigkeit aufweisen, wurde im Rahmen dieser Arbeit ein neues reaktives Verfahren entwickelt.

Das entwickelte verteilte reaktive Verkehrslenkungsverfahren, welches auf der Verwendung von Multi Protocol Label Switching MPLS basiert, kann mit einem Pfadumlegungsansatz und einem Mehrpfadansatz betrieben werden. Im Fall des Pfadumlegungsansatzes werden die Randknoten des MPLS-Netzes unter Verwendung von Label Switched Path LSPs einfach vollvermascht. Die Umverteilung der Last bei dem Auftreten von Überlastsituationen erfolgt über das Rerouting eines oder mehrerer LSPs. Die Verwendung des Mehrpfadansatzes basiert auf einer Vollvermaschung der Randknoten über zwei oder mehrere LSPs. In diesem Ansatz erfolgt die Reaktion auf Überlast durch die Umverteilung der Last auf die bestehenden LSPs.

Die im Rahmen dieser Arbeit durchgeführte Evaluierung des Verfahrens zeigt, dass gegenüber den Konkurrenzverfahren die Komplexität reduziert und die Konvergenzgeschwindigkeit gesteigert werden konnte. Simulative Untersuchungen der Konvergenzgeschwindigkeit weisen nach, dass das entwickelte Verfahren nach Erkennung einer Überlastsituation innerhalb weniger Sekunden konvergiert. Dabei wurde lediglich eine schwache Abhängigkeit der Konvergenzdauer vom Netz- und Auslastungsszenario festgestellt. Weitere Performanz-Analysen belegen, dass das Verfahren bei Verwendung des Mehrpfadansatzes unabhängig vom Netz- und Auslastungs-Szenario die Netzperformanz der optimalen Lastverteilung sehr gut annähert. Der Pfadumlegungsansatz zeigt ein vergleichbares Verhalten in großen Netzen, in welchen eine fein-granulare Verteilung der Last auf die LSPs vorliegt.

Um eine simulative Performanz-Analyse der Verfahren in realistischen Netzszenarien zu ermöglichen, wurde in dieser Arbeit zudem eine skalierbare Simulations-Methode entwickelt, welche auf der ratenbasierten Modellierung des Datenverkehrs basiert.

Abkürzungsverzeichnis

AMPL	A Modeling Language for Mathematical Programming
ATM	Asynchronous Transfer Mode
BGP	Border Gateway Protocol
CLIP	Classical IP over ATM
CR-LDP	Constraint-based Routed Label Distribution Protocol
CSR	Cell Switched Router
DiffServ	Differentiated Services
DNS	Domain Name Service
DUAL	Diffusing Update Algorithmus
ECMP	Equal Cost Multipath
EIGRP	Enhanced Interior Gateway Routing Protocol
FEC	Forwarding Equivalence Class
FTP	File Transfer Protocol
IETF	Internet Engineering Task Force
IGRP	Interior Gateway Routing Protocol
IntServ	Integrated Services
ILP	Integer Linear Program
IP	Internet Protocol
IPv6	Internet Protocol Version 6
ISIS	Intermediate System to Intermediate System
ISO	International Standards Organization
LAN	Local Area Network
LANE	LAN Emulation
LDP	Label Distribution Protocol
LP	Linear Program
LSP	Label Switched Path
LSR	Label Switching Router
MATE	MPLS Adaptive Traffic Engineering
MCF	Multi Commodity Flow
MIP	Mixed Integer Program
MPLS	Multiprotocol Label Switching
MPOA	Multiprotocol over ATM
NLP	Nonlinear Program
OpIATE	Optimization Integrated Adaptive Traffic Engineering
OSI	Open Systems Interconnect
OSPF	Open Shortest Path First
POP	Point of Presence
PPP	Point-to-Point Protocol
QoS	Quality of Service
RFC	Request For Comments
RIP	Routing Information Protocol
RSVP-TE	Resource Reservation Protocol with Traffic Engineering Extensions
RTO	Retransmission Timeout
RTT	Round Trip Time
SDH	Synchronous Digital Hierarchy
SMTP	Simple Mail Transport Protocol
TCP	Transport Control Protocol

TE Traffic Engineering
TTL Time To Live
UDP User Datagram Protocol
WDM Wavelength Division Multiplexing
WWW World Wide Web

Inhaltsverzeichnis

1	Einleitung	1
1.1	Technisches Umfeld	2
1.2	Beitrag der Arbeit	6
1.3	Aufbau der Arbeit	6
2	Grundlagen und Stand der Technik	8
2.1	Grundlegende Konzepte und Protokolle	8
2.1.1	Verkehrslenkung im Internet	8
2.1.1.1	Klassifizierung der IP-Routing-Protokolle	10
2.1.1.2	Bekannte IP-Routing-Protokolle	11
2.1.2	Multiprotocol Label Switching MPLS	13
2.1.2.1	Überblick über IP over ATM Netzkonzepte	13
2.1.2.2	Grundlagen der MPLS-Technologie	15
2.1.2.3	Grundlagen der Label Distribution Protokolle	18
2.2	Realisierungsvarianten lastabhängiger Verkehrslenkung	21
2.3	Definition der optimalen Lastverteilung	22
2.4	Proaktive Verkehrslenkungsverfahren	24
2.4.1	Anforderungen an proaktive Verkehrslenkungsverfahren	24
2.4.2	Grundlagen proaktiver Verkehrslenkungsverfahren	24
2.4.3	Bekannte proaktive Verkehrslenkungsverfahren	25
2.4.3.1	IP-basierte proaktive Verkehrslenkungsverfahren	25
2.4.3.2	MPLS-basierte proaktive Verkehrslenkungsverfahren	26
2.4.3.3	Hybride proaktive IP/MPLS Verkehrslenkungsverfahren	26
2.5	Reaktive Verkehrslenkungsverfahren	27
2.5.1	Anforderungen an reaktive Verkehrslenkungsverfahren	27
2.5.2	Grundlagen reaktiver Verkehrslenkungsverfahren	28
2.5.2.1	Lastabhängige Verkehrslenkung in verbindungslosen Netzen	29
2.5.2.2	Lastabhängige Routing-Algorithmen	30
2.5.2.3	Lastabhängige Mehrpfadansätze	31
2.5.3	Bekannte reaktive Verkehrslenkungsverfahren	32
2.5.3.1	MATE: MPLS Adaptive Traffic Engineering	33
2.5.3.2	Dynamic Traffic Partitioning in MPLS Networks	34
2.5.3.3	OpIATE: Optimization Integrated Adaptive Traffic Engineering	35
2.6	Zusammenfassung des Stands der Technik	35

3	Das verteilte reaktive Verkehrslenkungsverfahren	36
3.1	Das verteilte reaktive MPLS Traffic Engineering System	36
3.2	Das LSP-Monitoring-System	39
3.3	Der Lastverteilungsprozess	39
3.3.1	Ablauf des Lastumverteilungsprozesses	40
3.3.2	Definition einer Überlastsituation	41
3.3.3	Vorgehen zur Berechnung der neuen Lastverteilung	41
3.3.4	Zeitliche Koordination des Lastverteilungsprozesses	43
3.4	Lastverteilungsalgorithmen	45
3.4.1	Linkkosten-Funktion der beiden Lastverteilungsalgorithmen	46
3.4.2	Algorithmus des Pfadumlegungsansatzes	51
3.4.3	Algorithmus des Mehrpfadansatzes	53
3.5	Qualitative Analyse des Verkehrslenkungsverfahrens	55
3.5.1	Bewertung der Stabilität	55
3.5.1.1	Stabilitäts-Analyse bei konstantem Eingangsverkehr	55
3.5.1.2	Stabilitäts-Analyse bei dynamischem Eingangsverkehr	56
3.5.2	Bewertung der Netzperformanz	57
3.5.2.1	Netzperformanz im eingeschwungenen Zustand	57
3.5.2.2	Netzperformanz im transienten Zustand	58
3.5.2.3	Netzperformanz bei dynamischem Eingangsverkehr	59
3.5.3	Bewertung der Konvergenzdauer	59
3.5.4	Bewertung der Komplexität der Lastverteilungsberechnung	60
3.5.4.1	Komplexitäts-Analyse	61
3.5.4.2	Vergleich mit dem Konkurrenz-Verfahren OpIATE	62
3.6	Zusammenfassung wichtiger Aspekte des Verkehrslenkungsverfahrens	63
4	Eine skalierbare ratenbasierte Simulations-Methode	65
4.1	Methoden zur Simulation von IP-Netzen: Stand der Technik	65
4.1.1	Grundlagen zur ereignisgesteuerten Simulationstechnik	66
4.1.2	Simulations-Methode mit paketbasierter Verkehrsmodellierung	67
4.1.2.1	Modellierungsaspekte der Netzkomponenten	67
4.1.2.2	Analyse der Genauigkeit	67
4.1.2.3	Analyse der Skalierbarkeit	67

4.1.3	Simulations-Methode mit ratenbasierter Verkehrsmodellierung	68
4.1.3.1	Modellierung eines Puffers	69
4.1.3.2	Modellierung einer Quelle	71
4.1.3.3	Bestimmung der Performanzparameter	72
4.1.3.4	Analyse der Genauigkeit	73
4.1.3.5	Analyse der Skalierbarkeit	77
4.1.4	Zusammenfassende Bewertung bekannter Simulations-Methoden	81
4.2	Modifizierte ratenbasierte Simulations-Methode	82
4.2.1	Modellierung eines Puffers	83
4.2.2	Modellierung einer Quelle	84
4.2.3	Bestimmung der Performanzparameter	86
4.2.4	Analyse der Genauigkeit	87
4.2.5	Analyse der Skalierbarkeit	88
4.3	Zusammenfassung wichtiger Aspekte der Simulations-Methodik	89
5	Performanz-Analyse anhand ausgewählter Fallstudien	91
5.1	Analyse der Netzperformanz	91
5.1.1	Definition der optimalen Lastverteilung	92
5.1.2	Untersuchung eines 44Knoten-Netzes	94
5.1.2.1	Auswertung des Gleichlast-Szenarios	94
5.1.2.2	Auswertung des Szenarios <i>Lastgranularität</i>	95
5.1.2.3	Auswertung des Szenarios <i>Initiale Lastverteilung</i>	96
5.1.3	Untersuchung eines 14Knoten-Netzes	100
5.1.3.1	Auswertung des Gleichlast-Szenarios	100
5.1.3.2	Auswertung des Szenarios <i>Initiale Lastverteilung</i>	101
5.1.4	Untersuchung eines 45Knoten-Netzes	102
5.1.5	Untersuchung eines 16Knoten-Netzes	105
5.1.6	Zusammenfassende Bewertung der Netzperformanz	107
5.2	Analyse der Konvergenzdauer	108
5.2.1	Vorgehen zur Bestimmung der Timer-Werte	109
5.2.2	Modellierung des MPLS Adaptive Traffic Engineering MATE	110
5.2.3	Untersuchung eines 44Knoten-Netzes	111
5.2.3.1	Festlegung der Timer-Werte	111

5.2.3.2	Auswertung der Konvergenzdauer	112
5.2.4	Untersuchung eines 14Knoten-Netzes	113
5.2.4.1	Festlegung der Timer-Werte	114
5.2.4.2	Auswertung der Konvergenzdauer	114
5.2.5	Zusammenfassende Bewertung der Konvergenzdauer	116
5.3	Analyse des Konvergenzverhaltens bei dynamischem Eingangsverkehr . . .	116
5.3.1	Verkehrsmodellierung	116
5.3.2	Modellierung eines proaktiven Verkehrslenkungsverfahrens	117
5.3.3	Untersuchung eines 44Knoten-Netzes	117
5.3.3.1	Konvergenzverhalten bei großen Laständerungen in zwei Beispielszenarien	117
5.3.3.2	Auswertung des Konvergenzverhaltens in Abhängigkeit von der Verkehrsdynamik	118
5.4	Zusammenfassung der Ergebnisse der Performanz-Analyse	119
6	Zusammenfassung	122

1 Einleitung

Der Wandel zur Informationsgesellschaft in den letzten Jahren hat große Änderungen in der Telekommunikation hervorgerufen. War die klassische Telekommunikation nahezu ausschließlich durch die Telefonie bestimmt, so ist heute das Internet als neues Medium zum Austausch beliebiger digitaler Informationen in den Mittelpunkt gerückt. Ursprünglich geprägt durch die Dienste World Wide Web, Email und den Datei-Transfer entwickelt sich das Internet immer mehr zu einer Multi-Service-Plattform, über welche verschiedenste Telekommunikationsdienste abgewickelt werden [E01]. Hierzu zählen auch die klassischen Telekommunikationsdienste wie die Telefonie.

Im Gegensatz zu den klassischen Telefonnetzen stellen sich an die Netze, auf welchen das Internet basiert, völlig neue Anforderungen. Erstens bedeutet die Realisierung des Internets als Multi-Service-Plattform, dass unterschiedliche Anforderungen der Dienste an die Übertragungsqualität berücksichtigt werden müssen. Zweitens sind diese Netze einem stetigen Wachstum des Datenvolumens ausgesetzt. Drittens zeigt der Datenverkehr im Internet eine zunehmende Dynamik, was die Vorhersagbarkeit des Verkehrs erschwert.

Die Gründe für das anhaltende Wachstum des Datenverkehrs im Internet sind vielfältig. Da ist zunächst die zunehmende Anzahl an neuen Internet-Nutzern zu nennen. Zwar nutzen weltweit heute bereits über 600 Millionen Menschen das Internet, aber ein Ende des Anstiegs ist nicht in Sicht. Des weiteren wirkt sich das Hinzukommen neuer Dienste, Anwendungen und Inhalte auf das Wachstum des Datenverkehrs im Internet aus. Das steigende Angebot erhöht die Attraktivität des Internets und ermöglicht den Nutzern, immer mehr private wie geschäftliche Angelegenheiten über das Internet abzuwickeln. Das resultiert in einer im Durchschnitt steigender Online-Zeit der Nutzer, welche über die letzten Jahre festgestellt wurde. Schließlich stellt die Zunahme der Übertragungskapazität an den Netzzugängen einen weiteren wichtigen Grund für das anwachsende Datenvolumen dar. Das erlaubt den Internet-Nutzern zunehmend die Verwendung hochbitratiger Dienste. Der Einfluss eines hochbitratigen Diensts auf das Wachstum des Datenverkehrs lässt sich am Beispiel des File-Sharing-Diensts erkennen, welcher seit Einführung vor wenigen Jahren heute zeitweise bis zu 50% des gesamten Datenvolumens in einem Netz ausmacht.

Wie für das Ansteigen des Datenvolumens sind auch für die Zunahme der Verkehrsdynamik mehrere Gründe verantwortlich. Einer der Hauptgründe ist die zunehmende Mobilität der Internet-Nutzer. Auslöser hierfür ist einerseits die Wireless LAN (WLAN) Technologie, welche für die sich ausbreitenden drahtlosen Internet-Zugänge eingesetzt wird. Andererseits lassen die anlaufenden Mobilfunksysteme der dritten Generation wie das Universal Mobile Telecommunications System UMTS eine Zunahme der Mobilität erwarten. Diese ermöglichen im Gegensatz zu den aktuellen Mobilfunksystemen einen mobilen hochbitratigen Internet-Zugang.

Ein weiterer Grund für die zunehmende Dynamik ist die steigende Vielfalt der Dienste und Anwendungen. Insbesondere hochbitratige Dienste und Anwendungen können die Dynamik des Eingangsverkehrs merklich beeinflussen. In der Zukunft ist zudem eine weitere Steigerung der Verkehrsdynamik erwarten, wenn sich die neu angedachten Netzbetreibermodelle durchsetzen. Diese sollen einem Netzbetreiber das dynamische Anmieten bzw. Vermieten von Übertragungskapazität über kurze Zeiträume ermöglichen.

Um mit dem anwachsenden Datenvolumen Schritt zu halten, muss ein Betreiber die Übertragungskapazität seines Netzes von Zeit zu Zeit anpassen. Die Anpassung der Übertragungskapazität muss dabei unter dem Gesichtspunkt eines wirtschaftlichen Betriebs

des Netzes erfolgen, welcher durch den Wettbewerb mit anderen Netzbetreibern in den Mittelpunkt gerückt ist. Ein wirtschaftlicher Betrieb des Netzes erfordert eine optimale Ausnutzung der Netz-Ressourcen. Durch die hohe Dynamik und die daraus resultierende schlechte Vorhersagbarkeit des Internet-Datenverkehrs lässt sich dieses Ziel jedoch nur durch eine zeitlich flexible Verteilung der Last über die Netz-Ressourcen erreichen.

Aus diesem Grund nimmt die Bedeutung lastabhängiger Verkehrslenkungsverfahren im Internet immer weiter zu. Lastabhängige Verkehrslenkungsverfahren passen die Verteilung des zu übertragenden Datenverkehrs über die vorhandenen Netz-Ressourcen dynamisch an die aktuelle Auslastungssituation an. Eine temporäre Überlast in einem Teilbereich des Netzes kann durch die Nutzung freier Kapazität in anderen Netzbereichen ausgeglichen werden. Somit können die Netzbetreiber auf absehbare sowie unvorhersehbare Schwankungen des Verkehrs reagieren.

Diese Arbeit beschäftigt sich mit reaktiven Verkehrslenkungsverfahren, welche basierend auf der Messung der Netzauslastung über den Zeitraum von Sekunden bis Minuten die Lastverteilung im Netz anpassen. Diese Verfahren sind für die Reaktion auf unerwartete Lastschwankungen konzipiert. Zur Beschreibung des technischen Umfelds lastabhängiger Verkehrslenkungsverfahren im Internet beschreibt Abschnitt 1.1 den Gesamtkontext des Internet Traffic Engineering. Abschnitt 1.2 erläutert anschließend den Beitrag dieser Arbeit und Abschnitt 1.3 stellt den weiteren Aufbau vor.

1.1 Technisches Umfeld

Um den neuen Anforderungen an das Internet zu begegnen, welche durch die Realisierung als Multi-Service-Plattform sowie durch das Wachstum und die Dynamik des Datenverkehrs entstanden sind, wurden in den letzten Jahren eine Vielzahl an Technologien sowie wissenschaftlichen Prinzipien entwickelt. Diese werden unter dem Begriff Internet Traffic Engineering zusammengefasst. Dieser Abschnitt gibt einen Überblick über den Gesamtkontext des Internet Traffic Engineering und beschreibt das Zusammenwirken der einzelnen Prinzipien sowie Technologien aus der Sichtweise der lastabhängigen Verkehrslenkung.

Wie in den Request For Comments RFCs [ACE02] und [AMA99] definiert, umfasst der Begriff Internet Traffic Engineering die Analyse und Kontrolle des IP-Verkehrs. Das Ziel des Internet Traffic Engineering ist die möglichst effiziente Ausnutzung der Netz-Ressourcen unter Einhaltung der Anforderungen an die Übertragungsqualität einzelner Verkehrsflüsse. Abbildung 1 gibt einen Überblick über die Teilaspekte des Internet Traffic Engineering.

Der erste Hauptaspekt des Internet Traffic Engineering, die Analyse des IP-Verkehrs, setzt sich aus der Messung, der Charakterisierung und der Modellierung des IP-Verkehrs zusammen. Der Verkehrsmessung kommen dabei zwei verschiedene Aufgaben zu. Einerseits dienen die Ergebnisse der Verkehrsmessung als Basis, um das Verhalten des IP-Verkehrs zu charakterisieren und hieraus entsprechende Verkehrsmodelle für den IP-Verkehr abzuleiten. Diese Informationen werden wiederum bei der Untersuchung und Entwicklung der Mechanismen zur Kontrolle des IP-Verkehrs benötigt. Andererseits ist die Verkehrsmessung für den operativen Netzbetrieb relevant. Aus der Messung des aktuellen Verkehrs im Netz leiten sich die Aktionen der Kontroll-Mechanismen während des Netzbetriebs ab.

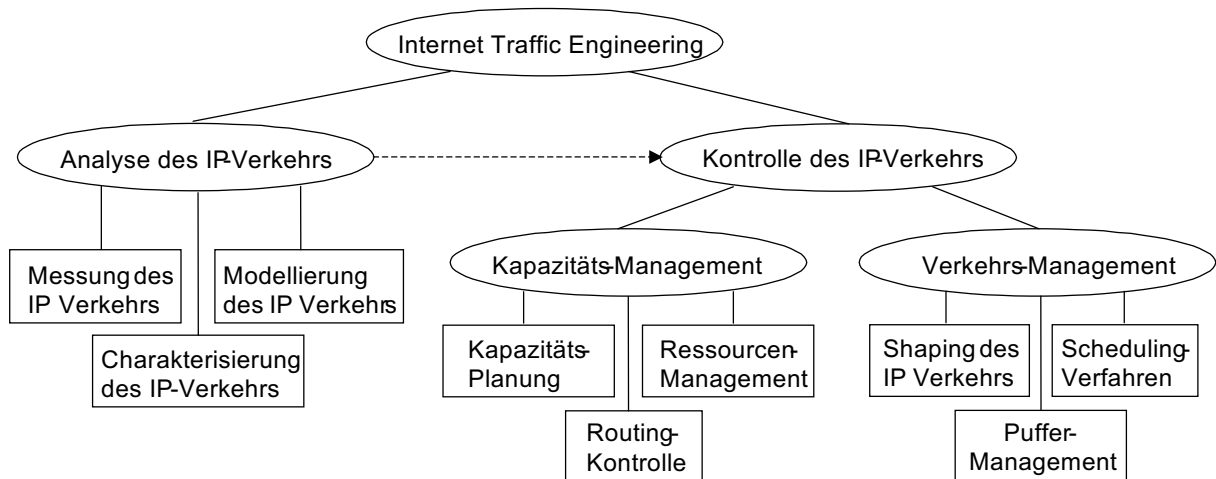


Abbildung 1: Überblick über Internet Traffic Engineering

Der zweite Hauptaspekt des Internet Traffic Engineering, die Kontrolle des IP-Verkehrs, umfasst Methoden des Kapazitäts- und des Verkehrs-Managements. Zum Kapazitäts-Management gehört die Kapazitäts-Planung von Netzen, die Kontrolle der Verkehrslenkung und die Verwaltung der Netz-Ressourcen.

Die Kapazitäts-Planung eines Netzes beschäftigt sich mit der Platzierung und Dimensionierung von Netzknoten und Links (siehe auch [E02]). Dieser Aspekt des Internet Traffic Engineering erfolgt bei dem initialen Aufbau eines neuen Netzes sowie bei der Erweiterung eines bestehenden Netzes. Mit der Erweiterung der Kapazität eines bestehenden Netzes reagiert ein Netzbetreiber auf langfristige Änderungen des Verkehrs. Die Zeitabstände zwischen den Anpassungen der Übertragungskapazität liegen typischerweise in der Größenordnung von Wochen bis Monaten.

Die Verkehrslenkungs-Kontrolle reagiert auf Verkehrsschwankungen, welche in Zeitabständen von Sekunden bis Wochen auftreten. Hierzu passt diese die Pfade, über welche die einzelnen Verkehrsströme durch das Netz geleitet werden, dynamisch an die aktuelle Auslastungssituation im Netz an. Auf die verschiedenen Realisierungsmöglichkeiten der Verkehrslenkungs-Kontrolle wird in dieser Arbeit noch näher eingegangen.

Das Ressourcen-Management als weiterer Aspekt des Kapazitäts-Managements realisiert die Steuerung der Annahme einzelner Verbindungen. Hierzu überprüft das Ressourcen-Management mit jeder Verbindungsanfrage, ob ausreichend Netz-Ressourcen zum Aufbau der Verbindung zur Verfügung stehen. Zu den Netz-Ressourcen zählen die Übertragungskapazität der Links sowie die Pufferkapazität und die Prozessorleistung der Netzknoten. Das Ressourcen-Management wird für Verkehrsströme mit Anforderungen an die Dienstgüte (Quality of Service QoS) benötigt. Für den klassischen Verkehr im Internet, den Best Effort Verkehr, ist kein Ressourcen-Management vorgesehen. Dieser wird ohne Prüfung der Netzauslastung über das Netz übertragen.

Während die Funktionen des Kapazitäts-Managements unter Berücksichtigung des gesamten Netzes ausgeführt werden, bearbeiten die Funktionen des Verkehrs-Managements autonom innerhalb eines Netzknotens den ankommenden Verkehr. Zu den Funktionen zählen das Shaping des IP-Verkehrs, das Puffer-Management und die Anwendung von Scheduling-Verfahren.

Die Shaping Funktion passt einen Verkehrsstrom an definierte Verkehrsparameter wie

z.B. eine maximale Senderate, eine maximale Burstlänge und mittlere Datenrate an. Das Puffer-Management teilt den aktiven Verkehrsströmen die Pufferkapazität an den Ausgängen dynamisch zu. Die Scheduling-Verfahren legen die Sendereihenfolge der Pakete in einem Netzknoten fest.

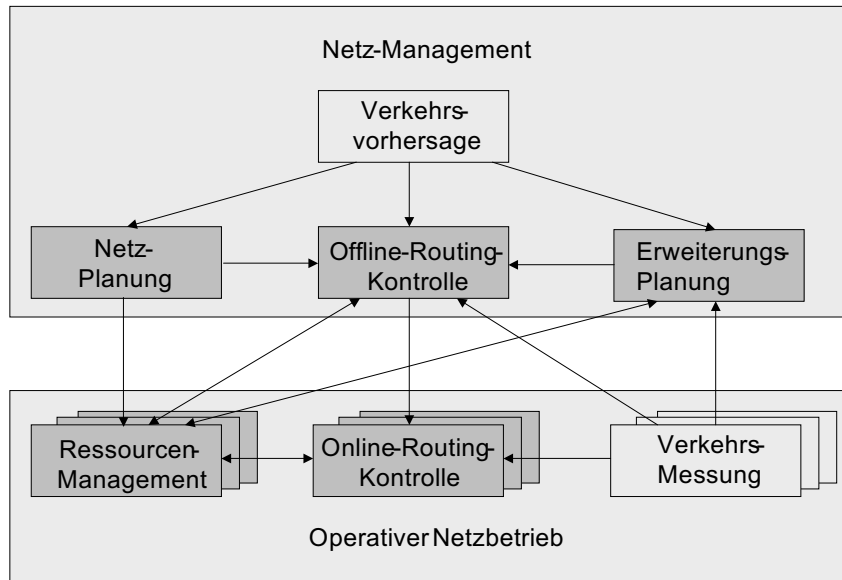


Abbildung 2: Zusammenspiel der TE Mechanismen des Kapazitäts-Managements

In der Folge wird das mögliche Zusammenwirken der Traffic Engineering TE Mechanismen bei dem Betrieb eines Netzes beschrieben (siehe Abbildung 2). Im Zentrum dieser Betrachtung, welche sich auf die Mechanismen des Kapazitäts-Managements beschränkt, steht dabei die Funktionsweise der dynamischen Routing-Kontrolle.

Zu Beginn des operativen Netzbetriebs steht die initiale Planung sowie der Aufbau des Netzes. Zudem erfolgt die Festlegung des initialen Routings. Während des operativen Netzbetriebs führt das Ressourcen-Management eine Verbindungsannahme für den QoS Verkehr durch. Das Ressourcen-Management kann zentral durch eine oder verteilt durch mehrere Instanzen durchgeführt werden. Auf Anfrage übergibt das Ressourcen-Management die aktuellen Blockierungswahrscheinlichkeiten an die Instanzen der Online und Offline Routing-Kontrolle.

In gleicher Weise liefern verteilte Mess-Instanzen Informationen über die aktuelle Netzauslastung an die Instanzen der Routing-Kontrolle. Zur Beschreibung der aktuellen Netzauslastung können die Mess-Instanzen verschiedene Parameter bestimmen. Hierzu zählt die mittlere Auslastung, die mittlere Paketverzögerung, der Paketverzögerungs-Jitter und die Paketverlustwahrscheinlichkeit. Diese Parameter können dabei auf einen Link oder einen Pfad bezogen sein. Basierend auf den Informationen der Instanzen des Ressourcen-Managements und der Verkehrsmessung entscheiden die Instanzen der Routing-Kontrolle über die Anpassung der Lastverteilung im Netz.

Die Instanz der Offline Routing-Kontrolle, welche Teil des Netz-Managements ist, erhält diese Informationen in Zeitabständen von Stunden bzw. Wochen. Basierend auf diesen Informationen und einer Verkehrsvorhersage berechnet die Offline Routing-Kontrolle die optimale Lastverteilung für die nächste Zeitperiode. Die Berechnung der neuen Lastverteilung ist dabei entkoppelt von dem Betrieb des Netzes. In Situationen schwacher

Netzauslastung (z.B. in der Nacht) erfolgt die Umstellung der Verkehrslenkung auf die neue Lastverteilung.

Die Online Routing-Kontrolle hingegen, welche durch eine zentrale oder mehrere verteilte Instanzen realisiert sein kann, fragt die Auslastungs-Informationen in kurzen Zeitabständen von einigen Sekunden bis wenigen Minuten ab. Erkennt die Online Routing-Kontrolle eine Schiefelast bzw. eine Überlastung des Netzes, führt diese unmittelbar während des Netzbetriebs eine Anpassung der Lastverteilung durch. Hierzu bestimmt die Online Routing-Kontrolle in der Regel nur eine lokale Optimierung der Lastverteilung.

Bei einer zu großen Gesamtlast im Netz kann auch durch die Verwendung der Routing-Kontrolle keine ausreichende Netzperformanz erzielt werden. In diesem Fall muss die Erweiterungsplanung zur Anpassung der Übertragungskapazität des Netzes angestoßen werden. Die Erweiterungsplanung basiert wie die Offline Routing-Kontrolle auf den Informationen des Ressourcen-Managements, der Verkehrsmessung und der Vorhersage über zukünftig erwarteten Verkehr.

Um die Umsetzung des Kapazitäts-Managements in IP-Netzen zu realisieren, wurden in den letzten Jahren einige neue Architekturen und Protokolle definiert. Der Hauptfokus lag hierbei einerseits darauf, aus dem Internet eine Multi-Service-fähige Plattform zur Übertragung von Verkehr mit unterschiedlichen Anforderungen an die Übertragungsqualität zu machen. Andererseits galten die Bestrebungen der Umsetzung der dynamischen lastabhängigen Routing-Kontrolle in den IP-Netzen.

Zur Übertragung von Verkehr mit unterschiedlichen Anforderungen an die Übertragungsqualität in IP-Netzen wurden die beiden Architekturen Integrated Services IntServ und Differentiated Services DiffServ entwickelt. IntServ [W97] basiert auf IP-Netzen und führt eine Verbindungsannahme und Kapazitäts-Reservierung auf der Ebene einzelner Datenflüsse durch. Zur Signalisierung wird das Resource Reservation Protocol RSVP [BZB97] verwendet. Während der Vorteil von IntServ die Garantie der Übertragungsqualität für die angenommenen Verbindungen ist, stellt diese Architektur auf Grund der zu großen Komplexität keine skalierbare Lösung zum Einsatz in Netzen beliebiger Größe dar.

Das DiffServ-Konzept [BBC98] lässt sich sowohl in IP-Netzen als auch in Multiprotocol Label Switching (MPLS) Netzen realisieren [FDD02]. DiffServ unterteilt den Verkehr in wenige Verkehrsklassen, welche an den Knoten im Netz mit unterschiedlichen Prioritäten behandelt werden. Die Anzahl der Verkehrsklassen sowie die Durchführung der Priorisierung wird dabei durch DiffServ nicht festgelegt. DiffServ stellt auf der einen Seite eine einfache skalierbare Architektur dar, welche sich in Netzen beliebiger Größe realisieren lässt. Auf der anderen Seite kann mit DiffServ im Gegensatz zur IntServ-Architektur keine Garantie für die Übertragungsqualität einzelner Verbindungen gegeben werden.

Basierend auf der IntServ und der DiffServ-Architektur sind weitere Architekturen zur Realisierung des Internets als Multi-Service-Plattform definiert worden. Eine interessante Architektur schlagen Glasmann et al. [GME02] [GM02] vor, in welcher ein zentraler Ressourcen-Manager eine Verbindungsannahme-Steuerung innerhalb einer DiffServ Netz-Domäne realisiert. Damit lassen sich innerhalb der Domäne die Übertragungsqualität einzelner Verbindungen garantieren und somit die Vorteile der DiffServ und der IntServ Architektur verbinden. Durch die Aufteilung eines Netzes in mehrere Netz-Domänen sowie durch die Definition der Kommunikation zwischen verschiedenen Ressourcen-Managern wird die Skalierbarkeit der Architektur erreicht.

Zur Umsetzung der dynamischen lastabhängigen Routing-Kontrolle in IP-Netzen werden zwei Realisierungsmöglichkeiten unterschieden. Die eine basiert auf einer modifizierten Anwendung der bekannten IP-Routing-Protokolle und die andere auf der Verwendung der MPLS-Technologie.

Die IP-Routing-Protokolle, welche die Verkehrslenkung und die daraus resultierende Lastverteilung im Netz festlegen, beruhen auf der Verwendung link-bezogener Metriken. Diese Link-Metriken sind in der bisherigen Realisierung statisch und bleiben unverändert. Die modifizierte Anwendung der IP-Routing-Protokolle erlaubt nun eine dynamische Festlegung der Link-Metriken, um die Verkehrslenkung an neu entstandene Lastsituationen im Netz anzupassen.

Bei Verwendung der verbindungsorientierten Technologie MPLS wird der IP-Verkehr über MPLS-Verbindungen durch das Netz übertragen. Dies erlaubt bei entsprechender Festlegung der Pfade der MPLS-Verbindungen eine an die Auslastungssituation angepasste Verteilung der Last über das Netz. Sowohl die IP-Routing-Protokolle als auch die MPLS-Technologie werden ausführlich in Abschnitt 2.1 vorgestellt.

1.2 Beitrag der Arbeit

Im Rahmen der Arbeit wurde ein neues verteiltes reaktives Verkehrslenkungsverfahren für MPLS-Weitverkehrsnetze entwickelt. Das Verfahren kann sowohl nach einem Mehrpfadansatz als auch nach einem Pfadumlegungsansatz betrieben werden. Das Verfahren dient zur Optimierung der Lastverteilung des klassischen Best Effort IP-Verkehrs. Die Hauptmerkmale des Verfahrens sind die geringe Konvergenzdauer nach Erkennung von Überlastsituationen sowie die flexible Einsatzmöglichkeit in beliebigen Netzszenarien.

Das Verfahren wurde simulativ anhand realistischer Netzszenarien evaluiert, welche die Größe und Topologie existierender IP-Netze widerspiegeln. Das stellt eine Neuerung gegenüber den bisherigen Untersuchungen reaktiver Verkehrslenkungsverfahren dar, welche sich auf kleine Testnetze beschränkt haben.

Zur Untersuchung des reaktiven Verkehrslenkungsverfahrens wurde in dieser Arbeit eine Methode entwickelt, welche die Simulation der Verfahren in realistischen Netzszenarien erlaubt. Diese Methode beruht auf der Modellierung des Verkehrs auf Ratenebene.

1.3 Aufbau der Arbeit

Die weitere Arbeit gliedert sich in fünf Kapitel. Das folgende Kapitel 2 beschreibt zunächst die Grundlagen lastabhängiger Verkehrslenkungsverfahren und gibt einen Überblick über den aktuellen Stand der Technik.

Kapitel 3 beschreibt das entwickelte reaktive Verkehrslenkungsverfahren für MPLS-Weitverkehrsnetze und führt eine qualitative Bewertung des Verfahrens durch.

Im Anschluss daran setzt sich Kapitel 4 mit der Simulations-Methode zur Performanz-Analyse reaktiver Verkehrslenkungsverfahren in großen Netzen auseinander. Neben einer Analyse des Stands der Technik zur Simulation von Paketnetzen präsentiert das Kapitel eine neu entwickelte skalierbare Simulations-Methode. Diese basiert auf der ratenbasierten Modellierung des Datenverkehrs.

Kapitel 5 führt unter Verwendung der neuen ratenbasierten Simulations-Methode eine quantitative Analyse des entwickelten reaktiven verteilten Verkehrslenkungsverfahrens anhand ausgewählter Netz- und Auslastungsszenarien durch.

Die wesentlichen Ergebnisse dieser Arbeit fasst schließlich das Kapitel 6 zusammen.

2 Grundlagen und Stand der Technik

Dieses Kapitel beschäftigt sich mit den Grundlagen und mit dem Stand der Technik lastabhängiger Verkehrslenkung in IP- und MPLS-Netzen.

Der erste Teil des Kapitels erläutert die Grundlagen und geht hierzu in Abschnitt 2.1 zunächst auf die Konzepte und protokoll-spezifischen Aspekte der IP-Verkehrslenkung und der MPLS-Technologie ein. Abschnitt 2.2 erläutert im Anschluss daran die verschiedenen Varianten zur Realisierung lastabhängiger Verkehrslenkung in IP-Netzen. Der Grundlagenteil wird durch die Definition der optimalen Lastverteilung in Abschnitt 2.3 abgeschlossen.

Der zweite Teil des Kapitels beschreibt den Stand der Technik lastabhängiger Verkehrslenkung. Abschnitt 2.4 setzt sich mit den proaktiven Verkehrslenkungsverfahren und Abschnitt 2.5 mit den reaktiven Verkehrslenkungsverfahren auseinander. Beide Abschnitte definieren jeweils die Anforderungen an die Verfahren, nennen grundlegende Berechnungsmethoden und Algorithmen und geben einen Überblick über bereits realisierte Verfahren. Der letzte Abschnitt 2.6 fasst den Stand der Technik zusammen und zeigt auf, weshalb die Entwicklung eines neuen reaktiven Verkehrslenkungsverfahrens notwendig ist.

2.1 Grundlegende Konzepte und Protokolle

Dieser Abschnitt geht auf die Konzepte und Protokolle lastabhängiger Verkehrslenkungsverfahren ein. Hierzu präsentiert Abschnitt 2.1.1 die Grundlagen der Verkehrslenkung im Internet und nennt die bekannten IP-Routing-Protokolle. Abschnitt 2.1.2 führt in die MPLS-Technologie ein.

2.1.1 Verkehrslenkung im Internet

Die Verkehrslenkung ist nach der Definition des ISO-OSI-Referenzmodells eine Aufgabe der Vermittlungsschicht (siehe auch [S99b]). Die Vermittlungsschicht legt den Pfad durch das Netz bzw. die Netze fest, auf welchem die Daten zwischen zwei Endsystemen übertragen werden. Das Internet Protocol IP als zentrales Protokoll der Vermittlungsschicht stellt einen verbindungslosen paketorientierten Übertragungsdienst zur Verfügung, bei welchem die Vermittlung von Datenpaketen auf weltweit eindeutigen IP-Adressen beruht. Diese Adresse trägt jedes IP-Paket in den Kontroll-Informationen mit sich. Die Berechnung der Pfade auf der Basis der IP-Adressen wird durch die IP-Routing-Protokolle realisiert, auf die in der Folge noch näher eingegangen wird.

Abbildung 3 gibt einen Überblick über die Struktur des Internets aus der Sicht der Verkehrslenkung. Das Internet setzt sich aus autonomen Systemen zusammen. Ein autonomes System bezeichnet ein IP-Netz innerhalb des Internets, welches von einem Netzbetreiber verwaltet wird.

Ein IP-Netz besteht aus IP-Routern, welche über Links mit vorgegebener Übertragungskapazität und Länge verbunden sind. Transport-Technologien wie z.B. SDH, ATM oder WDM realisieren die Datenübertragung auf den Links.

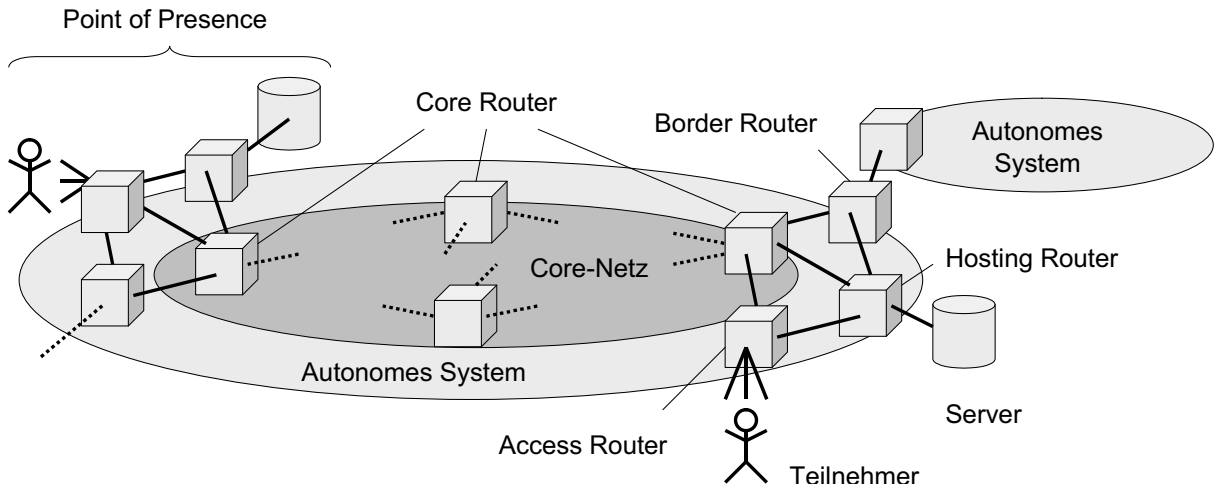


Abbildung 3: Struktur des Internets aus der Sicht der Verkehrslenkung

Die IP-Router in einem IP-Netz werden in Points of Presence PoPs gruppiert. Ein PoP fasst IP-Router mit unterschiedlichen Funktionen zusammen. Innerhalb eines PoPs existieren ein oder mehrere Core Router, die die Points of Presence eines IP-Netzes miteinander verbinden. Die Vernetzung der Core Router wird auch als das Core-Netz bezeichnet. Neben den Core Routern gibt es pro Point of Presence eine Kombination aus drei Klassen von Routern. Über diese Router tritt der Verkehr in das Netz ein und wieder aus. Hierzu zählen erstens die Access Router, welche den Netzzugang für die Teilnehmer bereitstellen. Zweitens gehören die Hosting Router dazu, welche die Anbindung der Server realisieren. Die dritte Klasse sind die Border Router, welche das IP-Netz mit den IP-Netzen anderer Netzbetreiber verbinden.

Die Haupt-Funktionen eines IP-Routers sind das IP-Forwarding und das IP-Routing. IP-Forwarding bezeichnet das Weiterleiten eines ankommenden IP-Pakets zu dem nächsten IP-Router, welcher auf dem Weg zum Ziel-Endsystem liegt. IP-Routing umschreibt die Berechnung und die Festlegung der Pfade, auf welchen die IP-Pakete mit der jeweiligen IP-Adresse weitergeleitet werden.

Den prinzipiellen Aufbau eines IP-Routers zeigt Abbildung 4. Nach Empfang eines IP-Pakets durch eine LineCard trifft eine Forwarding Einheit die Entscheidung über das Weiterleiten eines Pakets. Hierzu wertet der IP-Router die IP-Adresse aus und vergleicht diese mit den Einträgen in der Routing-Tabelle. Ein Eintrag in der Routing-Tabelle ordnet einer IP-Adresse einen Ausgangsport zu, über welchen der IP-Router das IP-Paket mit entsprechender IP-Adresse aussenden soll. Für alle nicht aufgeführten IP-Adressen wird ein Default-Ausgangsport definiert. Nach Festlegung des Ausgangsports übergibt die IP-Forwarding-Einheit das IP-Paket an die Switching-Matrix, welche das IP-Paket zu dem entsprechenden Ausgang weiterleitet. Die LineCard sendet das IP-Paket schließlich zum nächsten IP-Router.

Die Routing-Einheit legt die Einträge in der Routing-Tabelle fest. Hierzu bestimmt die Routing-Einheit für die jeweiligen IP-Adressen die kürzesten Pfade durch das Netz. Aus den kürzesten Pfaden leitet der IP-Router den nächsten IP-Router und den Ausgangsport für die jeweilige IP-Adresse ab. Nach jeder Neubestimmung der Einträge in die Routing-Tabelle übergibt die Routing-Einheit die neuen Informationen an die Routing-Tabellen der Forwarding-Einheiten. Zur Berechnung der kürzesten Pfade durch das Netz werden

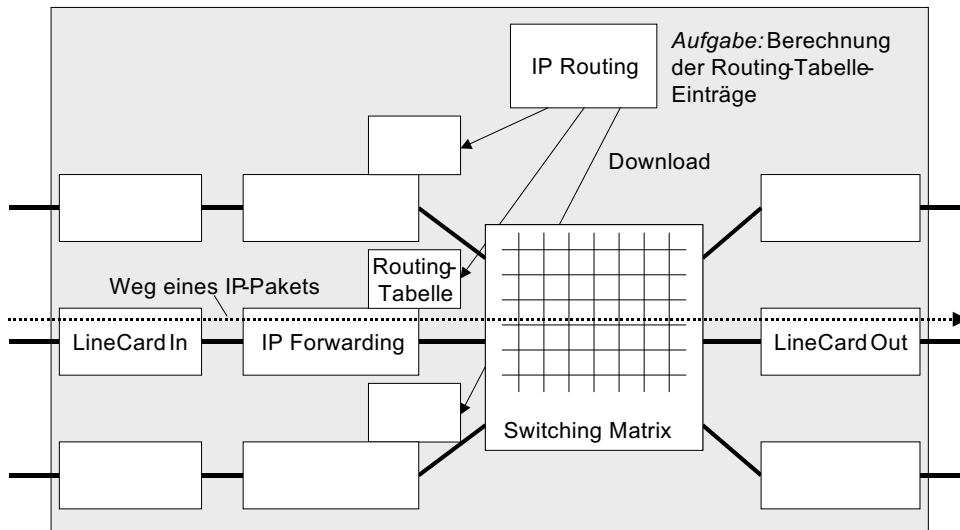


Abbildung 4: Aufbau eines IP-Routers

Routing-Protokolle verwendet, auf die in den beiden folgenden Abschnitten eingegangen wird.

2.1.1.1 Klassifizierung der IP-Routing-Protokolle Dieser Abschnitt beschreibt die grundlegende Funktionsweise der IP-Routing-Protokolle und führt die Klassifizierungsmerkmale dieser Protokolle ein.

Die Berechnung der Pfade durch das Netz basiert auf der Zuordnung einer oder mehrerer Metriken zu jedem Link in einem Netz. Der Wert der Metrik kann im einfachsten Fall gleich eins gesetzt werden oder von den charakteristischen Merkmalen des Links wie z.B. der Länge oder Übertragungskapazität abgeleitet werden. Ein Link kann durch eine oder mehrere Metriken charakterisiert werden. Zur Auswahl des günstigsten Pfades zwischen zwei Knoten wird entweder der Pfad mit der geringsten Summe der Link-Metriken oder mit der größten minimalen Link-Metrik festgelegt. Als dritte Möglichkeit zur Festlegung des Pfades besteht die Kombination der beiden Auswahl-Kriterien.

Die Routing-Protokolle, welche auf diesem Prinzip aufbauen, lassen sich mit folgenden Klassifizierungsmerkmalen charakterisieren.

Distance Vector Routing versus Link State Routing: Distance Vector Routing und Link State Routing beschreiben die beiden Prinzipien zur Bestimmung der Pfade zwischen zwei Knoten. Distance Vector Routing Protokolle basieren auf dem Austausch von Erreichbarkeitsinformationen, welche den Pfad zwischen zwei Routern charakterisieren. Erreichbarkeitsinformationen eines Pfades geben an, wie groß die Summe der Link-Metriken bzw. die minimale Link-Metrik des Pfades von dem sendenden Router aus sind. Durch Vergleich der Erreichbarkeitsinformationen, die ein Router für eine Ziel-Adresse von verschiedenen Nachbar-Routern erhält, und unter Berücksichtigung der lokalen Links wählt ein Router den günstigsten Pfad aus. Der Router seinerseits sendet diesen wiederum an seine Nachbar-Router. Den Distance Vector Routing Protokollen liegt der Bellmann-Ford-Algorithmus zu Grunde. Link State Routing-Protokolle basieren auf dem Austausch der Link-Metriken und der Zustände der Netzkomponenten. Aus diesen Informationen baut ein Router die Topologie-Datenbank auf, welche die Netzsicht eines Routers widerspiegelt. Auf dieser Basis berechnet ein Router die günstigsten Pfade durch das Netz. Zur

Berechnung der günstigsten Pfade wird der Dijkstra-Algorithmus verwendet.

Internes Routing versus externes Routing: Internes Routing und externes Routing unterscheiden das Einsatzgebiet des Routings. Das Routing innerhalb eines autonomen Systems wird als internes Routing und das Routing zwischen den autonomen Systemen als externes Routing bezeichnet.

Statisches Routing versus dynamisches Routing: Statisches und dynamisches Routing beziehen sich auf die Auslösung einer Neuberechnung der Routing-Tabellen-Einträge. Statisches Routing legt zu einem gegebenen Zeitpunkt die Routing-Tabellen-Einträge für die jeweiligen IP-Adressen fest und verändert diese unabhängig von möglichen Ereignissen im Netz nicht. Nur ein manueller Eingriff löst eine Neubestimmung der Routing-Tabellen-Einträge aus. Dynamisches Routing hingegen berechnet die Einträge in den Routing-Tabellen neu, wenn Änderungen der Topologie wie z.B. bei Ausfällen von Netzkomponenten erkannt werden. Dynamisches Routing reagiert jedoch nicht auf Änderungen der Lastverteilung.

	Internes Routing	Externes Routing
Distance Vector Routing	RIP, IGRP, EIGRP	BGP
Link State Routing	OSPF, ISIS	

Tabelle 1: Klassifizierung der Routing-Protokolle

Tabelle 1 klassifiziert die bekannten IP-Routing-Protokolle nach den zuvor beschriebenen Merkmalen. Alle genannten Routing-Protokolle sind dynamisch und werden in dem folgenden Abschnitt kurz erläutert.

2.1.1.2 Bekannte IP-Routing-Protokolle Dieser Abschnitt liefert eine Beschreibung der bekannten IP-Routing-Protokolle, welche in Tabelle 1 aufgeführt sind.

Routing Information Protocol RIP: RIP [H88] [M93] ist das älteste Routing Protokoll und wurde 1988 von der IETF standardisiert. Die RIP-Instanzen in den Routern tauschen in periodischen Zeitabständen (typischerweise 30 Sekunden) sämtliche Erreichbarkeitsinformationen untereinander aus. Hierzu verwendet RIP UDP-Verbindungen. RIP-Instanzen wählen die Pfade mit dem geringsten Hop Count aus.

Erhält ein Router von seinem Nachbar-Router über mehrere Perioden keine Erreichbarkeitsinformationen, dann erkennt er den Link zu dem Nachbar-Router als ausgefallen. Nach Erkennung eines Link-Ausfalls erklärt ein Router alle Ziel-Router, die zuvor über diesen Link erreichbar waren, als unerreichbar. Diese Nicht-Erreichbarkeitsinformation verteilt der Router an alle anderen Router.

Der große Vorteil des RIP ist der einfache Aufbau und die daraus resultierende Robustheit, was der Grund für die ehemals weite Verbreitung des Protokolls ist. Dem gegenüber stehen zwei große Nachteile von RIP. Erstens braucht RIP eine sehr lange Zeit (bis zu mehreren Minuten), um nach einem Ausfall wieder einen stabilen Zustand zu erreichen. Während der Übergangsphase können Routing-Schleifen entstehen. Der zweite Nachteil ist, dass der maximale Hop Count auf 15 begrenzt ist und das RIP somit nur in Netzen mittlerer Größe verwendet werden kann.

Interior Gateway Routing Protocol IGRP: Das Interior Gateway Routing Protokoll IGRP [H91] ist ein proprietäres Routing Protokoll, welches von dem Unternehmen CISCO

entwickelt wurde. Die grundsätzliche Funktionsweise entspricht der von RIP mit den in der Folge beschriebenen Erweiterungen.

Im Gegensatz zum RIP verwendet das IGRP vier Metriken pro Link. Eine Metrik steht für die Übertragungsverzögerung, eine weitere für die Übertragungskapazität, die dritte für die aktuelle Auslastung und die vierte für die Zuverlässigkeit des Links. Da durch die Berücksichtigung der aktuellen Auslastung und der Zuverlässigkeit des Links instabile Zustände im Netz entstehen können, werden in der Regel nur die ersten beiden Metriken verwendet. Die Verwendung mehrerer Metriken pro Link erhöht die Flexibilität des Routings.

Eine weitere Neuerung des IGRP ist die Definition weiterer Maßnahmen zur Stabilisierung des Routings. Zudem erlaubt das IGRP, den Verkehr zu einem Ziel-Router über mehrere Pfade zu übertragen.

Enhanced Interior Gateway Routing Protocol EIGRP: Das Enhanced Interior Gateway Routing Protocol EIGRP [AGB94] ist eine Weiterentwicklung des IGRP. Die wesentliche Neuerung ist die Verwendung des Diffusing Update Algorithmus DUAL, welcher ein stabiles Routing sowie eine Schleifenfreiheit nach Topologie-Änderungen garantiert. Nach [AGB94] erreicht EIGRP annähernd die Konvergenzzeiten der Link State Routing Protokolle bei deutlich geringerem Prozessoraufwand.

Open Shortest Path First OSPF und Intermediate System to Intermediate System ISIS: In den heutigen IP-Netzen finden hauptsächlich die beiden Link State Routing Protokolle Open Shortest Path First OSPF und Intermediate System to Intermediate System ISIS Anwendung. Beide Protokolle haben den gleichen Aufbau. Die Haupt-Komponenten beider Protokolle werden in der Folge kurz erläutert.

Mit dem Hello Protokoll führen die Router eine Nachbarschaftserkennung und fortlaufende Überprüfung der Verbindung zum Nachbar-Router durch. Jeder Router sendet dafür in periodischen Zeitabständen (der Standard-Wert ist auf 10 Sekunden eingestellt) Hello Nachrichten an alle Nachbar-Router. Bleiben mehrere Hello-Nachrichten von einem Nachbar-Router aus (hier ist der Standard-Wert 4 Sekunden), erkennt ein Router den Link zum Nachbar-Router als gestört und verteilt die Information über die Topologie-Änderung an die übrigen Router.

Mittels des Flutungs-Mechanismus sendet ein Router den Zustand und die Metriken der von ihm abgehenden Links an seine Nachbar-Router. Erhält einer Router eine Topologie-Information, überprüft dieser zunächst anhand einer Sequenznummer und der zugehörigen Router-ID, ob er die Nachricht bereits erhalten hat. Ist das der Fall, verwirft ein Router die Nachricht. Ansonsten sendet der Router die Topologie-Informationen weiter an seine Nachbar-Router mit Ausnahme des Routers, von dem der Router die Nachricht erhalten hat. Über den Flutungs-Mechanismus verteilen die Router die Zustandsinformation ihrer Links über das gesamte Netz.

Aus den erhaltenen Topologie-Informationen generiert ein Router seine aktuelle Netzsicht. Diese speichert ein Router in einer Topologie-Datenbank. Basierend auf der Topologie-Datenbank berechnet jeder Router mittels des Dijkstra-Algorithmus die kürzesten Pfade zu den Ziel-Adressen und trägt diese in die Routing-Tabelle ein. Hierfür verwenden OSPF und ISIS eine additive Metrik. Gibt es zu einem Zielknoten zwei oder mehrere kürzeste Pfade mit der gleichen Länge, dann erlaubt OSPF bei Einsatz des Equal Cost Multipath ECMP die Verwendung sämtlicher Pfade. ECMP verteilt die Datenpakete gleichmäßig

über die Pfade, ohne dabei die Zugehörigkeit der Datenpakete zu einzelnen IP-Flüssen zu berücksichtigen.

Border Gateway Protocol BGP: Das Border Gateway Protocol BGP [RWL95] wird im Gegensatz zu den anderen genannten Routing-Protokollen für das Routing zwischen den autonomen Systemen eingesetzt. BGP ist ein Pfadvektor Protokoll, was eine Erweiterung des Distance Vector Prinzips darstellt. Zusätzlich zu der Distanz zu den jeweiligen Zielen tragen die Erreichbarkeitsinformationen die Kennungen der durchlaufenen autonomen Systeme mit sich. Durch Auswertung der Pfadinformation kann ein BGP-Router Routing-Schleifen erkennen.

2.1.2 Multiprotocol Label Switching MPLS

Neben der zuvor beschriebenen derzeit verwendeten Verkehrslenkung im Internet stellt das Multiprotocol Label Switching MPLS eine Basis-Technologie zur Realisierung lastabhängiger Verkehrslenkung dar. Dieser Abschnitt erläutert die historische Entwicklung von MPLS und beschreibt die Grundlagen der Technologie. Hierzu gibt Abschnitt 2.1.2.1 zunächst einen Überblick über IP over ATM Netzkonzepte, welche die Vorläufer von MPLS sind. In der Folge beschreibt der Abschnitt 2.1.2.2 die grundlegenden Konzepte der MPLS Technologie und Abschnitt 2.1.2.3 stellt die Verbindungssteuerungs-Protokolle, die Label Distribution Protokolle, vor.

2.1.2.1 Überblick über IP over ATM Netzkonzepte Seit Beginn der neunziger Jahre sind viele Konzepte zur Kopplung von IP- und ATM-Netzen entstanden. Der ursprüngliche Grund der Entwicklung war, die große Übertragungskapazität der ATM-Netze für die Übertragung von IP-Verkehr zu nutzen. Während IP-Router traditionell software-basiert waren und die Übertragungskapazität der IP-Router gering war, verwendeten ATM-Switches bereits schnelle hardware-basierte Komponenten. Mit der Entwicklung hardware-basierter Hochgeschwindigkeits-IP-Router ist dieser Vorteil der ATM-Netze jedoch in den Hintergrund getreten.

Tabelle 2 gibt einen Überblick über die IP over ATM Netzkonzepte, welche in den letzten Jahren vom ATM-Forum und von der Internet Engineering Task Force IETF standardisiert worden sind.

IP over ATM Netzkonzepte	Gremium	Zeitpunkt
Classical IP over ATM CLIP	IETF	1994
LAN Emulation LANE	ATM-Forum	1995
IP Switching	IETF	1996
Multiprotocol over ATM MPOA	ATM-Forum	1997
Cell Switched Router CSR	IETF	1997
Tag Switching	IETF	1997
Multiprotocol Label Switching MPLS	IETF	2001

Tabelle 2: Übersicht über alle standardisierten IP over ATM Netzkonzepte

Als erstes Konzept wurde 1994 das Classical IP over ATM CLIP von der IETF standardisiert [L94]. CLIP teilt ein IP-Netz, welches auf einem ATM-Netz aufsetzt, in Subnetze auf und platziert pro Subnetz einen Server zur Adress-Auflösung. IP-Hosts aus

einem Subnetz geben dem Server sowohl die IP- als auch ATM-Adresse bekannt. Zur Vermittlung von IP-Daten zwischen zwei IP-Hosts existieren nach dem CLIP Ansatz zwei Varianten. Nach der ersten Variante vermitteln die Router die IP-Datenpakete auf der IP-Schicht und übertragen die Pakete anschließend auf zwischen den benachbarten Routern eingerichteten ATM-Verbindungen zum Nachbar-Router. Das entspricht der klassischen IP-basierten Vermittlung von Datenpaketen. Nach der zweiten Variante baut der Quell-Host eine direkte ATM-Verbindung zum Ziel-Host auf. Zur Auflösung der Ziel-IP-Adresse in die entsprechende ATM-Adresse stellt der Quell-Host eine Anfrage an den Adress-Auflösungs-Server und triggert anschließend die ATM-Schicht zum Aufbau einer ATM-Verbindung. Der Nachteil dieses Ansatzes ist, dass ATM-Verbindungen nur innerhalb eines IP-Subnetzes geschaltet werden können. Selbst wenn zwei IP-Subnetze auf demselben ATM-Netz aufsetzen, müssen die Datenpakete an den Grenzen der IP-Subnetze auf IP-Schicht vermittelt werden.

Die LAN Emulation LANE [A95] ist ein Vorschlag von dem ATM-Forum, welcher 1995 standardisiert wurde. Die Idee von LANE ist, räumlich voneinander getrennte LANs und einzelne ATM-Hosts über ein öffentliches ATM-Netz zu einem emulierten LAN zu verbinden. Bei diesem Ansatz setzen die Schicht2-Protokolle des Ethernet 802.3 und des Token Rings 802.5 auf der ATM-Schicht auf. LANE definiert Server, welche die Konfiguration beim Aufbau eines emulierten LANs steuern und die Adress-Auflösung sowie das Senden von Broadcast-Nachrichten zwischen einzelnen LANs und ATM-Hosts realisieren. Möchte ein LAN-Client Daten an einen anderen LAN-Client übertragen, stellt dieser zunächst die Anfrage zur Adress-Auflösung an den entsprechenden Server. Die Anfrage wird über vordefinierte ATM-Kontroll-Verbindungen übertragen. Anschließend wird der Aufbau einer ATM-Verbindung zur Übertragung der Daten angestoßen. Der Nachteil des Ansatzes ist die hohe Komplexität und die daraus resultierende geringe Skalierbarkeit.

Multiprotocol over ATM [A97], welches 1997 von dem ATM-Forum vorgeschlagen wurde, definiert ein Konzept zur Übertragung des Verkehrs beliebiger Protokolle der Vermittlungsschicht über ATM. Praktisch gesehen ist es auf Grund mangelnder Verbreitung anderer Protokolle ein Konzept zur Übertragung von IP-Verkehr über ATM-Netze. Ähnlich dem CLIP erlaubt MPOA die Übertragung von IP-Verkehr sowohl auf IP-Ebene als auch auf geschalteten ATM-Verbindungen, den sogenannten ATM-Shortcuts. Um die ATM-Adresse des Ziel-Hosts aufzulösen und eine ATM-Verbindung aufbauen zu können, verwendet MPOA das Next Hop Resolution Protocol NHRP [LKP98]. Ein Quell-IP-Host sendet hierzu eine Anfrage entlang des kürzesten Pfades, welcher durch die IP-Routing-Protokolle vorgegeben wird. Erreicht die Anfrage den Ziel-IP-Host, löst dieser die Adresse auf und die Antwort wird entlang des kürzesten Pfades wieder zum Quell-IP-Host gesendet. MPOA baut ATM-Shortcuts für einzelne IP-Flüsse auf, welche durch eine Flusserkennung automatisch erkannt werden. Die Realisierung der Flusserkennung ist nicht genauer spezifiziert.

Der Vorteil von MPOA gegenüber CLIP ist die Tatsache, dass Verbindungen über das gesamte ATM-Netz geschaltet werden können und keine IP-Subnetz-Grenzen existieren. Wie CLIP hat aber auch MPOA das Problem einer geringen Skalierbarkeit, da die Adress-Auflösung und die Schaltung von Verbindungen für einzelne IP-Flüsse erfolgt.

Der von Toshiba entwickelte und 1997 von der IETF standardisierte Ansatz des Cell Switched Routers CSR [KNE97] basiert auf einem mit MPOA vergleichbaren Ansatz. Der Ansatz ist nach dem zentralen Element, dem Cell Switched Router, benannt. Der Vorteil von CSR gegenüber MPOA ist die Tatsache, dass die Verwendung der Cell Switched Router

auch das Schalten von ATM-Verbindungen über ATM-Netz-Grenzen hinaus möglich ist. Der Nachteil ist wie bei den zuvor genannten Konzepten die geringe Skalierbarkeit.

Ein völlig anderes Konzept zur Kopplung von IP- und ATM-Netzen verfolgen die übrigen Ansätze. Während die zuvor genannten Konzepte auf der Nutzung des vollen Umfangs der ATM-Technologie basieren und die Hauptfunktionalität der Verfahren die Adress-Auflösung ist, verwenden die folgenden Ansätze lediglich die ATM-Hardware. Zur Steuerung der Datenübertragung definieren diese Ansätze eigene Protokolle.

IP Switching definiert zwei Protokolle [NEH96a] [NEH96b], welche einerseits das Management eines ATM-Switches übernehmen und andererseits eine Flusserkennung und eine Verbindungssteuerung realisieren. Als Flüsse werden zwei Typen unterschiedlicher Granularität definiert: Der Port Pair Flow Type und der Host Pair Flow Type. Während der erste Typ einzelne IP-Flüsse mit gleichen IP-Adressen und TCP/UDP-Portnummern bezeichnet, umfasst der zweite Typ sämtliche IP-Flüsse zwischen zwei IP-Hosts. Die beiden Protokolle des IP Switching wurden 1996 von der ITF standardisiert.

Tag Switching [RDK97] wurde ursprünglich von CISCO entwickelt und 1997 von der IETF standardisiert. Tag Switching basiert als direkter Vorläufer von MPLS auf den gleichen Ideen und wird daher nicht näher erläutert.

Multiprotocol Label Switching MPLS [RVC01] ist aus den Ansätzen zur Realisierung der Übertragung von IP-Verkehr über ATM-Netze entstanden. Wie im Namen verankert, ist MPLS nicht auf die Kopplung von IP- und ATM-Netzen beschränkt. MPLS ist unabhängig sowohl von den Transport-Technologien als auch von dem Vermittlungs-Protokoll. Neben ATM kann MPLS auf Frame Relay, Ethernet oder Point-To-Point-Protocol PPP aufsetzen. Für Synchronous Digital Hierarchy SDH und die optische Transportnetz-Technologie Wavelength Division Multiplexing WDM wird das Generalized MPLS [B03a], eine Weiterentwicklung von MPLS, verwendet. Als Vermittlungsprotokoll kann neben IP auch das Internet Protocol Version 6 IPv6 auf MPLS aufsetzen.

MPLS ist das einzige der vorgestellten IP over ATM Netzkonzepte, welches eine weite Verbreitung gefunden hat. Der Hauptgrund dafür ist die Tatsache, dass MPLS neben der Kopplung von IP und ATM für weitere Aufgaben im Netz eingesetzt wird. Neben der Anwendung zur lastabhängigen Verkehrlenkung in IP-Netzen, weshalb MPLS im Rahmen dieser Arbeit diskutiert wird, ist die Realisierung von virtuellen privaten Netzen ein weiteres Haupteinsatzgebiet der MPLS-Technologie. Virtuelle private Netze verbinden räumlich voneinander getrennte Netze über offene IP-Weitverkehrsnetze unter Wahrung der Authentizität, Integrität und Vertraulichkeit der Teilnehmer.

2.1.2.2 Grundlagen der MPLS-Technologie Dieser Abschnitt erläutert die grundlegenden Konzepte der MPLS-Technologie.

MPLS lässt sich wie IP nach der Definition des ISO-OSI-Referenzmodells der Vermittlungsschicht zuordnen. Da jedoch IP auf MPLS aufsetzt und MPLS zur Übertragung der IP-Datenpakete verwendet wird, wird MPLS auch als Sub-IP- oder Schicht2.5-Protokoll bezeichnet. Im Gegensatz zum Internet Protocol verfolgt MPLS einen verbindungsorientierten Ansatz. Datenpakete werden auf zuvor eingerichteten Verbindungen, den sogenannten Label Switched Paths, durch das Netz geleitet.

Abbildung 5 zeigt das Beispiel von durchgeschalteten LSPs in einem MPLS-Netz. Der Eingangs-LSR realisiert eine Zuordnung einer oder mehrerer IP-Adressen zu einem LSP.

Die Adresse ankommender IP-Pakete wertet der Eingangs-LSR aus und fügt das entsprechende Label des zugeordneten LSPs an das Paket. Anschließend schickt der Eingangs-LSR das MPLS-Paket über den entsprechenden Ausgangs-Port zum Nachbar-Router. Empfängt ein LSR ein MPLS-Paket, wertet dieser das ankommende MPLS-Label und weitere MPLS-Header-Informationen aus. Erkennt der LSR, dass das MPLS-Paket zum nächsten LSR gesendet werden soll, dann trägt der LSR das auf dem neuen Link gültige Label in den MPLS-Header ein und leitet das MPLS-Paket weiter. Kommt das MPLS-Paket bei dem Ausgangs-LSR an, dann entfernt dieser die MPLS-spezifischen Informationen des Pakets und sendet es als IP-Paket weiter.

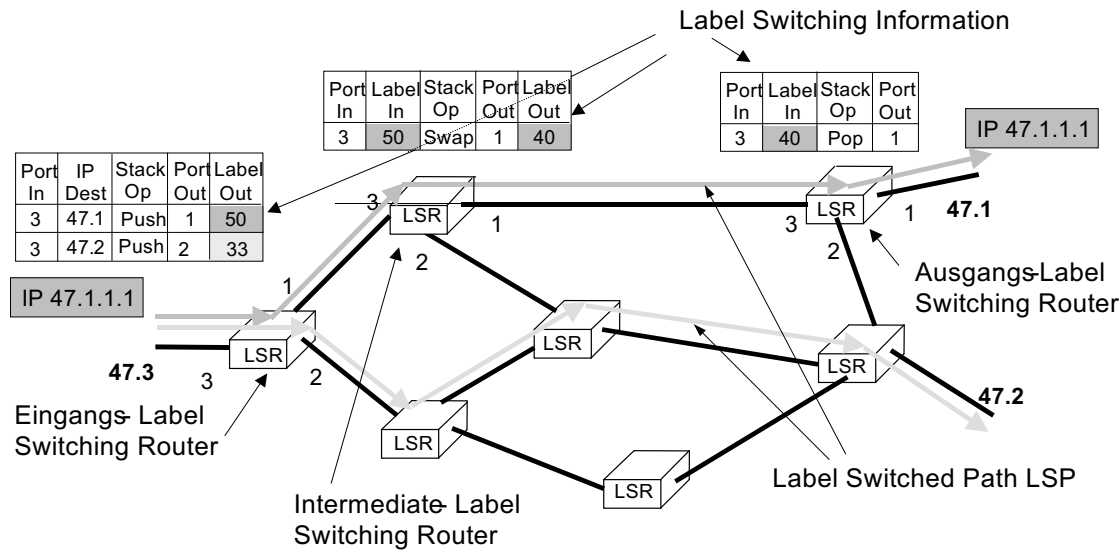


Abbildung 5: Realisierung des Label Switching in einem MPLS-Netz

Das Label kann in der verwendeten Übertragungstechnologie oder durch einen eigenen MPLS-Shim-Header (shim = Unterlegscheibe) transportiert werden (siehe Abbildung 6). Bei paketorientierten Transport-Technologien, die ein Feld für die Übertragung eines Labels bzw. einer Verbindungs-Kennung vorsehen, kann das MPLS-Label in diesen Feldern übertragen werden. Bei ATM können der Virtual Connection Identifier VCI und der Virtual Path Identifier VPI genutzt werden [DLM01], welche pro ATM-Zelle im ATM-Header übertragen werden. Bei Frame Relay Datenpaketen kann das MPLS-Label in dem Data Link Circuit Identifier DLCI Feld des Headers übertragen werden [CDM01].

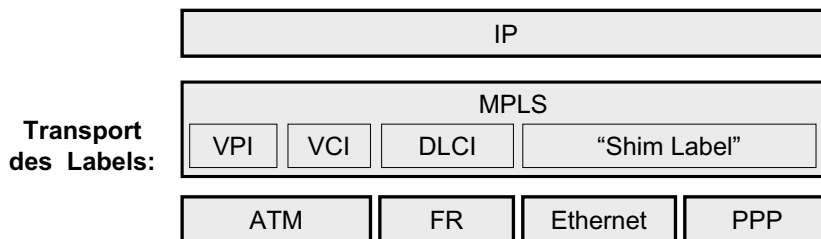


Abbildung 6: Transport des MPLS Labels

Die Nutzung von MPLS auf einer ATM-Infrastruktur schließt die parallele Nutzung von ATM nicht aus. Dieser Betriebsmodus wird auch mit dem Begriff Ships in the Night umschrieben. Bei gleichzeitigem Betrieb von MPLS und ATM muss der zur Verfügung stehende Label Bereich aufgeteilt werden.

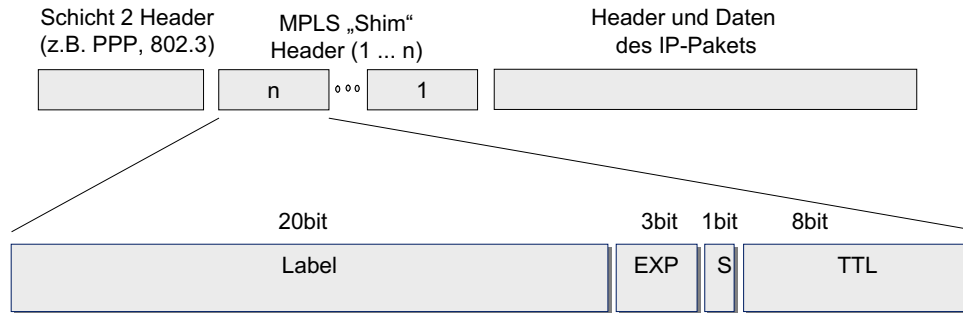


Abbildung 7: MPLS Shim Header

Bei dem Transport über Ethernet oder PPP, welche keine Verbindungs-Kennungen pro Datenpaket vorsehen, nutzt MPLS den Shim-Header, welcher zusätzlich an die IP-Pakete gehängt wird. Der Aufbau des Headers ist in Abbildung 7 dargestellt. Neben dem Label besteht der MPLS-Shim-Header aus einem Experimental Feld, einem Bottom of Stack Feld und einem Time To Live Feld. Die Nutzung des Experimental Feld ist für die Kodierung der Verkehrsklasse vorgesehen. Das TTL-Feld verhindert das endlose Kreisen von MPLS-Paketen auf Routing-Schleifen. Bei jedem Weiterleiten des MPLS-Pakets wird der Wert des Feldes um 1 dekrementiert und beim Erreichen des Wertes Null wird das Paket verworfen. Ein gesetztes Bottom of Stack Bit kennzeichnet den untersten MPLS-Shim-Header des Label Stacks.

Wie in Abbildung 7 dargestellt, bietet MPLS die Möglichkeit, mehrere Label pro Datenpaket zu übertragen. Diese Fähigkeit wird als hierarchisches Label Switching bezeichnet. Während das erste Label wie bereits erläutert sowohl als Shim-Header als auch innerhalb der verwendeten Übertragungstechnologie übertragen werden kann, werden alle weiteren Label als Shim-Header mit dem Datenpaket übertragen. Die Anzahl der Label pro Datenpaket ist nicht begrenzt.

Bei Verwendung von MPLS ist es dem Netzbetreiber überlassen, die Zuordnung zwischen dem ankommenden Verkehr und den LSPs durchzuführen. Typische Kriterien für die Zuordnung der ankommenden IP-Flüsse zu einzelnen LSPs sind z.B. die Portnummern und IP-Adressen des Quell-Hosts und des Ziel-Hosts oder der Eingangs- und Ausgangs-LSR, über welche die Datenpakete in das MPLS-Netz eintreten und das MPLS-Netz wieder verlassen.

Über die variable Aggregation hat ein Netzbetreiber die Möglichkeit, die Granularität des Verkehrs im Netz zu variieren. Wählt ein Netzbetreiber eine feine Granularität und baut z.B. für jeden einzelnen IP-Fluss ein LSP auf, hat der Netzbetreiber einerseits die höchste Flexibilität bei der Verteilung des Verkehrs im Netz. Andererseits bedeutet die Verbindungssteuerung vieler LSPs einen hohen Verarbeitungsaufwand in den Netzknoten. Wählt der Netzbetreiber hingegen eine grobe Granularität und baut z.B. einen LSP pro Eingangs-Ausgangs-LSR-Paar auf, reduziert sich der Verarbeitungsaufwand in den Netzknoten unter Inkaufnahme einer größeren Lastgranularität.

Zur Realisierung der variablen Aggregation definiert MPLS die Forwarding Equivalence Classes FECs. Entsprechend den oben erwähnten Parametern werden die Verkehrsflüsse in FECs zusammengefasst. Der Verkehr eines FECs kann über einen oder mehrere LSPs übertragen werden.

Abbildung 8 zeigt den Aufbau eines Label Switching Routers LSR. Das MPLS-Forwarding

basiert auf der Auswertung eines Labels, welches ein Datenpaket in den Kontrollinformationen mit sich führt. Das steht im Gegensatz zum IP-Forwarding, welches auf der Auswertung der IP-Adresse basiert. Ein Label hat lokale Gültigkeit auf einem Link und wird beim Verbindungsaufbau festgelegt. Die Verbindungssteuerungs-Einheit eines LSRs übergibt der jeweiligen MPLS-Forwarding-Einheit die Label Switching Informationen über neu eingerichtete LSPs. Mit Abbau eines LSPs werden die Label Switching Informationen wieder entfernt. Eine MPLS-Forwarding-Einheit hat ausschließlich die Label Switching Informationen der aktuellen LSPs.

Das Weiterleiten eines MPLS-Datenpakets über die Switching-Matrix und die LineCard entspricht dem Vorgang in einem IP-Router.

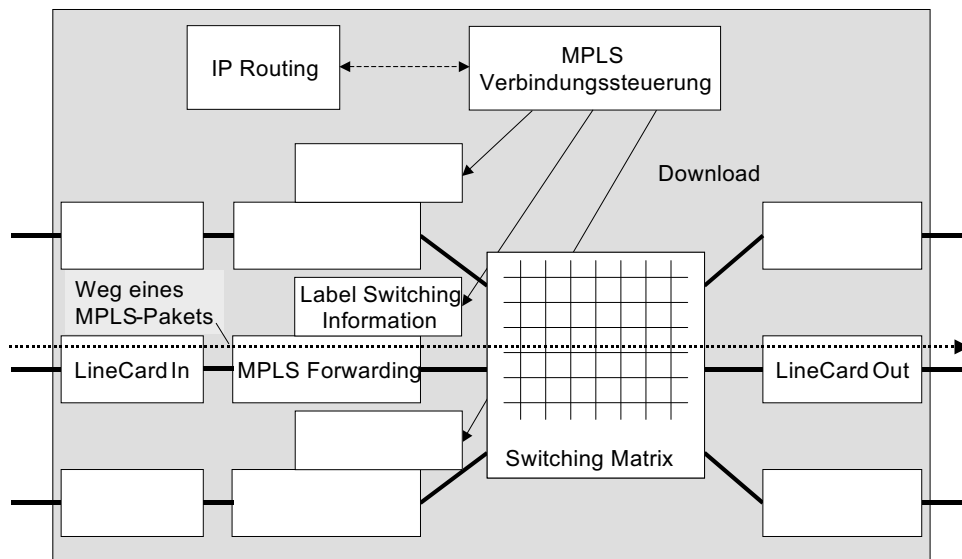


Abbildung 8: Aufbau eines MPLS Label Switching Routers

Für die Berechnung und Festlegung des Pfades beim Aufbau eines LSPs bieten sich bei Verwendung von MPLS mehrere Möglichkeiten. Die erste Möglichkeit beruht auf der Nutzung der IP-Routing-Protokolle. In diesem Fall befragt die Verbindungssteuerungs-Einheit die IP-Routing-Einheit, über welchen Pfad der Aufbau eines LSPs zum Zielsystem erfolgen soll. Dieser Sachverhalt ist in Abbildung 8 dargestellt. Die zweite Möglichkeit ist die Verwendung eines Quell-Routings. In diesem Fall legt die Quelle des LSPs den Pfad durch das Netz fest und übergibt die Pfadinformation in den Verbindungsaufbau-Nachrichten an die nachfolgenden LSRs. Die Verbindungssteuerungs-Einheit bestimmt durch die Auswertung der Pfadinformation den nächsten LSR, über welchen der LSP aufgebaut werden soll. Die Realisierung des Quell-Routings an dem Quell-LSR ist durch MPLS nicht festgelegt.

Die dritte Möglichkeit zur Festlegung des Pfades beruht auf der Kombination der beiden zuvor erwähnten Möglichkeiten. Der Quell-LSR kann einen Teil des Pfades vorgeben bzw. einige Stützpunkte des Pfades definieren. Für das Routing der restlichen Teile des Pfades wird die IP-Routing-Einheit befragt.

2.1.2.3 Grundlagen der Label Distribution Protokolle Dieser Abschnitt geht auf die Protokolle zur Verbindungssteuerung in MPLS-Netzen ein, welche den Aufbau und Abbau der LSPs realisieren. Hierzu zählen das Label Distribution Protocol LDP,

Verbindungsaufbau in verbindungsorientierten Netzen. Der Independent Modus erlaubt jedem LSR im Netz eine Zuordnung eines FECs zu einem Label durchzuführen und dies mit einem Nachbar-LSR festzulegen. Durch Verkettung dieser Label Zuweisungen an den einzelnen LSRs im Netz entstehen Label Switched Paths zwischen einem Eingangs-LSR und einem Ausgangs-LSR. Dieser Modus entspricht der Hop by Hop Weiterleitung von Datenpaketen in verbindungslosen Netzen.

Der Abbau eines LSPs erfolgt je nach Initiierung durch den sendenden oder den empfangenden LSR in unterschiedlicher Weise. Der sendende LSR sendet zum Abbau eines LSPs eine Label Release Nachricht zum nachfolgenden LSR. Im Ordered Control Modus (wie in Abbildung 10 dargestellt) geht die Label Release Nachricht vom Eingangs-LSR aus und wird bis zum Ausgangs-LSR weitergeleitet. Löst der Empfänger die Verbindung aus, sendet dieser eine Label Withdraw Nachricht an den Nachbar-LSR. Dieser antwortet mit einer Label Release Nachricht und sendet anschließend die Label Withdraw Nachricht zum nächsten LSR. Im Ordered Control Modus senden die LSRs die Label Withdraw Nachricht weiter, bis der Eingangs-LSR erreicht ist.

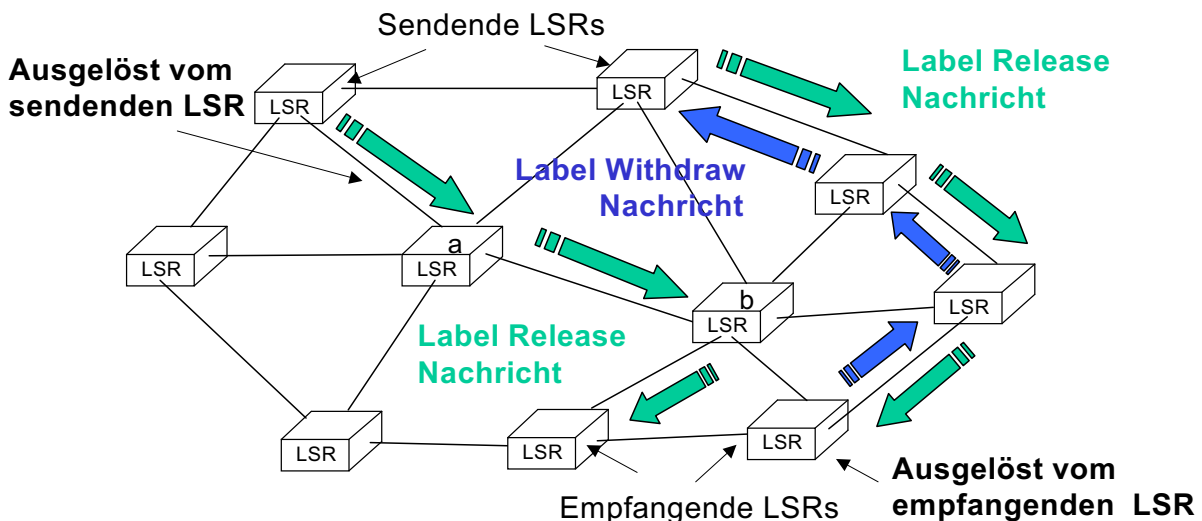


Abbildung 10: LSP Verbindungsabbau im Ordered Control Modus

Die Verwendung des LDP für lastabhängige Verkehrslenkung in MPLS-Netzen hat zwei große Nachteile. LDP erlaubt weder den Aufbau von Verbindungen mit Reservierung von Übertragungskapazität noch die Durchführung eines Quell-Routings. Zur Festlegung der Pfade der LSPs nutzt LDP ausschließlich die IP-Routingprotokolle.

Aus diesem Grund wurde das Constraint-based Routed Label Distribution Protocol CR-LDP [JAC02] entwickelt, welches das LDP um diese beiden Funktionen erweitert. Das CR-LDP verwendet den in Abbildung 9 beschriebenen Modus Downstream On Demand zum Aufbau der LSPs.

Neben dem CR-LDP ist das Resource Reservation Protocol mit Traffic Engineering Erweiterungen RSVP-TE [ABG01] entstanden, welches sich ebenfalls für den Einsatz von MPLS Traffic Engineering eignet.

Ursprünglich ist RSVP als Verbindungssteuerungs-Protokoll für die Integrated Service Architektur [W97] entwickelt worden, um Quality of Service Verbindungen in IP-Netzen zu realisieren. RSVP verwendet ebenfalls den in Abbildung 9 beschriebenen Modus Downstream On Demand zum Aufbau der LSPs. Die Weiterentwicklung von RSVP zum RSVP-

TE bestand im wesentlichen aus zwei Funktionen. Zum einen wurde das RSVP-TE um die zusätzliche Übertragung eines Labels erweitert, um den Aufbau von Verbindungen in MPLS-Netzen zu realisieren. Zum anderen wurde das RSVP-TE um die Funktion des Quell-Routings erweitert. Hierzu wird in den Signalisierungsnachrichten die Pfadinformation mit übertragen.

Das CR-LDP und das RSVP-TE weisen geringe Unterschiede auf und eignen sich beide für die Durchführung lastabhängiger Verkehrslenkung in MPLS-Netzen.

2.2 Realisierungsvarianten lastabhängiger Verkehrslenkung

Die lastabhängige Routing-Kontrolle in IP-Netzen kann in verschiedener Weise realisiert werden. Dieser Abschnitt gibt einen Überblick über die möglichen Varianten (siehe auch [ACE02]).

- IP-basiert versus MPLS-basiert
- Einzelpfad versus Mehrpfad
- (Quasi-)Statisch versus Dynamisch
- Ohne Kapazitätsreservierung versus mit Kapazitätsreservierung
- Zentral versus Dezentral
- Proaktiv versus Reaktiv
- Global versus Lokal
- Offline versus Online

IP-basiert versus MPLS-basiert: IP-basierte Verkehrslenkungsverfahren nutzen die aktuellen Routing-Protokolle in IP-Netzen. Die Verfahren passen die Link-Metriken in der Weise an, dass sich eine gewünschte Lastverteilung einstellt. MPLS-basierte Ansätze bestimmen die Pfade der LSPs durch das Netz und teilen den Eingangsverkehr über die LSPs auf.

Einzelpfad versus Mehrpfad: Je nach gewähltem Ansatz wird der Verkehr zwischen einem Netzeingang und einem Netzausgang über einen Pfad oder verteilt über mehrere Pfade übertragen. Im letzteren Fall unterscheiden sich die Realisierungen, den Verkehr auf Paketebene oder auf Verkehrsflussebene aufzuteilen. Die Verteilung auf Paketebene ist einfach zu realisieren, hat aber den Nachteil von möglichen Paketüberholungen innerhalb eines Verkehrsflusses. Das wiederum kann eine erhebliche Beeinträchtigung der Datenübertragung bedeuten (z.B. Durchsatzeinbruch bei TCP-Verkehrsflüssen). Eine Verteilung auf Verkehrsflussebene, welche die Übertragung sämtlicher Datenpakete eines Verkehrsflusses über einen Pfad garantiert, hat eine höhere Komplexität.

(Quasi-)Statisch versus Dynamisch: Die Unterscheidung zwischen statischem bzw. quasi-statischem und dynamischen Routing bezieht sich auf den Zeitpunkt der Festlegung der Pfade für ankommende Verkehrsflüsse. Bei statischem bzw. quasi-statischem Routing werden die möglichen Pfade einer Verbindung im voraus festgelegt. Neue ankommende

Verkehrsflüsse werden in diesem Fall den bestehenden Pfaden zugeordnet. Quasi-statisches Routing kann im Gegensatz zu rein statischem Routing die Pfade während des operativen Netzbetriebs auf Grund von Topologie-Änderungen oder Überlastsituationen anpassen. Bei dynamischem Routing hingegen wird für jeden neu ankommenden Verkehrsfluss ein neuer Pfad berechnet. Während das Routing in verbindungslosen Netzen und der Mehrpfadansatz auf statischer bzw. quasi-statischer Vorausberechnung der Pfade beruht, kann der verbindungsorientierte Ansatz mit einem Pfad sowohl (quasi-)statisch als auch dynamisch betrieben werden.

Zentral versus Dezentral: Die Steuerung der lastabhängigen Verkehrslenkung kann zentral oder dezentral erfolgen. Bei zentraler Ausführung benötigt die zentrale Instanz Auslastungsinformationen des Netzes. Basierend auf diesen Informationen berechnet die zentrale Instanz die neue Lastverteilung und steuert die Durchführung in den Routern. In der dezentralen Ausführung wird die Anpassung der Verkehrslenkung an die aktuelle Lastsituation von verteilten Einheiten durchgeführt.

Proaktiv versus Reaktiv: Lastabhängige Verkehrslenkung teilt sich in proaktive und reaktive Verfahren auf. Proaktive Verfahren werden in Zeitabständen von Tagen bis Wochen durchgeführt und basieren auf der Kenntnis des aktuellen Eingangsverkehrs. Diese wird aus der Messung des Verkehrs über die letzten Zeitperioden und auf einer Schätzung des zukünftigen Verkehrs gewonnen. Reaktive Verfahren hingegen bestimmen die Netzauslastung über kürzere Zeiträume im Bereich von einigen Sekunden bis wenigen Minuten. Im Gegensatz zu proaktiven Verfahren können diese schnell auf unerwartet auftretende Überlastsituationen reagieren.

Global versus Lokal: Globale Lastverteilung bezeichnet die Neubestimmung des Routings für die gesamten Eingangslast des Netzes. Diese wird in der Regel bei proaktiven Verkehrslenkungsverfahren eingesetzt. Bei lokaler Lastverteilung hingegen erfolgt eine Anpassung des Routings in Teilbereichen des Netzes. Diese wird typischerweise bei reaktiven Verkehrslenkungsverfahren eingesetzt.

Offline versus Online: Die Berechnung der neuen Lastverteilung kann offline unabhängig vom Netzbetrieb oder online während des aktiven Netzbetriebs erfolgen. Während proaktive Verfahren die Lastverteilung offline bestimmen, berechnen reaktive Verfahren online während des Netzbetriebs die neue Lastverteilung, um möglichst schnell Überlastsituationen aufzulösen.

2.3 Definition der optimalen Lastverteilung

Um die Lastverteilung in einem Netz zu optimieren, müssen die relevanten Güteparameter zunächst identifiziert und der optimale Zustand definiert werden. Das erfolgt in diesem Abschnitt.

Zur Festlegung eines Pfades in Abhängigkeit der Auslastung sind zwei Kriterien relevant. Auf der einen Seite sollte der Pfad einen möglichst geringen Hop Count haben und somit ein Minimum an Ressourcen belegen. Auf der anderen Seite gilt es, hohe Link-Auslastungen zu vermeiden. Für Verkehr mit Kapazitätsreservierung bedeuten hohe Link-Auslastungen eine Erhöhung der Blockierungswahrscheinlichkeit. Für Verkehr ohne Kapazitätsreservierung entsteht dadurch eine überproportionale Zunahme der Paketverzögerung und des Paketverlusts. Für TCP-Verkehr leitet sich daraus eine überproportional starke Reduktion des Durchsatzes ab.

Zur Durchführung eines lastabhängigen Routings müssen diese beiden Kriterien bei der Festlegung eines Pfades gegeneinander abgewogen werden. Dies wird in einer Zielfunktion festgelegt. Hierbei unterscheiden sich diverse Zielfunktionen, die in der Folge näher beschrieben werden.

Die erste Zielfunktion ist die Minimierung der Summe der Linkkosten. Die Linkkosten werden durch eine konvexe Funktion beschrieben, welche mit steigender Linkauslastung überproportional anwächst (siehe Abbildung 11). Hohe Link-Auslastungen werden durch hohe Linkkosten gewichtet, so dass die Verwendung der Links für weitere Verkehrsflüsse unwahrscheinlicher wird. Die Linkkostenfunktion kann z.B. von der mittleren Paketverzögerung bzw. Paketverlustwahrscheinlichkeit auf einem Link abgeleitet werden.

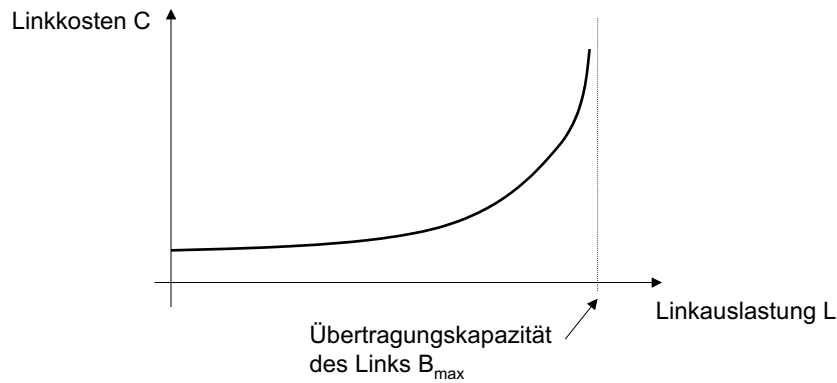


Abbildung 11: Beispiel einer konvexen Linkkostenfunktion

Gleichung 1 zeigt die Zielfunktion, wobei $C(L_{link}(i))$ die Linkkosten des Links i mit der Auslastung $L_{link}(i)$ angibt. E beschreibt die Menge aller Links eines Netzes.

$$\min \sum_{i \in E} C(L_{link}(i)) \quad (1)$$

Um hohe Link-Auslastungen stärker zu wichten, können die Linkkosten mit einem Wert größer eins potenziert werden. Diese Zielfunktion ist in Gleichung 2 beschrieben.

$$\min \sum_{i \in E} C(L_{link}(i))^n \quad \text{mit } n > 1 \quad (2)$$

Eine weitere Zielfunktion ist die Minimierung der maximalen Linkauslastung. Die Gleichung 3 definiert diese Zielfunktion, wobei $L_{link}(i)$ die Auslastung des Links i angibt und E die Menge aller Links im Netz ist. Diese Zielfunktion lässt sich aus der vorherigen Zielfunktion ableiten, in dem der Wert n auf unendlich gesetzt wird.

$$\min (\max_{i \in E} L_{link}(i)) \quad (3)$$

Die Algorithmen der in der Folge beschriebenen Verkehrslenkungsverfahren basieren direkt auf diesen Zielfunktionen oder sind von diesen abgeleitet.

2.4 Proaktive Verkehrslenkungsverfahren

Dieser Abschnitt gibt einen Überblick über die proaktiven Verkehrslenkungsverfahren. Darunter werden in der Folge solche proaktive Verfahren verstanden, welche eine globale Lastverteilung durchführen und diese offline parallel zum Netzbetrieb berechnen.

Zunächst beschreibt Abschnitt 2.4.1 die Anforderungen an die Realisierung von proaktiven Verfahren. Im folgenden Abschnitt 2.4.2 werden die Grundlagen zur Berechnung proaktiver Lastverteilung vorgestellt. Der abschließende Abschnitt beschreibt einige ausgewählte proaktive Verkehrslenkungsverfahren. Dabei werden IP-basierte, MPLS-basierte und hybride IP/MPLS-basierte Verfahren beschrieben.

2.4.1 Anforderungen an proaktive Verkehrslenkungsverfahren

Die Zielvorgabe bei der Realisierung eines proaktiven Verkehrslenkungsverfahrens ist die Bestimmung der optimalen Lastverteilung unter Berücksichtigung der gegebenen Randbedingungen. Eine solche Randbedingung kann beispielsweise die Begrenzung der Anzahl der Verbindungen pro Knotenpaar sein.

Da die in der Folge betrachteten proaktiven Verfahren offline betrieben werden, ist die Anforderung an die Berechnungskomplexität und die Berechnungsdauer gering. Es muss lediglich gewährleistet sein, dass die Berechnungsdauer die Größenordnung von einigen Stunden bzw. Tagen nicht überschreitet. Ferner muss die Komplexität der Berechnung durch die verwendete Hardware handhabbar sein.

Durch die Tatsache, dass die Umschaltung zu fest definierten Zeitpunkten in Phasen schwacher Netzauslastung durchgeführt wird, ergeben sich an den Umschaltvorgang keine großen Anforderungen. Es muss lediglich gewährleistet sein, dass aktive Datenübertragungen nicht gestört oder unterbrochen werden.

2.4.2 Grundlagen proaktiver Verkehrslenkungsverfahren

Dieser Abschnitt beschreibt die bekannten Methoden, die zur Berechnung einer globalen Lastverteilung eingesetzt werden.

Die häufigste Methode zur Bestimmung der globalen Lastverteilung ist die Formulierung eines mathematischen Programms. Hierzu wird zum Beispiel die Sprache AMPL (A Modeling Language for Mathematical Programming) [FGk93] genutzt. Ein mathematisches Programm setzt sich aus einer Zielfunktion und verschiedenen Nebenbedingungen zusammen. Als Zielfunktionen werden die Funktionen aus Abschnitt 2.3 verwendet. Die Nebenbedingungen ergeben sich unter anderem durch den Eingangsverkehr, die Topologie und die Link-Kapazitäten. Das Lastverteilungsproblem in einem Netz wird auch als Multi Commodity Flow MCF Problem bezeichnet.

Sind sowohl die Zielfunktion als auch die Nebenbedingungen lineare bzw. stückweise lineare Funktionen bzw. Gleichungen, dann kann die lineare mathematische Programmierung verwendet werden. Der Simplex-Algorithmus erlaubt für die Klasse der MFC Probleme eine effiziente Lösung selbst für große Netze. Lastverteilungsprobleme, welche die Anzahl und das Routing der Pfade nicht einschränken, lassen sich zum Beispiel auf lineare Programme LP abbilden.

Beinhalten die Zielfunktion oder die Nebenbedingungen Nicht-Linearitäten, so ist eine effiziente Bestimmung der Lösung des mathematischen Programms nicht möglich. Ein Beispiel einer nicht-linearen Zielfunktion ist Gleichung 1 mit einer konvexen nichtlinearen Linkkostenfunktion. Nicht-lineare mathematische Programme NLP lassen sich nur für kleine Netze mit sinnvollem Zeitaufwand lösen.

Eine Sonderklasse der nicht-linearen mathematischen Programme sind die Integer Programme ILP bzw. Mixed-Integer Linear Programme MIP. Während ein ILP ausschließlich aus Integer-Variablen besteht, sind die Variablen eines MIPs teilweise reeller und teilweise ganzzahliger Natur. Die Komplexität von ILPs bzw. MIPs ist NP-hart, was die Anwendung auf kleine Netze begrenzt. Ein MIP liegt z.B. vor, wenn die Anzahl der Pfade pro Verkehrsstrom eines Eingangs-Ausgangs-Knotenpaares begrenzt ist.

Um nicht-lineare bzw. Integer oder Mixed-Integer Lastverteilungsprobleme auch in großen Netzen lösen zu können, werden Heuristiken bzw. Meta-Heuristiken eingesetzt. Hierzu zählen einerseits nicht-deterministische Suchverfahren wie der genetische Algorithmus, Simulated Annealing oder Tabu-Search. Diese suchen in Teilen des Lösungsraums nach den lokalen Optima und wählen aus diesen die beste Lösung aus. Andererseits werden Heuristiken eingesetzt, welche auf den in Abschnitt 2.5.2 beschriebenen lokalen Routing- oder Mehrpfad-Algorithmen basieren.

2.4.3 Bekannte proaktive Verkehrslenkungsverfahren

Dieser Abschnitt präsentiert eine Auswahl von bekannten Vorschlägen zur Realisierung proaktiver Verkehrslenkungsverfahren. Hierzu werden in den folgenden Abschnitten IP-basierte, MPLS-basierte und hybride IP/MPLS-basierte Verfahren vorgestellt.

2.4.3.1 IP-basierte proaktive Verkehrslenkungsverfahren IP-basierte proaktive Verfahren beruhen auf der Anpassung der Link-Metriken zur Steuerung der Lastverteilung im Netz.

Das Verfahren von Fortz und Thorup [FT00] basiert auf OSPF und ECMP und hat die Zielfunktion, die Summe der Linkkosten zu minimieren. Der zur Bestimmung der optimalen Lastverteilung verwendete Ansatz basiert auf einem Hash-basierten Suchverfahren, welches eine Abwandlung von Tabu-Search darstellt.

Wang, Wang und Zhang beschreiben in [WWZ01] die Lastverteilung in verbindungslosen IP-Netzen als LP mit der Zielfunktion der Minimierung der maximalen Linkauslastung. Das Verfahren basiert auf der vereinfachenden Annahme, dass das verwendete Routing-Protokoll die Last beliebig über gleichlange Pfade aufteilen kann.

Die Lösung von Milbrandt et al. [MKS02] basiert auf der Unterteilung eines Netzes in Teilnetze. Die Teilnetze haben eine maximale Größe, so dass die optimale Lösung mittels eines MIPs bestimmt werden kann. Die Zielfunktion ist eine Kombination aus der Reduktion der maximalen Linkauslastung sowie der mittleren Linkauslastung unter Einhaltung einer maximalen Übertragungsverzögerung.

Ericsson et al. [ERP02] verwenden einen genetischen Algorithmus, um die optimale Lastverteilung zu bestimmen. Die Zielfunktion ist die Summe der Linkkosten.

Ein hochinteressantes Verfahren präsentiert Riedl [R02], welcher ebenso genetische Algorithmen zur Suche der optimalen Lösung verwendet. Im Gegensatz zu allen anderen

Verfahren basiert das Verfahren von Riedl auf zwei Metriken, welches die Flexibilität bei der Lastverteilung in IP-Netzen dramatisch erhöht. Die Zielfunktion ist die Minimierung der maximalen Linkauslastung. Das Routing-Protokoll, auf welches die Lösung von Riedl abgebildet werden kann, ist das EIGRP.

2.4.3.2 MPLS-basierte proaktive Verkehrslenkungsverfahren MPLS-basierte proaktive Verkehrslenkungsverfahren optimieren die Lastverteilung durch die gezielte Auswahl von einem oder mehreren Pfaden pro Eingangs-Ausgangsknotenpaar.

Wang und Wang [WW99] beschreiben zwei Verfahren. Das erste Verfahren basiert auf der Formulierung der Lastverteilung als LP mit der Zielfunktion, die maximale Linkauslastung zu minimieren. Entsprechend der sich ergebenden Lastverteilung werden LSPs durch das Netz geroutet. Wenn der Verkehr zwischen einem Eingangs-Ausgangsknotenpaar sich über mehrere Pfade aufteilt, müssen in diesen Fällen mehrere LSPs geschaltet werden.

Das zweite Verfahren von Wang und Wang schränkt die Anzahl der Verbindungen auf einen LSP pro Eingangs-Ausgangsknotenpaar ein. Ausgehend von der Lastverteilung des LPs wählt das zweite Verfahren für die über mehrere Pfade aufgeteilten Verkehrsströme unter Berücksichtigung der Minimierung der maximalen Linkauslastung einen Pfad aus.

Xiao et al. [XHB00] definieren ein proaktives Verfahren für Verkehr mit Kapazitätsanforderung (Constraint Based Routing). Hierzu sortiert der Algorithmus die Verkehrsströme mit absteigender Größe der mittleren Bandbreite und führt sequentiell für jeden Verkehrsstrom ein Shortest Path Routing durch. Nach jedem Routing wird die Kapazität auf den Links um die mittlere Senderate des Verkehrsstroms reduziert.

Mitra und Ramakrishnan [MR01] beschreiben die Lastverteilung für ein Multiservice Netz. Das Verfahren beruht auf einer Kapazitätsreservierung sowohl für die QoS-Verkehrsklassen als auch für die Best Effort Verkehrsklasse. Zur Berechnung der Lastverteilung formulieren Mitra und Ramakrishnan ein LP. Die Zielfunktion ist die Maximierung des Gewinns, welcher sich aus der Maximierung des Durchsatzes und dem Verdienst pro übertragener Bandbreite in der jeweiligen Verkehrsklasse ermittelt. Das Verfahren bestimmt zunächst den optimalen Durchsatz für die QoS-Verkehrsklasse. Anschließend minimiert das Verfahren den Ressourcen-Verbrauch unter der Nebenbedingung, dass der zuvor bestimmte optimale Durchsatz nicht verringert wird. Im dritten Schritt verteilt das Verfahren den maximalen Anteil des Best Effort Verkehrs auf die verbleibenden Netz-Ressourcen.

Das Verfahren von Schnitter und Hasslinger [SH02] legt wie das Verfahren von Xiao et al. eine Reihenfolge der Verkehrsströme fest. Die Bestimmung der Reihenfolge der möglichen Lösungen basiert auf der Verwendung von Simulated Annealing. Nach dem Routing eines Verkehrsstroms werden die Metriken der Links angepasst, welche eine bestimmte Grenze der Linkauslastung überschreiten. Diese Grenze wird definiert durch die maximale Linkauslastung, die sich durch die Lösung des LPs ohne Beschränkung auf einen LSP pro Verkehrsstrom ergibt, und einem zu definierenden Delta. Die beste Lösung ist die Lösung mit der geringsten maximalen Linkauslastung. Existieren mehrere Lösungen mit der gleichen maximalen Linkauslastung, wählt das Verfahren die Lösung mit der geringsten mittleren Linkauslastung.

2.4.3.3 Hybride proaktive IP/MPLS Verkehrslenkungsverfahren Neben den rein IP-basierten bzw. MPLS-basierten Verfahren verwenden einige Ansätze eine Kombination beider Varianten.

Das hybride Verfahren von Riedl [R03] basiert darauf, zunächst als rein IP-basiertes Verfahren die OSPF-Metriken zu optimieren. Hierzu verwendet das Verfahren Simulated Annealing. Die Zielfunktion ist die Minimierung der maximalen Linkauslastung. Anschließend verwendet das Verfahren ein Mixed Integer Programm, um durch das Schalten zusätzlicher LSPs die Lastverteilung zu optimieren. Die maximale Anzahl der zu schaltenden LSPs kann dabei vorgegeben werden. Riedl definiert zwei verschiedene MIPs, die hierfür eingesetzt werden können. Das erste MIP erlaubt die beliebige Auswahl von Pfaden für die zu schaltenden LSPs. Das zweite MIP gibt pro Eingangs-Ausgangsknotenpaar einige mögliche LSP-Pfade vor. Auf Grund der verringerten Komplexität des zweiten MIPs kann das Verfahren auch in großen Netzen eingesetzt werden.

Die Lösung von Köhler und Binzenhöfer [KB03] sieht vor, aus der durch OSPF (bzw. einem optimierten OSPF) erzielten Lastverteilung bestimmte Verkehrsflüsse auszuwählen, welche anschließend über LSPs durch das Netz geroutet werden. Das Hauptkriterium zur Auswahl eines Verkehrsflusses ist die Anzahl der überquerten überlasteten Links. Bei gleich großer Anzahl wird der Verkehrsfluss mit der geringeren mittleren Bandbreite ausgewählt. Anschließend werden die ausgewählten Verkehrsflüsse über LSPs optimal über das Netz mit den verbleibenden Linkkapazitäten verteilt. Die Berechnung der Lastverteilung realisiert ein LP. Die Zielfunktion des LPs ist die Minimierung der maximalen Linkauslastung unter Berücksichtigung der Minimierung der gesamten Netzauslastung.

2.5 Reaktive Verkehrslenkungsverfahren

Reaktive Verkehrslenkungsverfahren führen eine Anpassung der Lastverteilung im Netz auf der Basis der aktuellen Netzauslastung aus. Die Messung der Netzauslastung erfolgt über einen Zeitraum in der Größenordnung einiger Sekunden bzw. Minuten. Dieser Abschnitt definiert die Anforderungen an reaktive Verkehrslenkungsverfahren (Abschnitt 2.5.1) und gibt einen Überblick über die Grundlagen der Verfahren (Abschnitt 2.5.2). Im Anschluss daran erläutert Abschnitt 2.5.3 die in der Literatur erwähnten Verfahren und führt eine Bewertung der Verfahren durch.

2.5.1 Anforderungen an reaktive Verkehrslenkungsverfahren

Dieser Abschnitt erläutert die in der folgenden Aufzählung angegebenen Anforderungen an reaktive Verkehrslenkungsverfahren.

1. Stabilität
2. Netzperformanz
3. Konvergenzdauer
4. Berechnungskomplexität
5. Flexibilität bezüglich des Netz- und Auslastungsszenarios

Stabilität: Eine notwendige Voraussetzung für lastabhängige Verkehrslenkungsverfahren ist die Stabilität. Ein Verfahren ist stabil, wenn bei konstanter Eingangslast keine Routing-Oszillationen auftreten und keine andauernde Anpassung der Lastverteilung erfolgt. Nach

einer Änderung in der Eingangslast muss ein stabiles Verfahren einen neuen stabilen Zustand anstreben.

Netzperformanz:

Die Netzperformanz bewertet die Güte der Lastverteilung, welche das reaktive Verkehrslenkungsverfahren durch die Umverteilung der Last im Netz erzielt. Hierbei wird die Netzperformanz im eingeschwungenen und im transienten Zustand der reaktiven Lastverteilung unterschieden.

Die Netzperformanz im eingeschwungenen Zustand ergibt sich nach Beendigung eines reaktiven Lastverteilungsprozesses. Eine hohe Netzperformanz bedeutet in diesem Fall, dass das reaktive Verkehrslenkungsverfahren bei gegebener konstanter Eingangslast die optimale Verteilung dieser über das Netz erreicht bzw. annähert. Die optimale Lastverteilung ist dabei nach den in Abschnitt 2.3 gegebenen Zielfunktionen oder Abwandlungen dieser definiert.

Die Netzperformanz im transienten Zustand bewertet die Güte der Lastverteilung, die während des Lastverteilungsprozesses auftritt. Idealerweise verbessert sich die Netzperformanz kontinuierlich, und es treten keine Performanzeinbrüche während des Lastverteilungsprozesses auf. Einbrüche der Netzperformanz sind denkbar, wenn die einzelnen Lastumschaltvorgänge unkoordiniert durchgeführt werden und somit kurzzeitige Überlastsituationen auftreten.

Konvergenzdauer: Die Konvergenzdauer bezeichnet die Dauer, die ein reaktives Verkehrslenkungsverfahren ausgehend von einer beliebigen initialen Lastverteilung zum Erreichen der resultierenden Lastverteilung benötigt. Diese setzt sich aus den Dauern zur Erkennung einer Überlastsituation sowie zur Durchführung der Lastumverteilung zusammen. Die Güte eines Verkehrslenkungsverfahrens zeichnet sich durch eine möglichst geringe Konvergenzdauer aus.

Berechnungskomplexität: Reaktive Verkehrslenkungsverfahren berechnen neue Lastverteilungen während des Netzbetriebs. Um einerseits die Konvergenzdauer und andererseits den Prozessoraufwand in den TE-Einheiten so gering wie möglich zu halten, sollte die Berechnungsvorschrift eine möglichst geringe Komplexität aufweisen.

Flexibilität bezüglich des Netz- und Auslastungsszenarios: Idealerweise ist ein reaktives Verkehrslenkungsverfahren flexibel in beliebigen Netzen einsetzbar und erreicht unabhängig vom Auslastungsszenario eine definierte Güte des Konvergenzverhaltens. Eine notwendige Voraussetzung des Netzes für den Einsatz von Verkehrslenkungsverfahren muss allerdings die Existenz von zwei oder mehr Pfaden zwischen allen oder zumindest einem Teil der Eingangs-Ausgangs-Knotenpaare sein.

2.5.2 Grundlagen reaktiver Verkehrslenkungsverfahren

Dieser Abschnitt beschreibt die grundlegenden Prinzipien und Algorithmen, auf welchen die Realisierung reaktiver Verkehrslenkungsverfahren basiert. Hierbei werden drei verschiedene Realisierungs-Möglichkeiten unterschieden.

Abschnitt 2.5.2.1 beschreibt ein Prinzip zur Realisierung von reaktiven Verkehrslenkungsverfahren in verbindungslosen Netzen unter Verwendung der bekannten Routing-Protokolle und zeigt die Problematik dieser Variante auf. Abschnitt 2.5.2.2 stellt einige lastabhängige Routing-Algorithmen vor, welche Pfade unter Berücksichtigung der Netzauslastung

berechnen. Die grundlegenden Algorithmen zur Realisierung von reaktiven Verkehrslenkungsverfahren nach dem Mehrpfadansatz beschreibt Abschnitt 2.5.2.3.

2.5.2.1 Lastabhängige Verkehrslenkung in verbindungslosen Netzen Dieser Abschnitt beschreibt ein von Bertsekas und Gallagher [BG87] definiertes Prinzip zur Realisierung eines dezentralen lastabhängigen Routings für das verbindungslose Internet.

Die Grundlage des Prinzips ist die Abbildung der aktuellen mittleren Paketverzögerung auf einem Link auf die Metrik. Die mittlere Paketverzögerung auf den Links bestimmen die Router in kurzen periodischen Intervallen und verteilen die neuen Routing-Informationen an die Nachbar-Router. Die Routing-Protokoll-Instanzen bestimmen auf der Basis der additiven Link-Metriken dynamisch den Pfad mit der geringsten Paketverzögerung.

Der große Nachteil dieses Algorithmus ist das instabile Verhalten und das mögliche Auftreten von Routing-Oszillationen. Einen solchen Fall stellt Abbildung 12 dar. In dem links dargestellten Auslastungs-Szenario ergibt die Messung der Paketverzögerung, dass für den dargestellten Verkehrsstrom der untere Pfad die geringste Ende-zu-Ende-Paketverzögerung hat. Also wird in der Folge der Verkehrsstrom über den unteren Pfad geroutet. Dieser Fall ist rechts in der Abbildung dargestellt. Durch das Umlenken des Verkehrsflusses werden jedoch die Links des alten Pfades entlastet und die Links des neuen Pfades zusätzlich belastet. Das führt zu einer Verringerung der Paketverzögerung und folglich der Link-Metriken auf dem alten und zu einer Erhöhung der Paketverzögerung und der Link-Metriken auf dem neuen Pfad. In dem dargestellten Fall hat das zur Folge, dass der alte Pfad nun wieder der kürzeste Pfad ist und der Verkehrsstrom über den alten Pfad geroutet wird. Somit entsteht eine dauerhafte Oszillation der Lastverteilung ohne Änderung des Eingangsverkehrs.

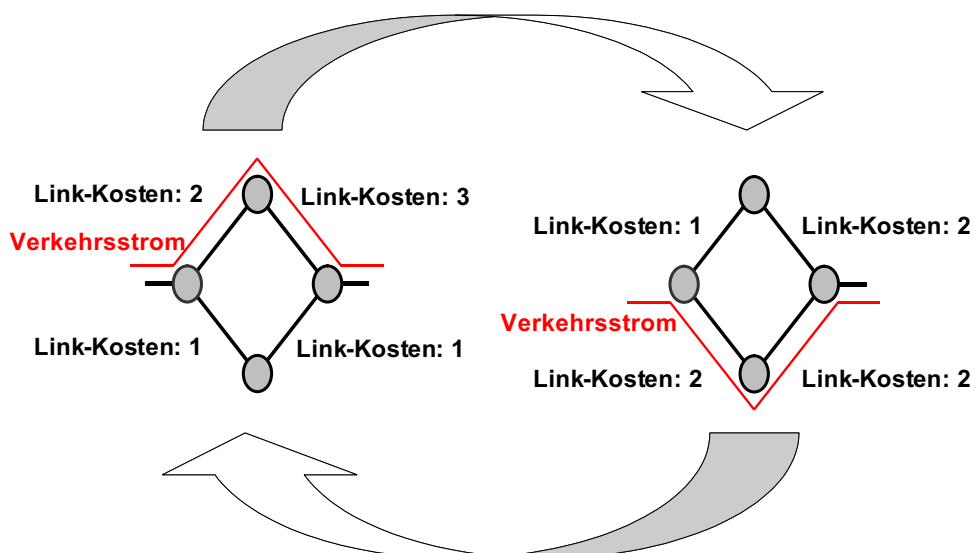


Abbildung 12: Routing Oszillationen in verbindungslosen Netzen mit additiven Metriken

Zur Verhinderung der Routing-Oszillationen schlagen Bertsekas und Gallagher einen konstanten additiven Term pro Link-Metrik vor. Dieser verbessert zwar die Stabilität, stellt aber keine Garantie für das Ausbleiben von Routing-Oszillationen dar. Zudem schränkt diese Erweiterung die Sensitivität gegenüber Überlastsituationen ein.

2.5.2.2 Lastabhängige Routing-Algorithmen Dieser Abschnitt gibt einen Überblick über einige Routing-Algorithmen, welche die Pfadberechnung in Abhängigkeit der Last durchführen.

Ma und Steenkiste stellen in [MSZ96] und [MS97b] grundlegende lastabhängige Algorithmen zur Pfadberechnung vor und vergleichen diese. Alle dort vorgestellten Algorithmen kombinieren mit unterschiedlicher Gewichtung die Minimierung des Hop Counts sowie die Minimierung der Link-Auslastungen.

Shortest Widest Path Algorithmus: Der Shortest Widest Path Algorithmus wählt zunächst den Pfad aus, welcher die größte verbleibende Kapazität zur Verfügung stellt. Existieren mehrere Pfade mit gleicher Kapazität, wählt der Algorithmus den kürzesten Pfad aus. Ursprünglich stammt der Algorithmus von Wang und Crawford [WC96].

Widest Shortest Path Algorithmus: Der Widest Shortest Path Algorithmus bestimmt zunächst den kürzesten Pfad zwischen der Quelle und der Senke. Existieren mehrere kürzeste Pfade der gleichen Länge, wählt der Algorithmus den Pfad mit der größten verbleibenden Kapazität aus. Dieser Algorithmus wurde zunächst von Guerin und Orda diskutiert [GOW97].

Shortest Distance Algorithmus: Der Shortest Distance Algorithmus bestimmt den kürzesten Pfad aus der Summe der Link-Metriken, welche der mittleren Paketverzögerung bzw. dem Kehrwert der verbleibenden Kapazität auf dem Link entsprechen. Diese Variante wurde bereits von [BG87] vorgeschlagen.

Das Routing eines Verkehrsflusses mit Kapazitätsreservierung basiert ausschließlich auf den Links, die die QoS-Anforderungen des Verkehrsflusses erfüllen. Das wird auch als QoS Routing oder Constraint Based Routing bezeichnet.

Ma und Steenkiste haben in [MS97b] die genannten Routing-Algorithmen für die Verteilung von Verkehr mit Kapazitätsreservierung verglichen. Dabei erzielte eine modifizierte Variante des Widest Shortest Path Algorithmus, bei dem zusätzlich die maximal erlaubte Pfadlänge auf die Länge des Shortest Path plus einen weiteren Hop begrenzt wird, die geringsten Blockierungswahrscheinlichkeiten. Das zeigt, dass der Hauptaspekt für Verkehr mit Kapazitätsreservierung auf der Minimierung des Ressourcenbedarfs liegt. Nur in Ausnahmefällen, wenn der bzw. die kürzesten Pfade belegt sind, ist die Verwendung längerer Pfade sinnvoll. Diese Variante ist eine Abwandlung des Verfahrens von Gibbens, Kelly und Key [GKK88], welches auf einer Reservierung von Link-Kapazität für die kürzesten Pfade basiert.

Für Best Effort Verkehr ohne Kapazitätsreservierung zeigen Ma, Steenkiste und Zhang in [MSZ96], dass der Shortest Distance Algorithmus die beste Netzperformanz erzielt. Dieser Algorithmus legt ein größeres Gewicht auf die Lastverteilung und die Vermeidung von lokalen Überlastsituationen. Damit wird der Algorithmus der Tatsache gerecht, dass sämtliche Best Effort Verkehrsflüsse der hoch ausgelasteten Links beeinträchtigt werden. Typischerweise wählt der Shortest Distance Algorithmus bereits alternative Pfade, wenn Links noch nicht zu 100% ausgelastet sind.

Der Shortest Widest Path Algorithmus erzielt bei beiden Verkehrstypen die geringste Performanz. Dieser Algorithmus schränkt den Ressourcen-Bedarf einzelner Verkehrsflüsse nicht ausreichend ein.

Eine Weiterentwicklung des Shortest Distance Path Algorithmus ist das in [BG87] beschriebene Routing im CODEX-Netz. Dies berücksichtigt die Tatsache, dass die Last des

zu routenden Verkehrsflusses einen nicht zu vernachlässigenden Einfluss auf die Linkkosten der jeweiligen Links hat.

Daher wählt der Algorithmus den Pfad durch das Netz aus, welcher die geringste Erhöhung der Summe der Linkkosten im Netz darstellt. Hierzu berechnet der Algorithmus zunächst für jeden Link die Linkkosten C_i , die der Link mit und ohne der zusätzlichen Last ΔL des zu routenden Verkehrsflusses hat. Die Differenz der Linkkosten ΔC_i verwendet der CODEX Routing Algorithmus als Link-Metrik zur Berechnung des kürzesten Pfades (siehe Gleichung 4). Der berechnete Pfad ist somit der Pfad, auf welchem ein Verkehrsstrom die geringsten zusätzlichen Linkkosten erzeugt.

$$\Delta C_i = C_i(L_i + \Delta L) - C_i(L_i) \quad (4)$$

2.5.2.3 Lastabhängige Mehrpfadansätze Die Mehrpfad-Lastverteilung basiert darauf, dass ein Eingangsknoten den Verkehr zu einem Ausgangsknoten über zwei oder mehrere fest eingerichtete Verbindungen verteilt und auf diese Weise die Lastverteilung im Netz optimiert. Neben der initialen Pfadbestimmung beschreibt dieser Abschnitt Lastverteilungsalgorithmen für Verkehr mit und ohne Kapazitätsreservierung.

Das wichtigste Kriterium beim Routing der Mehrpfad-Verbindungen eines Eingangs-Ausgangsknotenpaares ist, dass die Verbindungen unterschiedliche Pfade durch das Netz haben. Eine Möglichkeit ist die Verwendung des einfachen K-Shortest-Path-Algorithmus, um K kantendisjunkte Pfade zu ermitteln. Die Vorgehensweise des Algorithmus ist wie folgt (siehe auch [B99]):

1. Der K-Shortest-Path Algorithmus bestimmt basierend auf den Link-Metriken den kürzesten Pfad durch das gegebene Netz.
2. Anschließend blendet der Algorithmus die verwendeten Links aus dem Netz aus und geht zurück zum ersten Schritt.
3. Der Algorithmus endet, wenn entweder K Pfade bestimmt sind oder kein weiterer Pfad mehr existiert.

Weitere K-Shortest-Path-Algorithmen sind in [B99] definiert. Einer davon ist der Algorithmus von Suurballe, welcher zur Bestimmung von zwei Pfaden die kürzeste Schleife zwischen einem Netzeingang und Netzausgang bestimmt.

Mehrpfad-Verkehrslenkungsverfahren unterscheiden sich für die Anwendung bei Verkehr mit und ohne Kapazitätsreservierung. Für Verkehr mit Kapazitätsreservierung unterscheidet man zudem den Fall, ob für die Mehrpfad-Verbindungen Kapazität reserviert ist oder nicht. Im ersten Fall erfolgt die Verbindungsannahme für einzelne Verkehrsflüsse ausschließlich am Netzeingang, während im zweiten Fall jeder Knoten auf dem Pfad über die Verbindungsannahme einzelner Verkehrsflüsse entscheidet.

Es existieren verschiedene Verfahren, die für eine neue Verbindung die Auswahl eines der vorhandenen Pfade definieren. Diese können z.B. auf der Festlegung einer Absuchreihenfolge, auf der gleichmäßigen Verteilung über die Pfade oder der Bündelung auf einzelnen Pfaden basieren. Eine Übersicht über die Verfahren geben Girard [G90] und Siebenhaar [S96]. Siebenhaar schlägt zudem ein Verfahren vor, welches die reservierte Kapazität der

verschiedenen Mehrpfade gegeneinander austauscht, um die Blockierung im Netz zu reduzieren.

Das Mehrpfad-Verkehrslenkungsverfahren für Verkehr ohne Kapazitätsreservierung verteilt die Last unter dem Gesichtspunkt der Minimierung der gesamten Linkkosten. Das Verfahren beschreiben Bertsekas und Gallager in [BG87].

Das als optimales Routing bezeichnete Verfahren beruht auf dem Kuhn-Tucker-Theorem. Das besagt, dass die Summe der Linkkosten dann minimal ist, wenn ausschließlich die Pfade mit minimaler erster Ableitung der Pfadkosten die Last tragen. Voraussetzung für dieses Theorem ist, dass die Linkkosten einer über der Linkauslastung konvexen Funktion entsprechen.

Als Konsequenz aus dem Theorem verringert sich die Summe der Linkkosten dann, wenn die Last von einem Pfad mit größerer erster Ableitung der Pfadkosten auf einen Pfad mit geringerer erster Ableitung der Pfadkosten verlagert wird. Diese Tatsache wird anschaulich durch die Gleichung 5 beschrieben.

$$\begin{aligned}
 C_{totneu} &= C_{totalt} + \frac{\delta C_{PfadB}(L)}{\delta L} \cdot \Delta L - \frac{\delta C_{PfadA}(L)}{\delta L} \cdot \Delta L < C_{totalt} & (5) \\
 \text{mit} \quad &\frac{\delta C_{PfadB}(L)}{\delta L} < \frac{\delta C_{PfadA}(L)}{\delta L} \\
 \text{mit} \quad &\frac{\delta C_{Pfad}(L)}{\delta L} = \sum_i^{Links} \frac{\delta C(L_i)}{\delta L_i}
 \end{aligned}$$

Die üblichste Methode zur Bestimmung der optimalen Lastverteilung basierend auf dem Kuhn-Tucker-Theorem ist der Gradienten-Projektions-Algorithmus [BG87]. Der in Gleichung 6 beschriebene Algorithmus beruht darauf, dass die Lastverteilung $L_{lsp}(t)$ aller Pfade in die entgegengesetzte Richtung des Gradienten der Pfadkosten $\nabla C_{Pfad}(L_{lsp}(t))$ angepasst wird. Die Größe der Anpassung wird dabei durch die Schrittweite γ vorgegeben. Die Schrittweite γ muss so gering gewählt werden, dass keine Lastoszillationen auftreten können (siehe hierzu auch Abschnitt 2.5.3.1). Berechnet der Gradienten-Projektions-Algorithmus negative Eingangslasten, projiziert der Algorithmus diese in den gültigen Bereich auf den Wert null. Diese Projektion beschreibt Gleichung 6 durch die Schreibweise $[]^+$.

$$L_{lsp}(t+1) = [L_{lsp}(t) - \gamma \cdot \nabla C_{Pfad}(L_{lsp}(t))]^+ \quad (6)$$

2.5.3 Bekannte reaktive Verkehrslenkungsverfahren

Dieser Abschnitt stellt die in der Literatur vorgeschlagenen relevanten Verfahren zur reaktiven Lastverteilung vor. Sämtliche Verfahren basieren auf dem Mehrpfadansatz. Es existieren keine relevanten Vorschläge für reaktive Verkehrslenkungsverfahren, welche auf einem Pfadumlegungsansatz bzw. auf einem verbindungslosen Ansatz beruhen.

2.5.3.1 MATE: MPLS Adaptive Traffic Engineering Das meist zitierte reaktive MPLS Verkehrslenkungsverfahren ist MPLS Adaptive Traffic Engineering MATE [EJS01] [EJS02]. MATE ist ein verteilter Mehrpfadansatz, bei dem die Eingangs-LSRs die Lastverteilung berechnen und durchführen. MATE basiert auf dem optimalen Routing von Bertsekas und Gallager [BG87] und verwendet zur Berechnung der Lastverteilung den Gradienten-Projektions-Algorithmus.

Eine MATE-Instanz in einem Eingangs-LSR verteilt die IP-Pakete eines Verkehrstroms so über die LSPs, dass die Datenpakete eines IP-Verkehrsflusses über einen Pfad übertragen werden. Dadurch werden Paketüberholungen innerhalb eines IP-Verkehrsflusses verhindert, welche insbesondere bei Verwendung von TCP erhebliche Performanz-Einbrüche bedeuten können. Hierzu berechnet MATE auf einem oder mehreren Feldern des IP-Headers (z.B. Quell-IP-Adresse und/oder Ziel-IP-Adresse) mittels einer Hash-Funktion (z.B. Cyclic Redundancy Check) einen Zahlenwert. Eine Modulo-N Operation teilt die sich ergebenden Werte in N Untergruppen auf. Sämtliche IP-Pakete mit dem gleichen Wert werden über denselben Pfad durch das Netz geroutet. Je nach berechneter Lastverteilung am Netzeingang werden die N Untergruppen auf die Pfade aufgeteilt. Die Verwendung einer Hash-Funktion hat den Vorteil, dass die IP-Adressen zufällig den N Untergruppen zugeordnet werden und somit keine ungleichmäßige Verteilung der Last über die Untergruppen bedingt durch bevorzugt ausgewählte IP-Adress-Bereiche entsteht.

Die Messung des Verkehrs beruht auf dem aktiven Senden von Mess-Paketen durch den Eingangs-LSR. Der zugehörige Ausgangs-LSR des LSPs sendet die empfangenen Mess-Pakete mit Zeitstempel und der Anzahl der bisher erhaltenen Mess-Pakete zurück an den Eingangs-LSR. Basierend auf dem Zeitstempel bestimmt der Eingangs-LSR die Verzögerung der Pakete. Um auf eine zeitliche Synchronisierung der Eingangs- und Ausgangs-LSR zu verzichten, beruht die Bestimmung einer Metrik auf der Differenz der gemessenen Paketverzögerungen.

Um die Paketverlustwahrscheinlichkeit zu bestimmen, enthält ein Mess-Paket ein Feld für eine Sequenznummer und ein Feld, in welches der Ausgangs-LSR die Anzahl an empfangenen Paketen einträgt. Anhand dieser beiden Werte kann der Eingangs-LSR die Paketverlustwahrscheinlichkeit ermitteln. Aus den Messergebnissen ermittelt MATE mittels der Bootstrap-Methode die mittlere Paketverzögerung und die Paketverlustrate. Aus den Untersuchungen von MATE in [EJS01] lässt sich auf eine verwendete Messperiode in der Größenordnung von 30 Sekunden schließen.

Die Eingangs-LSRs des MATE-Verfahrens arbeiten autonom und passen die Last auf den abgehenden LSPs ohne Koordination mit den anderen Eingangs-LSRs an. Um die Oszillation der Verteilung der Last zu verhindern, muss die Schrittweite, mit der der Gradienten-Projektions-Algorithmus die Lastanpassung berechnet, ausreichend klein gewählt werden. Hierbei muss zudem berücksichtigt werden, dass die insgesamt im Netz durchgeführten Lastanpassungen die Messungen beeinflussen. In [EJS01] wird gezeigt, dass die Schrittweite γ entsprechend der Gleichung 7 gewählt werden muss.

Die Lipschitz-Konstante $K_{lipschitz}$ bezieht sich auf die Kostenfunktion und definiert die maximale Grenze der Steigung der Kostenfunktion (siehe Gleichung 8). Der Parameter t_0 gibt die Dauer der Messperiode an. Die gesamte Anzahl an LSPs π , der maximale Hop Count h sowie die maximale Anzahl an LSPs auf einem Link λ variieren mit der Größe des Netzes.

$$\gamma < \frac{1}{K_{lipschitz} \cdot (1 + \pi \cdot h \cdot \lambda \cdot (2 \cdot t_0 + 1))} \quad (7)$$

mit $K_{lipschitz} = Lipschitz - Konstante$

mit $t_0 = Dauereiner Messperiode$

mit $\pi = Gesamte Anzahl an LSPs im Netz$

mit $h = Maximale Anzahl an Hops eines LSPs$

mit $\lambda = Maximale Anzahl an LSPs pro Link$

$$\left| \frac{\delta C_l(x_1)}{\delta x} - \frac{\delta C_l(x_2)}{\delta x} \right| \leq K_{lipschitz} \cdot |x_1 - x_2| \quad (8)$$

Gleichung 7 zeigt, dass mit zunehmender Netzgröße und dem daraus folgenden Anstieg der Werte π , h und λ die Schrittweite γ kleiner wird. Bei großen Netzen ist aus diesem Grund eine geringe Konvergenzgeschwindigkeit zu erwarten. Da das Verhalten von MATE lediglich in kleinen Netzen (8 Knoten und 12 Links bzw. 12 Knoten und 15 Links) und einer geringen Anzahl an LSP (6 LSPs bzw. 9 LSPs) untersucht wurde, muss hierzu jedoch eine quantitative Analyse durchgeführt werden (siehe hierzu auch Kapitel 5).

2.5.3.2 Dynamic Traffic Partitioning in MPLS Networks Ein zu MATE vergleichbares Verfahren definieren Dinan, Awduche und Jabbari in [DAJ00a]. Wie MATE basiert das Verfahren auf einem verteilten Mehrpfadansatz, bei welchem die Eingangs-LSRs autonom ohne Koordination mit den anderen Eingangs-LSRs die Lastverteilungsaktionen durchführen. Ebenso arbeitet das Verfahren auf Basis der Messung der Paketverlustwahrscheinlichkeit und der mittleren Paketverzögerung.

Um die Auslastung dynamisch über die Pfade zu verteilen, bestimmt das Verfahren die Sensitivität der mittleren Paketverzögerung auf den Ende-zu-Ende-Pfaden sowie den Links gegenüber Laständerungen. Die Sensitivität entspricht der ersten Ableitung der mittleren Paketverzögerung in Abhängigkeit von der Auslastung. Die Last wird von den Pfaden mit der größten Sensitivität auf die Pfade mit der geringsten Sensitivität verteilt. Dieser Ansatz ist mit dem in [BG87] beschriebenen optimalen Routing von Bertsekas und Gallagher vergleichbar, welches ebenfalls die Last auf die Pfade mit minimaler erster Ableitung der Pfadkosten verteilt.

Dinan et al. verwenden eine analytische Methode zur Bewertung der Verfahren. Die Pfade werden dabei durch eine Verkettung von M/M/1/K-Puffern modelliert, welche jeweils die Ausgangspuffer in den Netzknoten darstellen.

Die Stabilität des Systems wird in [DAJ00b] nachgewiesen. Der Beweis beruht auf der Begründung, dass die Sensitivität eines Pfades gegenüber einer Laständerung immer größer als die Sensitivität eines auf dem Pfad liegenden Links ist. Daraus lässt sich folgern, dass die Laständerung eines Pfades geringer als die Summe der Laständerungen auf den einzelnen Links ist. Einer initialen Laständerung im Netz folgen somit immer geringer werdende Laständerungen auf den Pfaden. Damit strebt das System in einen stabilen Zustand.

Allerdings wird bei diesem Stabilitätsbeweis der Fall vernachlässigt, dass mehrere Pfade gleichzeitig die Lastverteilung basierend auf dem gleichen Netzzustand durchführen. Für diesen Fall ist das Auftreten von instabilen Zuständen durchaus denkbar.

2.5.3.3 OpIATE: Optimization Integrated Adaptive Traffic Engineering Wie das zuvor beschriebene Verfahren MATE basiert OpIATE [SP02] auf einem verteilten Mehrpfadansatz. Im Gegensatz zu MATE tauschen die Eingangs-LSRs nach dem OpIATE-Verfahren Auslastungs-Informationen der einzelnen LSPs aus und koordinieren so die Lastverteilungsaktionen der verteilten Eingangs-LSRs. Hierzu misst jeder Eingangs-LSR die mittlere Auslastung der abgehenden LSPs und verteilt diese Informationen periodisch an alle anderen Eingangs-LSRs. Alle Eingangs-LSRs tauschen die Informationen über die mittleren Auslastungen der LSPs gleichzeitig aus und führen anschließend die Lastumverteilungsaktionen durch. Die Realisierung der Methode zur gleichzeitigen Lastverteilung wird nicht näher spezifiziert.

Mit der Kenntnis der Pfade der LSPs durch das Netz hat jeder Eingangs-LSR die Sicht über die gesamte Netzauslastung. Auf Grundlage dieser Information formuliert OpIATE ein NLP mit der Zielfunktion, die Summe der Linkkosten zu reduzieren. Die Linkkosten werden von der mittleren Verzögerung eines M/M/1 Wartesystems abgeleitet. Zur Lösung des NLP und Berechnung der optimalen Lastverteilung wendet OpIATE den Gradienten-Projektions-Algorithmus an.

Die Lastanpassung an die berechnete optimale Lösung erfolgt mit einer empfohlenen Schrittgröße von $\alpha \in [0.01, 0.2]$. Damit werden zu starke Lastschwankungen in der Übergangsphase vermieden.

Die grundsätzliche Problematik des Verfahrens ist, dass die Berechnung der Lösung des nicht-linearen Programms mit Hilfe des Gradienten-Projektions-Algorithmus eine hohe Komplexität aufweist. Das lässt eine eingeschränkte Skalierbarkeit des Verfahrens vermuten. Da in der in [SP02] vorgestellten Untersuchung auf diesen Punkt nicht eingegangen wird, kann jedoch an dieser Stelle keine quantitative Aussage hierzu getroffen werden.

2.6 Zusammenfassung des Stands der Technik

Die Analyse des Stands der Technik proaktiver Verkehrslenkungsverfahren zeigt, dass hierfür in der Literatur eine Vielzahl an Vorschlägen existiert. Dabei unterscheiden sich Lösungsvorschläge für IP-Netze, für MPLS-Netze und für kombinierte Ansätze, welche sowohl auf IP als auch auf MPLS basieren.

Die Auswertung des Stands der Technik reaktiver Verkehrslenkungsverfahren zeigt, dass lediglich drei relevante Ansätze existieren. Die Analyse dieser Ansätze hat gezeigt, dass diese nicht zu vernachlässigende Schwächen aufweisen. Das Verfahren MATE lässt eine geringe Konvergenzgeschwindigkeit bei großen Netzen erwarten. Da hierfür keine quantitative Bewertung vorliegt, muss diese Aussage noch in einer entsprechenden Untersuchung verifiziert werden (siehe hierzu Kapitel 5). Das Verfahren von Dinan et al. lässt in bestimmten Fällen das Auftreten instabiler Zustände im Netz zu. Das Verfahren OpIATE beruht auf einer komplexen Berechnungs-Vorschrift zur Bestimmung der Lastverteilung und lässt somit eine eingeschränkte Skalierbarkeit erwarten. Eine quantitative Analyse der Berechnungs-Komplexität ist jedoch ebenfalls noch erforderlich.

Diese Analyse zeigt den Bedarf nach einem Verfahren, welches einerseits unabhängig von der Netzgröße schnell konvergiert und welches andererseits auf einer verringerten Berechnungs-Komplexität beruht.

3 Das verteilte reaktive Verkehrslenkungsverfahren

Die Analyse des Stands der Technik in Kapitel 2 zeigt den Bedarf an skalierbaren verteilten reaktiven Verkehrslenkungsverfahren in IP- bzw. MPLS-Netzen. Aus diesem Grund wurde im Rahmen dieser Arbeit ein MPLS-basiertes Verfahren entwickelt, welches dieses Kapitel präsentiert. Das Verfahren kann sowohl nach dem Mehrpfadansatz als auch nach dem Pfadumlegungsansatz betrieben werden.

Zunächst gibt Abschnitt 3.1 eine Übersicht über das System, in welches das Verkehrslenkungsverfahren eingebettet ist. Die Beschreibung der Hauptaspekte des Verfahrens erfolgt in den folgenden drei Abschnitten. Abschnitt 3.2 beschreibt die Funktionsweise des Monitoring-Systems, Abschnitt 3.3 geht auf den Ablauf des Lastverteilungsprozesses und die Koordination der verteilten Traffic Engineering Einheiten ein und Abschnitt 3.4 erläutert die verwendeten Algorithmen für den Mehrpfad- und den Pfadumlegungsansatz. Im anschließenden Abschnitt 3.5 erfolgt eine qualitative Bewertung des reaktiven verteilten Verkehrslenkungsverfahrens. Die wesentlichen Aspekte dieses Kapitels fasst Abschnitt 3.6 zusammen.

3.1 Das verteilte reaktive MPLS Traffic Engineering System

Dieser Abschnitt beschreibt das System, auf welchem das entwickelte reaktive Verkehrslenkungsverfahren basiert. Abbildung 13 zeigt die Komponenten des Systems.

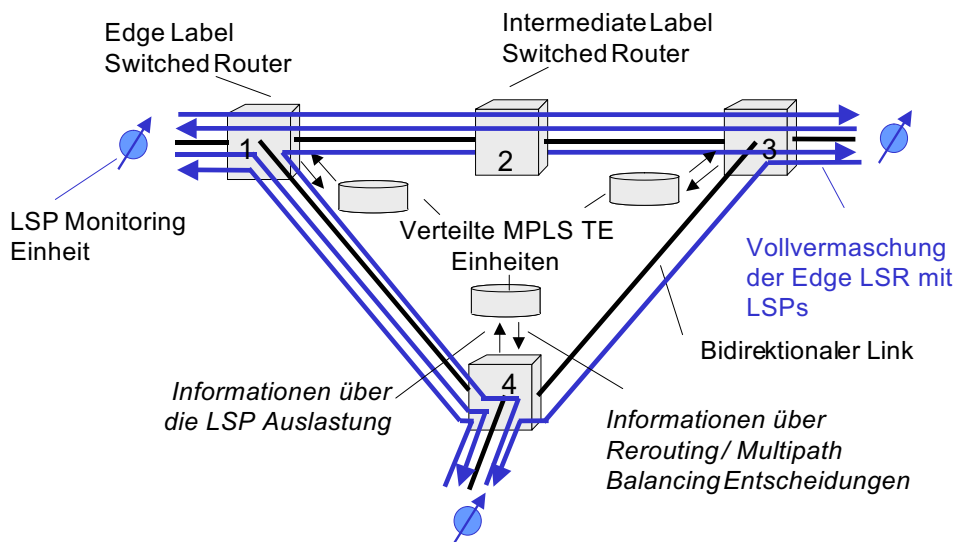


Abbildung 13: Das verteilte reaktive MPLS Traffic Engineering System

Das System basiert auf einem MPLS-Netz, welches sich aus Edge LSRs und Intermediate LSRs zusammensetzt. Die LSRs sind über bidirektionale Links miteinander verbunden. Über die Edge-LSRs tritt der Datenverkehr in das Netz ein und wieder aus. Der Datenverkehr wird durch direkt angeschlossene Teilnehmer, durch Server oder durch die angebundenen Nachbar-Netze erzeugt (siehe hierzu auch Abschnitt 2.1.1). Ein Edge LSR ist sowohl ein Eingangs- als auch ein Ausgangs-LSR. Intermediate LSRs transportieren den Datenverkehr zwischen den Edge-LSRs.

Zwischen benachbarten LSRs ist ein Signalisierungskanal mit ausreichend großer Bandbreite und hoher Übertragungspriorität reserviert. Hierüber tauschen die LSRs ihre Signalisierungs-Nachrichten aus.

Die Edge LSRs sind über LSPs vollvermascht. Die LSPs sind unidirektional und verbinden zwei Edge LSRs in einer Senderichtung. Bei dem Pfadumlegungsansatz existiert ein LSP pro Edge LSR Paar und Senderichtung und im Mehrpfadansatz sind es zwei oder mehr LSPs pro Edge LSR Paar und Senderichtung.

Jeder Edge-LSR ist mit einer LSP-Monitoring-Einheit verbunden, welche die mittlere Eingangslast der abgehenden LSP in periodischen Zeitabständen bestimmt. Die LSP-Auslastungs-Informationen flutet ein Edge-LSR an alle anderen Edge-LSRs. Aus den LSP-Auslastungs-Informationen und der Kenntnis der Pfade der LSPs berechnen die Edge-LSRs die Netzauslastung. Diese Informationen dienen zur Bestimmung der Lastumverteilung. Eine detaillierte Beschreibung des Monitoring-Systems erfolgt in Abschnitt 3.2.

Die TE-Einheiten befinden sich ebenfalls an den Edge-LSRs. Die TE-Einheiten führen die Berechnung der Lastumverteilung auf Basis der gemessenen Auslastungs-Informationen aus. Im Pfadumlegungsansatz bestimmen diese einen neuen Pfad für die jeweiligen LSPs und im Mehrpfadansatz berechnen sie die Gewichte, mit denen die Last auf die LSPs aufgeteilt wird. Anschließend triggern die TE-Einheiten die Edge-LSRs zur Durchführung der Lastverteilungsaktionen.

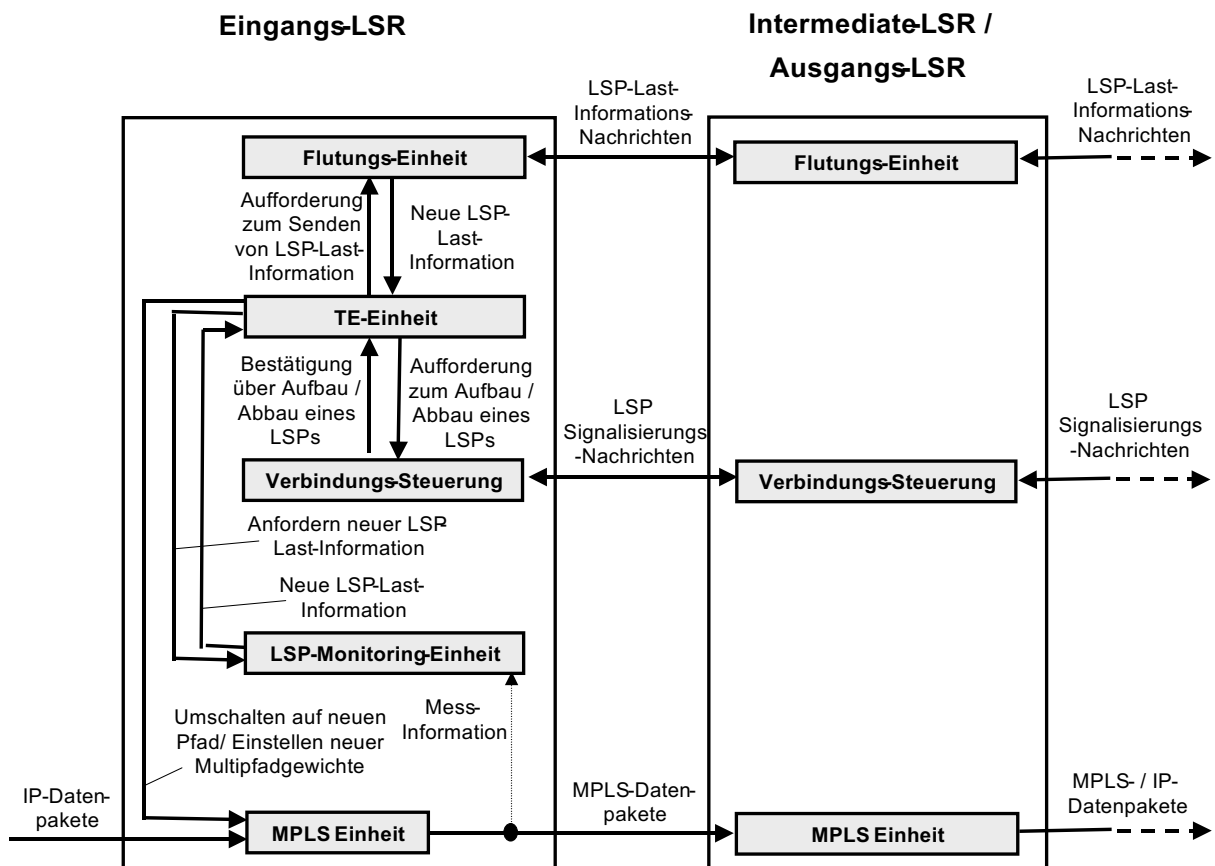


Abbildung 14: Funktionsblöcke der LSRs mit integrierter Monitoring- und TE-Einheit

Den Aufbau der LSRs mit integrierter Monitoring- und TE-Einheit stellt Abbildung 14 dar. Die Abbildung zeigt die Funktionsblöcke und deren Zusammenspiel innerhalb ei-

nes LSRs. Die Hauptfunktionalität des Systems ist in den Eingangs-LSRs vereint. Hier ist zu beachten, dass ein LSR gleichzeitig sowohl Eingangs-LSR, Intermediate LSR und Ausgangs-LSR jeweils für bestimmte LSPs sein kann.

An einem Eingangs-LSR verarbeitet die MPLS-Einheit die ankommenden IP-Datenpakete. Die MPLS-Einheit wertet die Zieladresse des IP-Datenpakets aus und bestimmt den LSP bzw. die LSPs, die das IP-Paket zum entsprechenden Ausgangs-LSR durch das MPLS-Netz übertragen. Im Pfadumlegungsansatz existiert ein LSP zwischen einem Eingangs-LSR und einem Ausgangs-LSR, so dass die Auswahl des LSPs eindeutig ist.

Bei dem Mehrpfadansatz stehen zwei oder mehrere LSPs für die Übertragung zur Auswahl. Um eine Reihenfolge der Pakete innerhalb eines Verkehrsflusses zu gewährleisten, verwendet dieses Verfahren das Hashing-Verfahren, welches in [EJS01] für das MATE-Verfahren vorgeschlagen wird (siehe auch Abschnitt 2.5.3.1).

Die MPLS-Einheit im Eingangs-LSR ergänzt das IP-Datenpaket um die MPLS-Kontroll-Informationen und sendet das Paket zum nächsten LSR. Die Intermediate LSRs schicken das MPLS-Datenpaket auf Basis des MPLS-Labels weiter bis zum Ausgangs-LSR, welcher schließlich die MPLS-Kontroll-Informationen entfernt und das Datenpaket als IP-Datenpaket weiterleitet.

Die Monitoring-Einheit im Eingangs-LSR misst die Last auf den abgehenden LSPs. Die aktuellen LSP-Auslastungen werden von der Traffic-Engineering-Einheit abgefragt und über die Flutungs-Einheit an alle Eingangs-LSR im Netz verteilt. Die Flutungs-Einheit verwendet den Mechanismus, welchen die Routing-Protokolle OSPF und ISIS verwenden (siehe Abschnitt 2.1.1.2). Zur Anwendung dieser Protokolle in diesem System müssten sowohl OSPF als auch ISIS um die Fähigkeit erweitert werden, die Auslastung einzelner LSP zu übertragen. Die Flutungs-Einheiten der Intermediate- und Ausgangs-LSRs senden die LSP-Auslastungs-Informationen weiter, ohne diese auszuwerten.

Erhält die TE-Einheit eines Eingangs-LSRs neue LSP-Auslastungs-Informationen, wertet diese die neue Auslastungssituation aus und berechnet im Fall einer Überlast die Lastumverteilung. Neue Auslastungs-Informationen liefert entweder die eigene Monitoring-Einheit oder die Flutungs-Einheit.

Im Pfadumlegungsansatz triggert die TE-Einheit die Verbindungssteuerung, um für den ausgewählten LSP einen neuen Pfad aufzubauen. Die neu berechnete Pfadinformation übergibt die TE-Einheit an die Verbindungssteuerung. Die Verbindungssteuerung können die in Abschnitt 2.1.2.3 beschriebenen Label Distribution Protokolle CR-LDP oder RSVP-TE realisieren. Nach erfolgreichem Aufbau des neuen Pfades benachrichtigt die Verbindungssteuerung die TE-Einheit. In gleicher Weise steuert die TE-Einheit den Abbau eines LSPs, nachdem die Last von dem alten auf den neuen LSP umgeschaltet wurde. Sowohl der Verbindungsaufbau als auch der Verbindungsabbau unter Verwendung von RSVP-TE bzw. CR-LDP beschreibt Abschnitt 2.1.2.3.

Um die Last von dem alten auf den neuen LSP umzuschalten, triggert die TE-Einheit die MPLS-Einheit. Diese ändert die Eintragungen in der Label Information Tabelle und ordnet die entsprechenden IP-Adressen dem neuen LSP zu.

Im Mehrpfadansatz übergibt die TE-Einheit die neu bestimmten Verteilgewichte für den ausgewählten LSP-Fächer an die MPLS-Einheit. Diese stellt die Verteilgewichte unter Verwendung des in [EJS01] beschriebenen Hashing-Verfahrens ein.

3.2 Das LSP-Monitoring-System

Das LSP-Monitoring-System beruht auf der passiven Messung der Auslastung jedes LSPs. Die Messungen führen die Mess-Einheiten an den Eingangs-LSRs durch, an welchen die LSPs beginnen. Hierzu zählt eine Mess-Einheit über die Dauer eines Mess-Intervalls die Anzahl der übertragenen Bytes auf den einzelnen LSPs.

Die TE-Einheit fordert in periodischen Zeitabständen die aktuelle Messung bei der Mess-Einheit an. Allerdings existieren auch Ausnahmen, in der die TE-Einheit von der periodischen Abfrage abweicht. Diese Ausnahmen beschreibt Abschnitt 3.3. Die Mess-Einheit bestimmt die mittlere Auslastung eines LSPs über dem Zeitraum zwischen der letzten Abfrage und dem aktuellen Zeitpunkt.

Jeder Eingangs-LSR kennt den Pfad von jedem LSP im Netz. Diese Information erhalten die Eingangs-LSRs während der Initialisierungsphase des Netzes. Im Pfadumlegungsansatz werden zusätzlich nach jeder Pfadumlegung eines LSPs die Pfadinformationen aktualisiert und über den Flutungs-Mechanismus an alle Eingangs-LSRs verteilt.

Basierend auf der Kenntnis der mittleren Auslastung und der Pfade der LSPs kann jeder Eingangs-LSR die mittlere Auslastung der Links im Netz berechnen. Dieser Ansatz basiert auf der Voraussetzung, dass der gesamte Verkehr im Netz über die LSPs übertragen wird.

Abbildung 15 zeigt ein Beispiel für die Berechnung der Link-Auslastungen für die Topologie aus Abbildung 13. Da jedem Eingangs-LSR die gleichen Informationen zu Grunde liegen, hat jeder Eingangs-LSR die gleiche Sicht der Netzauslastung.

	LSP 1	LSP 2	LSP 3	LSP 4	LSP 5	LSP 6	Link Auslastung
Link 1-2	13.4%	-	-	-	-	15.6%	29.0%
Link 2-1	-	12.5%	-	-	-	-	12.5%
Link 2-3	13.4%	-	-	-	-	15.6%	29.0%
Link 3-2	-	12.5%	-	-	-	-	12.5%
Link 1-4	-	-	20.5%	-	-	15.6%	36.1%
Link 4-1	-	-	-	24.5%	-	-	24.5%
Link 3-4	-	-	-	-	12.6%	-	12.6%
Link 4-3	-	-	-	-	-	-	0.0%

Abbildung 15: Berechnung der Link-Auslastungen aus Pfad- und Auslastungsinformation

3.3 Der Lastverteilungsprozess

Der Lastverteilungsprozess kann in einem sequentiellen und einem parallelen Durchführungsmodus betrieben werden. Zum Lastverteilungsprozess gehören die Auslösung des Lastverteilungsprozesses, die Berechnung der neuen Lastverteilung, die Durchführung der Lastumverteilung sowie die Koordinierung der Eingangs-LSRs.

Zunächst gibt Abschnitt 3.3.1 einen Überblick über den Ablauf eines Lastverteilungsprozesses. Anschließend geht Abschnitt 3.3.2 auf die Definition einer Überlastsituation ein. Das Vorgehen zur Berechnung der neuen Lastverteilung beschreibt Abschnitt 3.3.3. Abschnitt 3.3.4 geht auf die zeitlichen Bezüge innerhalb eines Lastverteilungsprozesses ein und erläutert die Koordinierung der Eingangs-LSRs.

3.3.1 Ablauf des Lastumverteilungsprozesses

Am Ende einer Messperiode flutet ein Eingangs-LSR die neu bestimmten mittleren Auslastungen der abgehenden LSPs an alle anderen Eingangs-LSRs. Die TE-Einheiten der Eingangs-LSRs berechnen aus diesen Informationen die aktuelle Netzauslastung. Stellen die TE-Einheiten bei der neu berechneten Netzauslastung eine Überlast im Netz fest, stoßen diese einen Lastverteilungsprozess an. Die Definition einer Überlast gibt Abschnitt 3.3.2.

Hierzu fordern alle TE-Einheiten zunächst die aktuellen Auslastungen der von ihrem Eingangs-LSR abgehenden LSPs bei den LSP-Monitoring-Einheiten an. Das sind die in 3.2 beschriebenen Fälle, bei denen abweichend von den periodischen Messintervallen die LSP-Auslastungen bestimmt werden. Die TE-Einheiten triggern die Flutungs-Einheiten, die aktuell bestimmten LSP-Auslastungen an alle Eingangs-LSRs zu verteilen. Nachdem die TE-Einheiten sämtliche LSP-Auslastungen empfangen und erneut die aktuelle Netzauslastung bestimmt haben, prüfen diese das Netz ein weiteres Mal auf Überlastsituationen.

Ist das Netz weiterhin im überlasteten Zustand, beginnen die TE-Einheiten die Neuberechnung der Lastverteilung. Hierzu wählen die TE-Einheiten einen LSP aus. Bei dem Pfadumlegungsansatz bestimmen die TE-Einheiten einen neuen Pfad für den LSP. Bei dem Mehrpfadansatz legen die TE-Einheiten neue Gewichte zur Verteilung der Last über den LSP-Fächer fest. Da alle TE-Einheiten die gleiche Netzsicht haben und den gleichen Algorithmus zur Auswahl eines LSP verwenden, wählen alle TE-Einheiten den gleichen LSP bzw. LSP-Fächer aus und berechnen die gleiche Lastumverteilung. Das ist der Hauptaspekt der Koordination zwischen den Eingangs-LSRs. Falls sich für den ausgewählten LSP bzw. LSP-Fächer kein besserer Pfad findet bzw. sich die Last nicht besser über den LSP-Fächer verteilen lässt, wählen die TE-Einheiten den nächsten LSP bzw. LSP-Fächer aus. Abschnitt 3.3.3 beschreibt den Algorithmus zur Festlegung der Reihenfolge der LSPs. Abschnitt 3.4 geht auf die verwendeten Lastverteilungs-Algorithmen des Mehrpfadansatzes und des Pfadumlegungsansatzes ein.

Der Eingangs-LSR des ausgewählten LSPs führt die Lastumverteilung durch. Bei dem Pfadumlegungsansatz steuert die TE-Einheit den Aufbau des neu berechneten Pfades, die Umschaltung der Last von dem alten auf den neuen Pfad und den Abbau des alten LSP. Bei dem Mehrpfadansatz steuert die TE-Einheit die Einstellung der neu berechneten Verteilgewichte.

Im **sequentiellen Durchführungsmodus des Lastumverteilungsprozesses** fordert die TE-Einheit die Flutungs-Einheit auf, die anderen Eingangs-LSRs über die erfolgreiche Durchführung der Lastumverteilung zu informieren. Hierzu sendet die Flutungs-Einheit im Pfadumlegungsansatz den neu berechneten Pfad an alle anderen Eingangs-LSRs. Bei dem Mehrpfadansatz übermittelt die Flutungs-Einheit den anderen Eingangs-LSRs die neuen Verteilgewichte des LSP-Fächers. Alle anderen Eingangs-LSRs warten die Information über die erfolgreiche Lastumverteilung ab. Nach Erhalt dieser Information bewerten

alle Eingangs-LSRs die neue Lastverteilung. Befindet sich das Netz weiterhin im überlasteten Zustand, wählen die Eingangs-LSRs den nächsten LSP aus und der Prozess wiederholt sich. Ist das Netz nicht mehr im überlasteten Zustand, endet der Lastverteilungsprozess.

Der sequentielle Durchführungsmodus des Lastverteilungsprozesses garantiert die Einhaltung der bei der Berechnung festgelegten Reihenfolge der LSP-Lastverteilungsaktionen.

Im **parallelen Durchführungsmodus des Lastverteilungsprozesses** gehen alle Eingangs-LSRs von der erfolgreichen Durchführung der berechneten Lastumverteilung des LSPs durch den aktiven Eingangs-LSR aus. Ohne das Abwarten einer Nachricht bestimmt jeder Eingangs-LSR die neue Netzauslastung unter Berücksichtigung der zuvor berechneten Lastumverteilung. Bestehen weiterhin Überlastsituationen im Netz, wählen die TE-Einheiten den nächsten LSP aus und berechnen die Lastumverteilung. Der Eingangs-LSR des LSPs führt diese dann durch. Dieser Schritt wiederholt sich, bis die TE-Einheiten keine Überlast im Netz mehr feststellen. Abschließend informieren die Eingangs-LSRs sich gegenseitig über sämtliche durchgeführte Lastumverteilungen.

Im parallelen Durchführungsmodus führen die Eingangs-LSRs die sequentiell berechneten Lastverteilungsaktionen parallel aus, was eine Verringerung der Konvergenzdauer zur Folge hat.

3.3.2 Definition einer Überlastsituation

Ein Netz wird als überlastet bezeichnet, wenn die Auslastung auf einem Link einen definierten Schwellwert überschreitet. Dieser Wert wird in der Folge als Überlastschwelle bezeichnet. Die Einführung einer Überlastschwelle beschränkt die Durchführung von Lastumverteilungen auf kritische Auslastungszustände, in denen die Übertragungsqualität beeinträchtigt ist oder die Gefahr der Beeinträchtigung besteht. Damit reduzieren die Verfahren die umzuschaltende Verkehrskapazität in Situationen geringer bis mittlerer Netzauslastung.

3.3.3 Vorgehen zur Berechnung der neuen Lastverteilung

Wie in Abschnitt 3.3.1 beschrieben, wählen die TE-Einheiten die LSPs sequentiell aus und berechnen für den ausgewählten LSP bzw. LSP-Fächer die Lastumverteilung. Dieser Abschnitt beschreibt den Algorithmus, welcher die Reihenfolge der LSPs festlegt. Der Algorithmus beruht dabei auf der Kenntnis der Pfade und der gemessenen mittleren Senderaten der LSPs.

Das Ziel des Algorithmus zur Festlegung der Reihenfolge der LSPs ist eine möglichst schnelle Auflösung einer Überlastsituation. Dabei berücksichtigt der Algorithmus zwei Nebenbedingungen. Die erste fordert eine möglichst geringe Anzahl an Lastverteilungsaktionen, wodurch die Konvergenzdauer und der Berechnungsaufwand minimiert werden. Die zweite zielt auf die Verringerung der Wahrscheinlichkeit von zukünftigen Überlastsituationen. Hierzu sollte der Algorithmus die LSPs unter Berücksichtigung der Minimierung des gesamten Ressourcen-Bedarfs auswählen.

Der in dem Verkehrslenkungsverfahren verwendete Algorithmus wählt einen LSP nach den in den folgenden vier Punkten definierten Kriterien aus. Nach der Berechnung bzw.

Durchführung der Lastumverteilung des ausgewählten LSPs erfolgt die Auswahl des nächsten LSPs. Dieser Prozess wird so lange wiederholt, bis entweder kein Link den Überlast-Schwellwert überschreitet oder jeder LSP auf den überlasteten Links einmal ausgewählt wurde.

1. **Auswahl des Links mit der größten relativen Auslastung:** Im ersten Schritt wählt der Algorithmus den Link mit der größten relativen Auslastung. Haben zwei Links die gleiche relative Auslastung, wird der mit der größeren Übertragungskapazität verwendet. Haben zwei Links mit gleicher Übertragungskapazität die gleiche relative Auslastung, wird der mit der höheren Link-ID verwendet.
2. **Auswahl des LSPs mit der größten gemessenen mittleren Übertragungsrates, der nicht über den kürzesten Pfad geroutet ist:** Im zweiten Schritt wählt der Algorithmus einen der LSPs aus, welcher über den ausgewählten Link geroutet wird. Hierzu betrachtet der Algorithmus zunächst die LSPs, welche nicht über den kürzesten Pfad (Pfad mit dem geringsten Hop Count) geroutet sind. Bei der Auswahl eines solchen LSPs ist die Wahrscheinlichkeit groß, dass die Last auf den kürzesten Pfad verteilt wird und somit der gesamte Ressourcen-Bedarf im Netz reduziert wird.

Von den beschriebenen LSPs wählt der Algorithmus den LSP mit der größten gemessenen mittleren Übertragungsrates aus. Damit wird im Erfolgsfall die größte Reduktion der mittleren Auslastung des überlasteten Links erzielt. Der positive Nebeneffekt hierbei ist, dass eine Umverteilung einer großen Last auf den kürzesten Pfad auch eine große Reduktion des gesamten Ressourcen-Bedarfs bedeutet.

3. **Auswahl eines LSPs, wenn alle LSPs über den kürzesten Pfad geroutet sind:** In dem Fall, dass alle LSPs über den kürzesten Pfad geroutet sind, unterscheidet sich der Algorithmus für den Pfadumlegungs- und den Mehrpfadansatz.

Im Pfadumlegungsansatz wählt der Algorithmus den LSP aus, ohne dessen mittlere Übertragungsrates die Auslastung des Links mit geringstem Abstand unterhalb der definierten Überlastschwelle liegt. Damit schaltet der Algorithmus eine möglichst geringe Verkehrskapazität, um die Überlastsituation aufzulösen. Falls die mittlere Auslastung von keinem LSP größer als die Differenz zwischen der aktuellen Link-Auslastung und der Überlastschwelle ist, wählt der Algorithmus den LSP mit der größten mittleren Übertragungsrates aus. Der Algorithmus zur Berechnung eines neuen Pfades beschreibt Abschnitt 3.4.

Im Gegensatz zum Pfadumlegungsansatz wählt der Algorithmus im Mehrpfadansatz immer den LSP mit der größten gemessenen mittleren Übertragungsrates aus. Somit ist die Wahrscheinlichkeit am größten, dass der Mehrpfadansatz einen möglichst großen Teil auf andere LSPs des LSP-Fächers verteilt und so die Überlast möglichst schnell auflöst. Würde wie im Pfadumlegungsansatz ein LSP mit einer mittleren Auslastung gewählt, welche nur geringfügig größer als die Differenz zwischen der aktuellen Auslastung und der Überlastschwelle ist, wird die Überlastsituation nur dann aufgelöst, wenn die gesamte Last auf die anderen LSPs des LSP-Fächers verteilt wird. Der Algorithmus zur Berechnung der neuen Verteilgewichte beschreibt Abschnitt 3.4.

4. **Begrenzung der Auswahl auf die LSPs bzw. LSP-Fächer, welche noch nicht ausgewählt wurden:** Der Algorithmus wählt jeden LSP bzw. LSP-Fächer

pro Lastverteilungsprozess maximal einmal aus. Zum einen verhindert diese Maßnahme ein mehrfaches Umschalten der Last eines LSPs bzw. LSP-Fächers während eines Lastverteilungsprozesses. Zum anderen begrenzt das die maximale Anzahl der Lastverteilungsaktionen pro Lastverteilungsprozess.

3.3.4 Zeitliche Koordination des Lastverteilungsprozesses

Dieser Abschnitt beschreibt den zeitlichen Ablauf eines Lastverteilungsprozesses. Dieser basiert auf der Annahme, dass keine Synchronisation der Start- und Endzeitpunkte der Mess-Intervalle der verschiedenen LSRs vorliegt. Der Lastverteilungsprozess basiert auf der Verwendung mehrerer Timer, welche die Koordination der Lastverteilungsaktionen der verschiedenen LSRs garantieren.

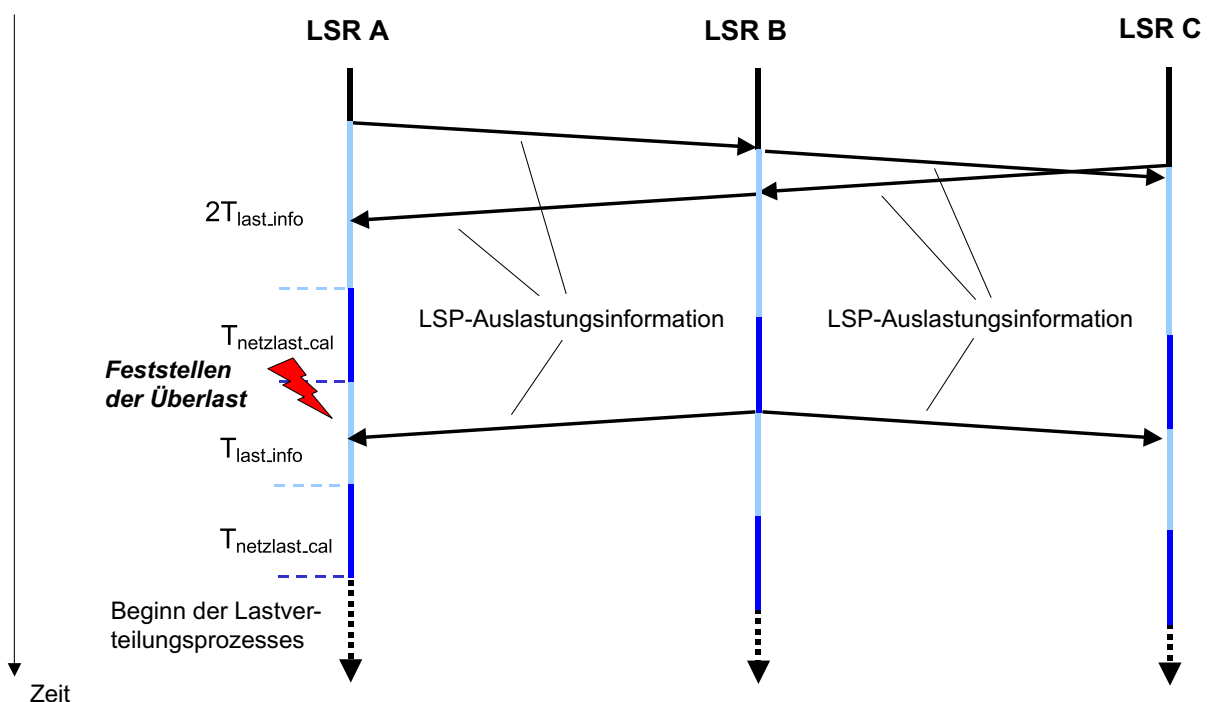


Abbildung 16: Aktualisierung der Auslastung bei Erkennung einer Überlast

Abbildung 16 zeigt das Nachrichtenflussdiagramm des Prozesses zur Aktualisierung der LSP-Auslastungs-Informationen im Fall einer erkannten Überlastsituation. Eingangs-LSR A sendet am Ende eines Messintervalls die aktualisierten mittleren Auslastungen der LSPs an alle anderen Eingangs-LSRs. Nach dem Senden bzw. nach dem Empfangen einer LSP-Auslastungs-Information wartet jeder Eingangs-LSR ohne eine weitere Aktion zunächst die zweifache Dauer $T_{last-info}$ ab. $T_{last-info}$ entspricht der maximalen Übertragungsdauer einer LSP-Auslastungs-Information zwischen zwei Netzknoten. Durch das Abwarten der zweifachen Dauer $T_{last-info}$ empfängt jeder LSR zufällig gleichzeitig gesendete bzw. kurzzeitig später gesendete LSP-Auslastungs-Informationen vor der Berechnung der Netzauslastung. Das garantiert jedem LSR bei der Berechnung die gleichen LSP-Auslastungs-Informationen und somit die gleiche Sicht der Netzauslastung. In dem Beispiel in Abbildung 16 sendet LSR C seine aktuellen LSP-Auslastungs-Informationen kurz vor Erhalt der LSP-Auslastungs-Information von LSR A. Ohne das Abwarten des Zeitintervalls hat LSR C zum Zeitpunkt des Erhalts der LSP-Auslastungs-Information von LSR A eine andere Netzsicht als LSR A und LSR B.

Anschließend berechnet jeder Eingangs-LSR auf der Basis der aktuellen LSP-Auslastungs-Informationen die Netzauslastung. Die maximale Dauer der Berechnung wird mit $T_{netzlast-cal}$ bezeichnet. Durch Vergleich der berechneten Link-Auslastungen mit der Überlastschwelle prüft jeder LSR das Netz auf Überlast. Im Fall einer Überlast fluten alle weiteren Eingangs-LSRs im Netz die aktuellen LSP-Auslastungs-Informationen an die anderen Eingangs-LSRs. Vor ihrer Auswertung wartet jeder Eingangs-LSR erneut die Zeit $T_{last-info}$ ab, um konsistente LSP-Auslastungs-Informationen an allen Eingangs-LSRs zu garantieren.

Nach Ablauf der Zeit $T_{last-info}$ aktualisiert jeder LSRs die Sicht der Netzauslastung und überprüft das Netz erneut auf Überlast. Erkennen die Eingangs-LSRs das Netz nach der aktualisierten LSP-Auslastungs-Information nach wie vor als überlastet, starten die Eingangs-LSRs die Durchführung der Lastumverteilung. In der Folge wird zunächst die sequentielle und anschließend die parallele Durchführung des Lastverteilungsprozesses beschrieben.

Sequentielle Durchführung der Lastumverteilung: Das Nachrichtenflussdiagramm der sequentiellen Durchführung der Lastumverteilung der LSPs ist in Abbildung 17 dargestellt. Basierend auf der aktuellen Netzauslastung und den aktuellen LSP-Auslastungs-Informationen wählt jeder Eingangs-LSR mit dem in Abschnitt 3.3.3 beschriebenen Algorithmus einen LSP aus und führt die Berechnung zur Lastumverteilung durch. Falls sich die Last von diesem LSP nicht umverteilen lässt, wählen die LSRs den nächsten LSP aus. Die maximale Zeitdauer zur Berechnung der Lastumverteilung gibt der Wert $T_{lastverteilung-cal}$ an.

Anschließend wartet der Eingangs-LSR des ausgewählten LSPs die Zeit $T_{last-info}$ ab, um zu garantieren, dass jeder Eingangs-LSR die Berechnung abgeschlossen hat und auf die Lastumverteilung des LSP wartet. Im Pfadumlegungsansatz stößt der Eingangs-LSR des ausgewählten LSPs den Aufbau des neuen Pfades an und wartet auf die Nachricht zur Bestätigung des LSP-Verbindungsaufbaus. Hierdurch ergibt sich für jeden LSR die zusätzliche Wartedauer $T_{lsp-sig}$, was der maximalen Dauer eines LSP-Verbindungsaufbaus entspricht. Anschließend schaltet der Eingangs-LSR des ausgewählten LSP die Last vom alten auf den neuen Pfad um und stößt den Abbau des alten LSP an. Im Mehrpfadansatz schaltet der Eingangs-LSR die Last des LSP-Fächers entsprechend der neu berechneten Verteilgewichte direkt um. Der Umschaltvorgang eines LSPs bei dem Mehrpfadansatz ist somit um die Zeitdauer $T_{lsp-sig}$ kürzer.

Nach dem Umschaltvorgang flutet der Eingangs-LSR bei beiden Ansätzen eine Nachricht mit der neuen LSP-Auslastungs-Information an alle anderen Eingangs-LSRs. Im Pfadumlegungsansatz teilt der Eingangs-LSR den neuen Pfad mit und im Mehrpfadansatz flutet der LSR die neuen Verteilgewichte. Anschließend wartet er erneut die Dauer $T_{last-info}$ ab, um zu gewährleisten, dass alle LSRs konsistente Auslastungs-Informationen haben. Basierend auf diesen berechnet jeder Eingangs-LSR die aktuelle Netzauslastung und überprüft die Links auf Überlast. Falls das Netz weiterhin überlastet ist, wiederholt sich die sequentielle Durchführung der Lastverteilung erneut.

Der Lastverteilungsprozess endet, wenn entweder kein überlasteter Link mehr im Netz vorhanden ist oder wenn die Lastverteilung von jedem LSP bzw. jedem LSP-Fächer einmal angepasst wurde.

Parallele Durchführung der Lastumverteilung: Abbildung 18 stellt das Nachrichtenflussdiagramm für die parallele Durchführung der Lastverteilung dar. Auf der aktualisierten Netzsicht wählt jeder Eingangs-LSR gemäß dem in Abschnitt 3.3.3 beschriebenen

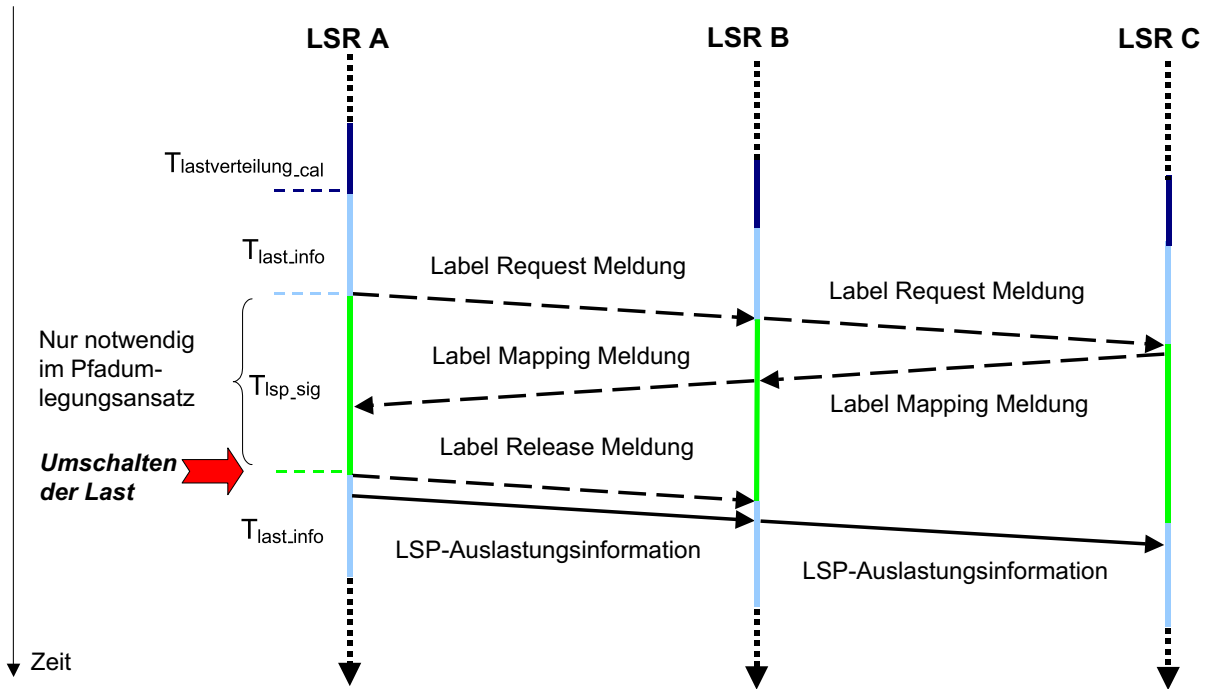


Abbildung 17: Sequentielle Durchführung der Lastumverteilungsaktionen

Algorithmus einen LSP aus, berechnet die Lastumverteilung und passt die eigene Sicht der Netzauslastung entsprechend an. Diesen Prozess wiederholt jeder Eingangs-LSR, wenn sich nach der Lastumverteilung das Netz weiterhin in einem überlasteten Zustand befindet. Die Eingangs-LSRs der ausgewählten LSPs führen die Lastumverteilungsaktionen durch.

$$T_{lastverteilung-par-cal} = N_{lsp} \cdot (T_{lastverteilung-cal} + T_{netzlast-cal}) \quad (9)$$

Um sämtliche notwendigen Lastumverteilungen durchführen zu können, wartet jeder LSR die in Gleichung 9 definierte Zeit $T_{lastverteilung-par-cal}$ ab, bevor er die neuen LSP-Auslastungs-Informationen überträgt. $T_{lastverteilung-par-cal}$ entspricht der Multiplikation aus der Anzahl N_{lsp} der LSPs bzw. LSP-Fächer im Netz und der Summe der maximalen Dauer zur Berechnung einer Lastumverteilung $T_{lastverteilung-cal}$ und der Dauer zur Neuberechnung der Netzauslastung $T_{netzlast-cal}$. Im Pfadumlegungsansatz verlängert sich die Dauer bis zum Senden der LSP-Auslastungs-Information um die Zeit $T_{lsp-sig}$. Damit warten die LSRs zusätzlich die Dauer des Verbindungsaufbaus ab. Anschließend senden die LSRs LSP-Auslastungs-Informationen über alle durchgeführten Lastverteilungsaktionen. Nach der Dauer $T_{last-info}$ haben alle Eingangs-LSRs alle LSP-Auslastungs-Informationen erhalten. Damit endet der Lastverteilungsprozess für die parallele Durchführung.

3.4 Lastverteilungsalgorithmen

Dieser Abschnitt geht auf die Algorithmen ein, welche der Pfadumlegungsansatz und der Mehrpfadansatz zur Lastverteilung einsetzen. Der Pfadumlegungsansatz verwendet das Routing im Codex-Netz (siehe Abschnitt 2.5.2.2) und wird in Abschnitt 3.4.2 beschrieben. Der Mehrpfadansatz basiert auf dem in Abschnitt 2.5.2.3 erläuterten optimalen Routing

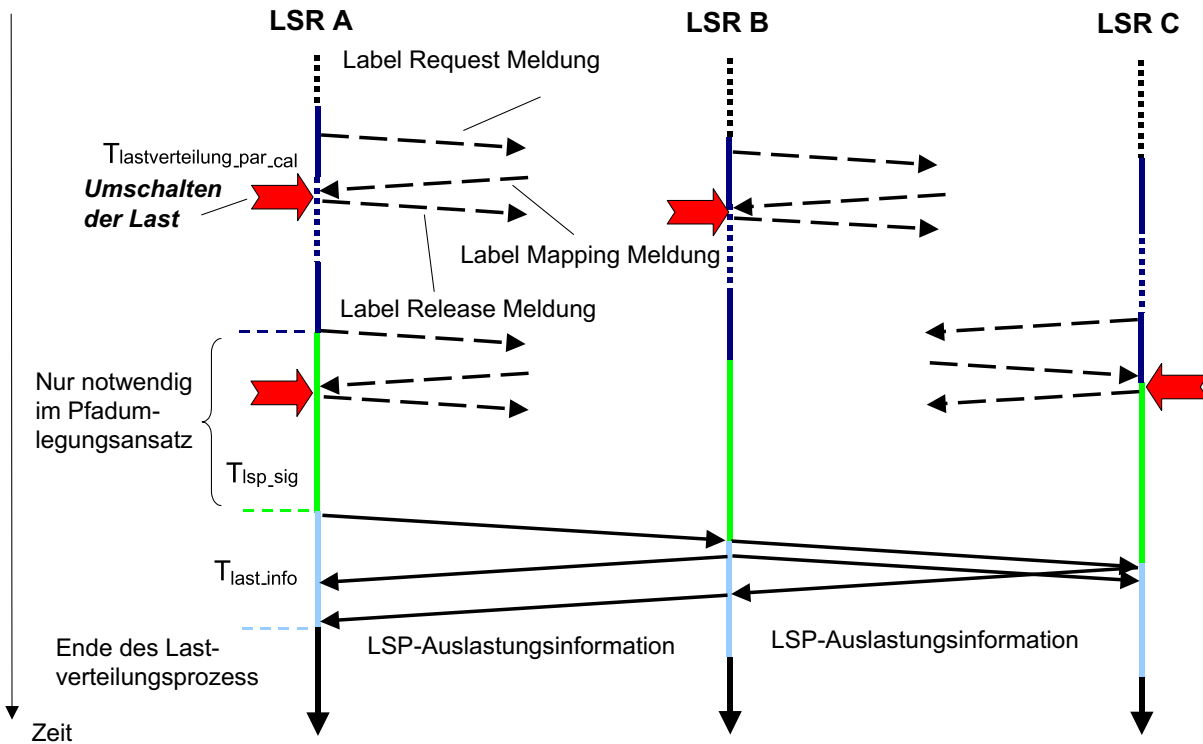


Abbildung 18: Parallele Durchführung der Lastumverteilungsaktionen

von Bertsekas und Gallager. Abschnitt 3.4.3 gibt einen Überblick über die Umsetzung des Mehrpfad-Algorithmus. Beide Algorithmen verwenden als Zielfunktion die Minimierung der gesamten Linkkosten. Die verwendete Linkkosten-Funktion beschreibt der folgende Abschnitt 3.4.1.

3.4.1 Linkkosten-Funktion der beiden Lastverteilungsalgorithmen

Die Linkkosten-Funktion definiert die Linkkosten in Abhängigkeit der Auslastung eines Links. Die verwendete Linkkosten-Funktion leitet sich aus den Verkehrsgüteparametern Paketverzögerung sowie Paketverlust eines Puffers ab. Die Funktion berücksichtigt zum einen die Paketverzögerung und den Verlust am Ausgangspuffer des sendenden Netzknotens. Diese Parameter werden mit einem M/M/1/S Wartesystem mit S Warteplätzen berechnet. Zum anderen fließt die Signallaufzeit auf dem Link in den Kostenterm der Paketverzögerung ein. Der Paketverlust auf dem Link wird vernachlässigt.

Die Linkkosten-Funktion C_{ges} beschreibt die Gleichung 10. Diese besteht aus den beiden Kostentermen des Paketverlusts C_{pvr} und der Paketverzögerung C_{pv} , welche über die Gewichte G_{pvr} und G_{pv} variabel zueinander gewichtet werden können.

$$C_{ges} = G_{pvr} \cdot C_{pvr} + G_{pv} \cdot C_{pv} \quad (10)$$

Über die Gewichte kann die Kostenfunktion zum einen an die Anforderungen des Verkehrs angepasst werden. Während für die Übertragungsgüte von Verkehrsflüssen mit Echtzeitanforderungen wie z.B. einer IP-Telefonie-Verbindung eine geringe Paketverzögerung das Hauptkriterium ist, ist für die Garantie eines hohen Durchsatzes von Datenübertragungen wie z.B. einer FTP-Verbindung ein geringer Paketverlust das Hauptziel.

Zum anderen kann die Kostenfunktion an das Netz angepasst werden. Je nach der Länge und Kapazität der Links hat die Auswahl eines Pfades für einen Verkehrsfluss einen unterschiedlichen Einfluss auf die Verkehrsparameter. In einem IP-Netz in Nordamerika beispielsweise beeinflusst die Auswahl eines Pfades abweichend vom kürzesten Pfad die Paketverzögerung stärker als in einem IP-Netz in Deutschland.

Der Kostenterm des Paketverlusts C_{pvr} entspricht einer auf die maximale Linkkapazität normierten Paketverlustrate. Der Kostenterm berücksichtigt, wie bereits erwähnt, ausschließlich den Paketverlust am Ausgangspuffer des Netzknotens. Die Paketverlustwahrscheinlichkeit P_{vw} eines M/M/1/S Wartesystems mit S Warteplätzen, welches den Ausgangspuffer modelliert, zeigt Gleichung 11. Der Kostenterm des Paketverlusts ergibt sich durch die Normierung der Paketverlustwahrscheinlichkeit mit der mittleren Auslastung ρ des Puffers sowie dem reziproken Wert des Faktors R (siehe Gleichung 13). Der Faktor R normiert den Kostenterm eines Links mit Kapazität B_{link} auf die maximale Linkkapazität B_{max} im Netz (siehe Gleichung 12). Eine Minimierung des Kostenterms C_{pvr} bedeutet die Minimierung der gesamten Paketverlustrate im Netz.

$$P_{vw} = \frac{\rho^{S+1} \cdot (1 - \rho)}{1 - \rho^{S+2}} \quad (11)$$

$$R = \frac{B_{max}}{B_{link}} \quad (12)$$

$$C_{pvr} = P_{vw} \cdot \rho \cdot \frac{1}{R} = \frac{\rho^{S+2} \cdot (1 - \rho)}{1 - \rho^{S+2}} \cdot \frac{1}{R} \quad (13)$$

Der Kostenterm der Paketverzögerung C_{pv} entspricht der Summe der auf die maximale Linkkapazität normierten Verzögerung der Pakete pro Zeiteinheit und teilt sich in drei Komponenten auf (siehe Gleichung 14). Der Kostenterm C_{wd} leitet sich aus der mittleren Wartedauer in einem Ausgangspuffer ab, der Kostenterm C_{ad} berücksichtigt die Aussen-dauer des Pakets und der Kostenterm C_{sv} bezieht die Signalverzögerung des Pakets auf dem Link ein. Eine Minimierung des Kostenterms der Paketverzögerung minimiert die Summe der Paketverzögerungen aller Pakete im Netz.

$$C_{pv} = C_{wd} + C_{ad} + C_{sv} \quad (14)$$

Der Kostenterm der mittleren Wartedauer im Puffer C_{wd} leitet sich aus der mittleren Warteschlangenlänge Ω des M/M/1/S Wartesystems ab, welche Gleichung 15 beschreibt. Die Multiplikation der mittleren Warteschlangenlänge Ω mit der Auslastung ρ des Puffers normiert die Warteschlangenlänge auf die Pakete pro Zeiteinheit, die die Wartedauer im Puffer erfahren. Durch die Multiplikation der Warteschlangenlänge mit dem Term $(1 - P_{vw})$ berücksichtigt die Wartedauer ausschließlich die Pakete, welche an dem Puffer nicht verloren gehen. Der resultierende Kostenterm C_{wd} ist in Gleichung 16 dargestellt.

$$\Omega = \frac{\rho^2 \cdot (1 - \rho^S)}{1 - \rho^{S+2}} \cdot \left(\frac{1}{1 - \rho} - \frac{S \cdot \rho^S}{1 - \rho^S} \right) \quad (15)$$

$$C_{wd} = \Omega \cdot \rho \cdot (1 - P_{vw}) = \frac{\rho^3 - \rho^{S+3} - \rho^{S+4} + \rho^{2 \cdot S+4}}{(1 - \rho^{S+2})^2} \cdot \left(\frac{1}{1 - \rho} - \frac{S \cdot \rho^S}{1 - \rho^S} \right) \quad (16)$$

Der Kostenterm C_{ad} bezieht sich auf die Dauer des Aussendens eines Pakets. Die Normierung der Aussendedauer entspricht der Normierung der Wartedauer im Puffer. Den Kostenterm zeigt Gleichung 17.

$$C_{ad} = \rho \cdot (1 - P_{vw}) = \frac{\rho - \rho^{S+2}}{1 - \rho^{S+2}} \quad (17)$$

Die Signalverzögerung T_{sv} des Pakets auf dem Link bestimmt sich aus der Länge des Links $x_{länge}$ und der Signalausbreitungsgeschwindigkeit v_{sv} (siehe Gleichung 18). Die Signalverzögerung T_{sv} wird auf die Anzahl der Pakete mit einer Länge N_{paket} normiert, die über den Link mit der Kapazität B_{link} übertragen werden. Durch die Multiplikation mit der mittleren Auslastung ρ , dem reziproken Wert des Faktors R und dem Wert $(1 - P_{vw})$ erfolgt die Normierung der Signalverzögerung. Den zugehörigen Kostenterm C_{sv} zeigt Gleichung 19.

$$T_{sv} = \frac{x_{länge}}{v_{sp}} \quad (18)$$

$$C_{sv} = T_{sv} \cdot \frac{B_{link}}{N_{paket}} \cdot \rho \cdot \frac{1}{R} \cdot (1 - P_{vw}) = T_{sv} \cdot \frac{B_{link}}{N_{paket}} \cdot \frac{1}{R} \cdot \frac{\rho - \rho^{S+2}}{1 - \rho^{S+2}} \quad (19)$$

In der Abbildung 19 ist der Verlauf einer Kostenfunktion dargestellt. Die Gewichte sind auf die Werte $G_{pv} = 1$ und $G_{pww} = 1e07$ gesetzt, der Puffergröße entspricht 75 Paketen, die Paketgröße N_{paket} ist 576Byte, der Link hat eine Länge von 100km und eine Kapazität von $2500 \frac{Mbit}{s}$. Den logarithmischen Verlauf der Kostenfunktion zeigt Abbildung 20. Der Verlauf der Linkkosten-Funktion zeigt bis zu einer mittleren Auslastung von 0.8 einen annähernd linearen Anstieg der Linkkosten. Diesen Anstieg verursacht der Kostenterm der Paketverzögerung. Mit zunehmender Linkauslastung steigt die Anzahl der Pakete, welche alle annähernd die gleiche Paketverzögerung erfahren. Die Verzögerungen der Pakete werden zum Kostenterm der Paketverzögerung aufsummiert. Den steilen Anstieg der Linkkosten-Funktion im hohen Auslastungsbereich ruft der Kostenterm des Paketverlusts hervor. Die in den Abbildungen 21 und 22 ohne Berücksichtigung der Gewichte G_{pv} und G_{pww} dargestellten Kostenterme der Paketverzögerung und der Paketverlustwahrscheinlichkeit unterstreichen die Aussagen.

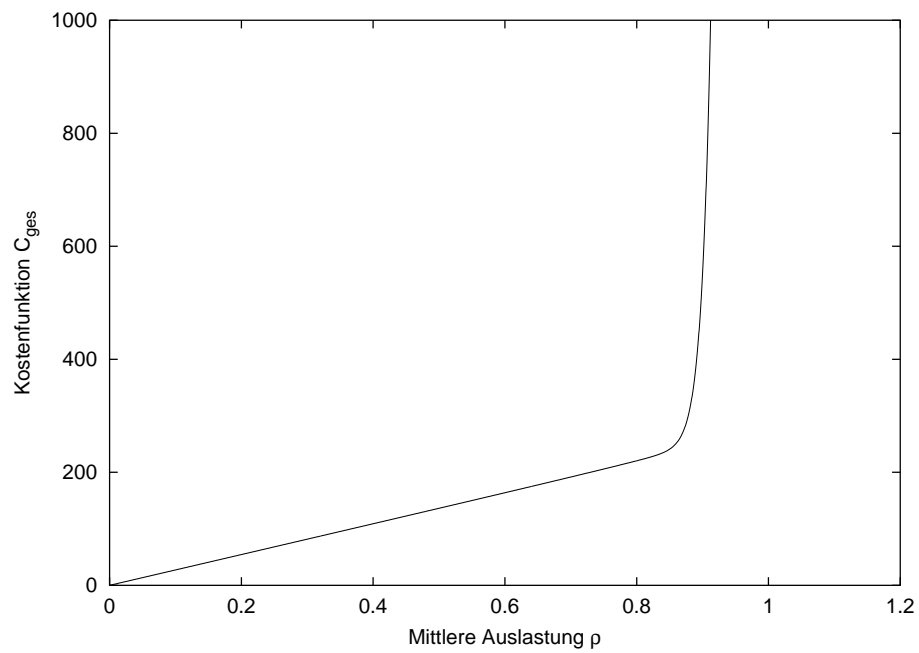


Abbildung 19: Darstellung der Linkkosten-Funktion

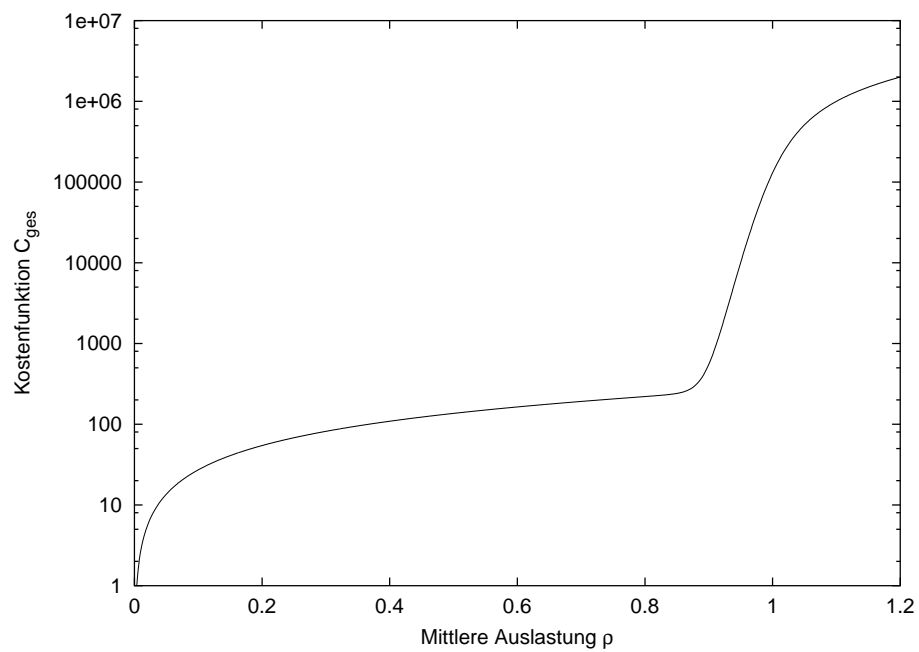


Abbildung 20: Logarithmische Darstellung der Linkkosten-Funktion

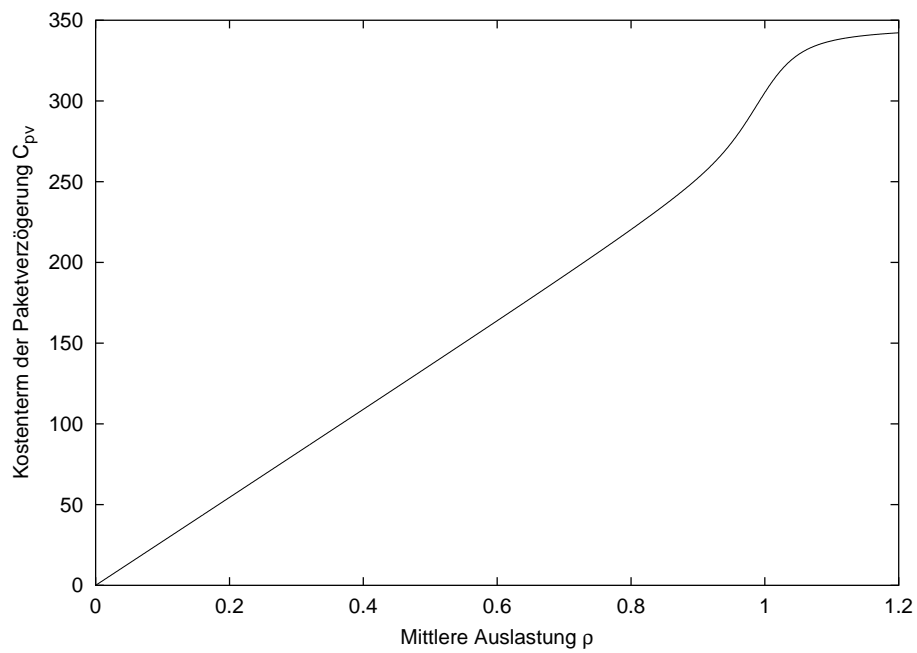


Abbildung 21: Darstellung des Kostenterms der Paketverzögerung

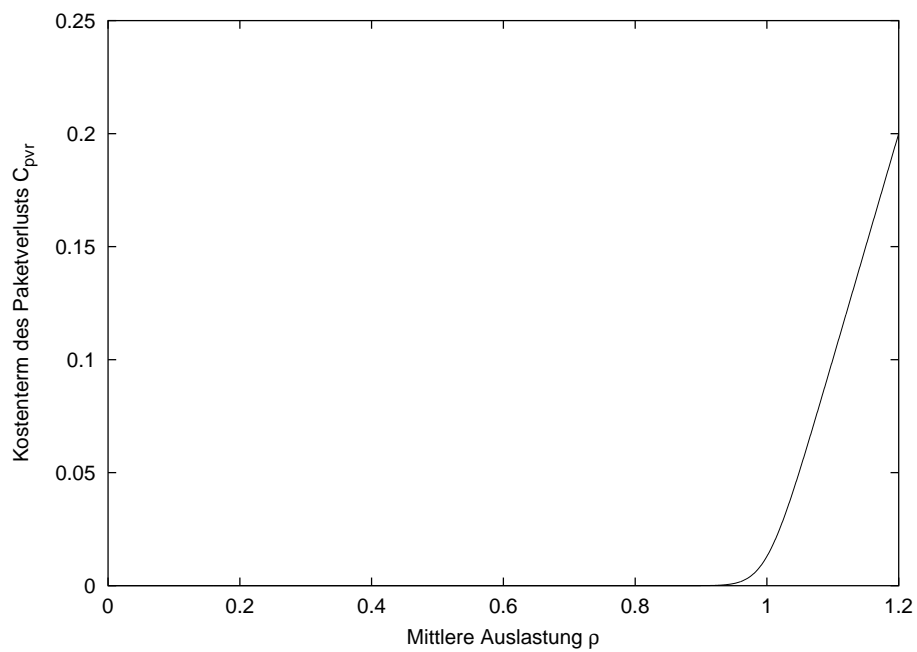


Abbildung 22: Darstellung des Kostenterms des Paketverlusts

Nachweis der Konvexität der Kostenfunktion: Eine Bedingung der Linkkostenfunktion ist ein konvexer Verlauf über der Linkauslastung. Das ist die Voraussetzung für die Anwendung der Lastverteilungsalgorithmen.

Nach der mathematischen Definition ist eine Funktion dann konvex, wenn die 1. Ableitung der Funktion monoton steigend ist [BS79]. Die Definition zeigt Gleichung 20.

$$\frac{\delta F(x_1)}{\delta x} \leq \frac{\delta F(x_2)}{\delta x} \quad \text{für} \quad x_1 < x_2 \quad (20)$$

Die Auswertung der ersten Ableitung der Kostenfunktion, deren Verlauf in Abbildung 23 dargestellt ist, weist die Konvexität der Funktion für den Auslastungsbereich zwischen 0 und 1 nach. Da die Kostenfunktion sich auf den Eingangsverkehr an einem Puffer bezieht, kann die Auslastung auch einen Wert größer 1 annehmen. In dem Auslastungsbereich größer 1 nähert die Funktion den Wert G_{pww} an.

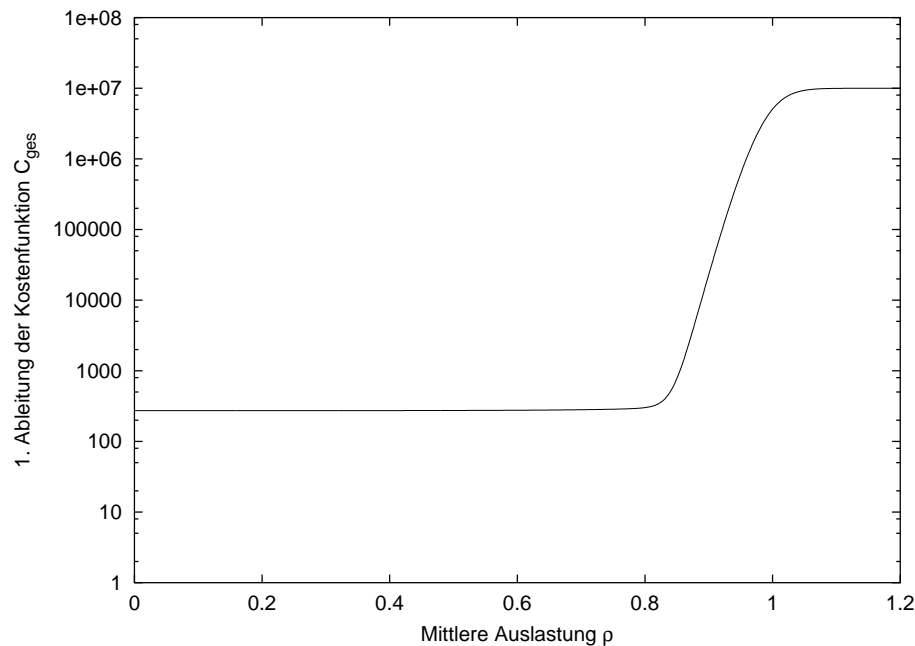


Abbildung 23: Logarithmische Darstellung der 1. Ableitung der Linkkosten-Funktion

3.4.2 Algorithmus des Pfadumlegungsansatzes

Das Ziel des Algorithmus des Pfadumlegungsansatzes ist die Bestimmung eines neuen Pfades für einen ausgewählten LSP. Die Auswahl eines LSPs beschreibt Abschnitt 3.3.3. Der verwendete Routing-Algorithmus bestimmt auf der Basis der aktuellen Sicht der Netzauslastung den Pfad, welcher die größte Reduktion der gesamten Linkkosten erzeugt. Die Realisierung des Routing-Algorithmus basiert auf der Idee des Routings im Codex-Netz, welcher in Abschnitt 2.5.2.2 erläutert wird.

Der Routing-Algorithmus bestimmt zunächst die Differenz der Linkkosten, die der Link mit und ohne Last L_{lsp} des ausgewählten LSPs hat. Basierend auf den Differenz-Linkkosten

bestimmt der Routing-Algorithmus unter Verwendung des Dijkstra-Algorithmus den kürzesten Pfad. Dieser Pfad entspricht dem Pfad, auf welchem der LSP die geringsten zusätzlichen Kosten erzeugt. Damit entspricht der Pfad ebenfalls dem Pfad, welcher die größte Reduktion der Summe der Linkkosten gegenüber dem alten Pfad erzielt.

Wenn der LSP bereits über den günstigsten Pfad geroutet ist, entspricht der neu berechnete Pfad dem alten Pfad. Wenn der Routing-Algorithmus einen abweichenden Pfad berechnet, der die gleichen zusätzlichen Linkkosten wie der alte Pfad erzeugt, wird als zweites Kriterium der Hop Count des Pfades in Betracht gezogen. Wenn der neue Pfad einen geringeren Hop Count als der alte Pfad hat, reduziert dieser die gesamte Netzauslastung L_{ges} (siehe Gleichung 21) und wird somit gegenüber dem alten Pfad bevorzugt.

Die gesamte Netzauslastung, welche Gleichung 21 zeigt, ergibt sich aus der Summe der Auslastungen aller Links L_i im Netz. E definiert die Menge aller Links. Die Auslastung eines Links L_i berechnet sich aus der Summe der mittleren Auslastungen der LSPs $L_{lsp j}$, welche über den Link i geroutet werden (siehe Gleichung 22). $N_{lsp i}$ ist die Menge aller LSPs des Links i .

$$L_{ges} = \sum_{i \in E} L(i) \quad (21)$$

$$L_i = \sum_{j \in N_{lsp i}} L_{lsp j} \quad (22)$$

Der Algorithmus des Pfadumlegungsansatzes läuft in der folgenden Weise ab:

1. Die Last L_{lsp} und der aktuelle Pfad des ausgewählten LSPs sowie die Auslastung aller Links L_i ohne die Last des ausgewählten LSPs sind dem ausführenden Eingangs-LSR durch den in Abschnitt 3.3.1 beschriebenen Austausch der LSP-Auslastungs-Information bekannt.
2. Für jeden Link i im Netz werden die Linkkosten C_i ohne Berücksichtigung der Last des LSPs und mit Berücksichtigung der Last des LSPs bestimmt.
3. Aus den beiden pro Link bestimmten Kosten werden die Differenzkosten $C_i(L_i + L_{lsp}) - C_i(L_i)$ für jeden Link berechnet.
4. Die Differenz der Linkkosten wird mit einem Faktor $F_{genauigkeit}$ multipliziert. Damit kann die Genauigkeit, mit der die Differenzkosten aufgelöst werden, variiert werden. Je größer $F_{genauigkeit}$, desto genauer werden die Differenzkosten aufgelöst.
5. Durch das Abrunden auf den nächsten ganzzahligen Wert werden die Differenzkosten ΔC_i für jeden Link bestimmt. Sind die Differenzkosten ΔC_i gleich null, werden diese auf den Wert eins gesetzt. Die Formeln zur Bestimmung der Linkkosten zeigen die Gleichungen 23 und 24.

$$\Delta C_i = \text{int}(F_{genauigkeit} \cdot (C_{ij}(L_{ij} + \Delta L) - C_{ij}(L_{ij}))) > 0 \quad (23)$$

$$\Delta C_i = 1 \quad \text{wenn} \quad \text{int}(F_{genauigkeit} \cdot (C_i(L_i + \Delta L) - C_i(L_i))) = 0 \quad (24)$$

6. Basierend auf den berechneten Differenzkosten ΔC_i wird unter Verwendung des Dijkstra-Algorithmus der Pfad berechnet, der die geringsten zusätzlichen Kosten erzielt.
7. Anschließend erfolgt eine Evaluierung des neu berechneten Pfades. In den beiden folgenden Fällen wird eine Pfadumlegung angestoßen:
 - (a) Der LSP erzeugt auf dem neuen Pfad geringere zusätzliche Linkkosten als auf dem alten Pfad.
 - (b) Der LSP erzeugt auf dem neuen Pfad die gleichen zusätzlichen Linkkosten wie auf dem alten Pfad und der neue Pfad hat einen geringeren Hop Count als der alte Pfad.

3.4.3 Algorithmus des Mehrpfadansatzes

Das Ziel des Algorithmus des Mehrpfadansatzes ist die Berechnung der Gewichte, mit denen die Last eines Eingangs-Ausgangs-Paars über den ausgewählten LSP-Fächer verteilt wird. Die Auswahl eines LSP-Fächers beschreibt Abschnitt 3.3.3. Der verwendete Mehrpfad-Algorithmus bestimmt auf der Basis der aktuellen Sicht der Netzauslastung die Lastverteilungsgewichte, welche die größte Reduktion der gesamten Linkkosten erzeugt. Die Realisierung des Mehrpfad-Algorithmus basiert auf dem Kuhn-Tucker-Theorem (siehe auch Gleichung 5). Dieses Theorem besagt, dass die Verteilung der Last über einen LSP-Fächer mit zwei oder mehreren LSPs dann die minimalen Gesamtkosten erreicht, wenn die Last sich ausschließlich auf die LSPs mit der geringsten ersten Ableitung der Pfadkosten verteilt.

Die erste Ableitung der Pfadkosten berechnet sich aus der Summe der ersten Ableitung der Kosten der Links, über den der Pfad geroutet ist (siehe auch Gleichung 5). Die erste Ableitung der Linkkosten-Funktion aus Abschnitt 3.4.1 beschreiben die Gleichungen 25 bis 34:

$$\frac{\delta C_{ges}}{\delta \rho} = G_{pvr} \cdot \frac{\delta C_{pvr}}{\delta \rho} + G_{pv} \cdot \frac{\delta C_{pv}}{\delta \rho} \quad (25)$$

$$\frac{\delta C_{pvr}}{\delta \rho} = \frac{(S+2) \cdot \rho^{S+1} - (S+3) \cdot \rho^{S+2} + \rho^{2 \cdot S+4}}{(1 - \rho^{S+2})^2} \quad (26)$$

$$\frac{\delta C_{pv}}{\delta \rho} = \frac{\delta C_{wd}}{\delta \rho} + \frac{\delta C_{ad}}{\delta \rho} + \frac{\delta C_{sv}}{\delta \rho} \quad (27)$$

$$\frac{\delta C_{wd}}{\delta \rho} = (K_1 + K_2) \cdot K_3 + K_4 \quad (28)$$

$$K_1 = \frac{3 \cdot \rho^2 - (S+3) \cdot \rho^{S+2} - (S+4) \cdot \rho^{S+3} + (2 \cdot S+1) \cdot \rho^{S+4}}{(1 - \rho^{S+2})^3} \quad (29)$$

$$K_2 = \frac{(2 \cdot S+4) \cdot \rho^{2 \cdot S+3} - (S+1) \cdot \rho^{2 \cdot S+4} - S \cdot \rho^{2 \cdot S+5}}{(1 - \rho^{S+2})^3} \quad (30)$$

$$K_3 = \left(\frac{1}{1-\rho} - \frac{S \cdot \rho^S}{1-\rho^S} \right) \quad (31)$$

$$K_4 = \frac{\rho^3 - \rho^{S+3} - \rho^{S+4} + \rho^{2 \cdot S+4}}{(1-\rho^{S+2})^2} \cdot \left(\frac{1}{(1-\rho)^2} - \frac{S^2 \cdot \rho^{S-1}}{(1-\rho^S)^2} \right) \quad (32)$$

$$\frac{\delta C_{ad}}{\delta \rho} = \frac{1 - (S+2) \cdot \rho^{S+1} + (S+1) \cdot \rho^{S+2}}{(1-\rho^{S+2})^2} \quad (33)$$

$$\frac{\delta C_{sv}}{\delta \rho} = \frac{1}{R} \cdot T_{sv} \cdot \frac{B_{max}}{N_{packet}} \cdot \frac{1 - (S+2) \cdot \rho^{S+1} + (S+1) \cdot \rho^{S+2}}{(1-\rho^{S+2})^2} \quad (34)$$

Zur Berechnung der optimalen Gewichte setzt der im Rahmen dieser Arbeit entwickelte Mehrpfadansatz den folgenden Algorithmus ein. Der Algorithmus teilt die Last des ausgewählten LSP-Fächers zunächst in Teile einer zuvor festgelegten Granularität auf. Anschließend bestimmt der Mehrpfad-Algorithmus den LSP mit der geringsten ersten Ableitung der Pfadkosten und ordnet diesem LSP den ersten Lastanteil zu. Diesen Schritt wiederholt der Mehrpfad-Algorithmus unter Berücksichtigung den zuvor verteilten Lastanteilen solange, bis die gesamte Last über die LSPs des ausgewählten LSP-Fächers verteilt ist. Der Mehrpfad-Algorithmus funktioniert für eine beliebige Anzahl an LSPs pro LSP-Fächer.

Der Algorithmus des Mehrpfadansatzes läuft in der folgenden Weise ab:

1. Die aktuellen Pfade und die Last $L_{lsp\ j}$ der LSPs des ausgewählten LSP-Fächers sowie die Auslastung L_i aller Links i ohne Berücksichtigung der Last des gewählten LSP-Fächers sind dem ausführenden Eingangs-LSR durch den in Abschnitt 3.3.1 beschriebenen Austausch der LSP-Auslastungs-Informationen bekannt.
2. Die Last des ausgewählten LSP-Fächers $L_{lsp-fächer}$ wird entsprechend der gewählten Granularität L_g in N Teile aufgeteilt.

$$N = \frac{L_{lsp-fächer}}{L_g} \quad (35)$$

3. Auf Basis der Linkauslastungen L_i und der bereits zugeordneten Teile der Last des LSP-Fächers $L_{lsp-fächer}$ werden die ersten Ableitungen der Pfadkosten der LSPs berechnet.
4. Die erste Ableitung der Pfadkosten wird mit dem Faktor $F_{genauigkeit}$ multipliziert, um so die Genauigkeit der Auflösung variabel einstellen zu können. Der vom Algorithmus verwendete Wert der ersten Ableitung der Pfadkosten $\frac{\delta C_{lsp\ j}}{\delta \rho}$ ergibt sich durch das Abrunden auf den nächsten ganzzahligen Wert. Diesen Wert zeigt Gleichung 36.

$$\frac{\delta C_{lsp\ j}}{\delta \rho} = int(F_{genauigkeit} \cdot \sum_{i \in P_{lsp\ j}} \frac{\delta C_i(L_i)}{\delta \rho}) \quad (36)$$

5. Im nächsten Schritt wird das Minimum der ersten Ableitungen der Pfadkosten $(\frac{\delta C_{isp}}{\delta \rho})_{min}$ der LSPs des LSP-Fächers ermittelt.
6. Jedem LSP, dessen erste Ableitung der Pfadkosten minimal sind, wird ein Teil der Last zugeordnet. Wenn die Anzahl der LSPs mit minimaler erster Ableitung größer als die Anzahl der restlichen zu verteilenden Lastanteile sind, dann werden die Lastanteile zufällig über diese verteilt.
7. Wenn der Algorithmus noch nicht alle Lastanteile den LSPs zugeordnet hat, fährt der Algorithmus bei Schritt 3 fort. Ansonsten endet der Mehrpfad-Algorithmus.

3.5 Qualitative Analyse des Verkehrslenkungsverfahrens

Dieser Abschnitt führt eine qualitative Analyse des entwickelten verteilten reaktiven MPLS Traffic Engineering Verfahrens durch. Die in den folgenden Abschnitten vorgestellte Bewertung erfolgt anhand der Anforderungen an reaktive Verkehrslenkungsverfahren, welche Abschnitt 2.5.1 definiert.

3.5.1 Bewertung der Stabilität

Im ersten Schritt geht die Betrachtung der Stabilität von konstantem Eingangsverkehr und idealer Kenntnis der Netzauslastung aus. Im zweiten Schritt geht der Abschnitt auf den Einfluss von dynamischem Verkehr auf die Stabilität der Verfahren ein. In beiden Fällen ist die fehlerfreie Funktionsweise des Netzes vorausgesetzt.

3.5.1.1 Stabilitäts-Analyse bei konstantem Eingangsverkehr Beide Verfahren wählen die LSPs bzw. die LSP-Fächer sequentiell aus. Für einen ausgewählten LSP bzw. LSP-Fächer berechnen beide Verfahren einen neuen Pfad bzw. neue Lastverteilungsgewichte. Sowohl das Verfahren mit Pfadumlegung als auch das Verfahren mit Mehrpfad-Lastverteilung führen eine Lastumverteilung dann aus, wenn diese die Summe der gesamten Linkkosten reduziert (siehe Gleichung 37). Der Fall, dass bei gleichbleibenden gesamten Linkkosten aber reduzierter gesamter Netzlast ebenfalls eine Lastverteilung erfolgt (siehe hierzu Abschnitt 3.4.2), wird zunächst nicht beachtet und im Anschluss diskutiert.

$$\text{Kriterium 1 : } C_{ges\ alt} > C_{ges\ neu} \geq C_{ges\ min} \quad (37)$$

Für eine konstante Eingangslast existiert sowohl bei dem Pfadumlegungsansatz als auch bei dem Mehrpfadansatz eine optimale Lastverteilung mit minimaler Summe der Linkkosten $C_{ges\ min}$. Mit jeder Lastumverteilung nähern sich beide Verfahren den minimalen Linkkosten an. Das zeigt einerseits, dass beide Verfahren in die Richtung eines stabilen Zustands streben. Haben die Verfahren die optimale bzw. eine durch den Algorithmus nicht verbesserbare Lastverteilung erreicht, endet bei beiden Verfahren der Lastverteilungsprozess, da keine weitere Reduktion der gesamten Linkkosten erreicht werden kann.

Andererseits lässt sich anschaulich zeigen, dass beide Verfahren keine Oszillationen zwischen zwei oder mehreren Auslastungszuständen zulassen. Verlässt das Netz einen Auslastungszustand mit bestimmten gesamten Linkkosten in einen Auslastungszustand mit

geringeren gesamten Linkkosten, erreicht das Netz den alten Auslastungszustand nicht wieder. Das würde bei einer oder mehrerer der folgenden Lastumverteilungen eine Erhöhung der gesamten Linkkosten bedeuten, was die Lastverteilungs-Algorithmen nicht zulassen.

Bei dem Verfahren mit Pfadumlegung existiert zusätzlich der Fall, dass dieses bei gleichbleibenden gesamten Linkkosten aber reduzierter Netzlast L_{ges} die Last eines LSPs auf einen neuen Pfad umlegt. Diese Bedingung zeigt Gleichung 38. In diesem Fall wird mit jeder Pfadumlegung eines LSPs bei gleichbleibenden Netzkosten die minimale Netzauslastung $L_{ges\ min}$ angenähert. Diese entspricht unabhängig von der Eingangslast dem Ergebnis eines Shortest Path Routings.

$$\textit{Kriterium 2} : (C_{ges\ alt} == C_{ges\ neu}) \wedge (L_{ges\ alt} > L_{ges\ neu} \geq L_{ges\ min}) \quad (38)$$

Das Verfahren mit Pfadumlegung beendet den Lastverteilungsprozess, wenn sich bei konstanter Eingangslast weder die Linkkosten noch die gesamte Netzlast reduzieren lassen. Eine Oszillation der Lastverteilung zwischen zwei oder mehreren Auslastungszuständen mit gleichen gesamten Linkkosten lässt das Verfahren mit Pfadumlegung nicht zu. Das würde ansonsten bedeuten, dass eine Pfadumlegung eine Erhöhung der Netzlast bewirkt.

Diese Betrachtungen der Stabilität zeigt, dass beide Verfahren kein oszillierendes Verhalten zwischen Lastverteilungszuständen zulassen und einen stabilen Auslastungszustand anstreben. Die Begrenzung der maximalen Anzahl der Lastverteilungsaktionen, welche zum Erreichen eines stabilen Zustands notwendig sind, wird in der Folge betrachtet.

Die Anzahl der Lastumverteilungen pro Lastverteilungsprozess ist bei beiden Verfahren auf die Anzahl der LSPs bzw. LSP-Fächer begrenzt. Wie in Abschnitt 3.3.3 definiert, verteilen beide Verfahren die Last eines LSPs bzw. LSP-Fächers nur einmal pro Lastverteilungsprozess. Allerdings wird nach Beendigung eines Lastverteilungsprozesses, welcher die Überlastsituation nicht auflöst, der nächste Lastverteilungsprozess angestoßen.

Die gesamte Anzahl der Lastverteilungsaktionen im Pfadumlegungsansatz ist begrenzt auf Grund der Verwendung von ganzzahligen Kostenwerten pro Link. Die Anzahl der Pfadumlegungen, welche die Kosten reduzieren, ist begrenzt auf die Differenz der initialen gesamten Linkkosten und der minimalen gesamten Linkkosten. Über den Faktor $F_{genauigkeit}$ variiert neben der Genauigkeit der Auflösung der Linkkosten auch die maximale Anzahl an Pfadumlegungen. Zusätzlich kommen im Pfadumlegungsansatz die Pfadumlegungen dazu, welcher bei gleichen Linkkosten die Netzlast reduzieren. Die maximale Anzahl dieser Pfadumlegungen ist abhängig von der Topologie und den initialen Pfaden der LSPs.

Im Mehrpfadansatz begrenzt ebenfalls die Verwendung ganzzahliger Werte für die erste Ableitung der Linkkosten die Anzahl der Lastverteilungsaktionen. Je weiter sich die gesamten Linkkosten mit jeder Lastverteilungsaktion an das Minimum der gesamten Linkkosten annähern, desto geringer werden die Unterschiede der ersten Ableitungen der Pfadkosten der LSPs eines LSP-Fächers. Je nach der Genauigkeit der Auflösung der Kostenwerte über die Wahl des Faktors $F_{genauigkeit}$ können die Unterschiede zwischen den ersten Ableitungen der Pfadkosten aufgelöst werden.

3.5.1.2 Stabilitäts-Analyse bei dynamischem Eingangsverkehr Das Ziel der entwickelten Verfahren ist die Optimierung der Lastverteilung bei unerwartet auftretenden

Verkehrsschwankungen. Treten ausschließlich Verkehrsschwankungen in Zeitabständen auf, welche groß sind verglichen mit der Dauer eines Mess-Intervalls bzw. der Dauer eines Lastverteilungsprozesses, kann nach einer Laständerung der Eingangsverkehr als konstant angenommen werden. In diesem Fall gelten für die Stabilität der Verfahren die Annahmen aus dem ersten Schritt. Die Dauer eines Mess-Intervalls bzw. der Dauer der Lastverteilung liegt in der Größenordnung von einigen Sekunden bis einigen Minuten. Für die Reaktion auf Verkehrsschwankungen, welche im Verhältnis dazu einen großen Zeitabstand haben, sind die reaktiven Verkehrslenkungsverfahren konzipiert.

Verkehrsschwankungen mit geringen Zeitabständen können andauernde Anpassungen der Lastverteilung erzeugen. Dieses quasi-oszillierende Verhalten eines reaktiven Verkehrslenkungsverfahrens lässt sich jedoch nicht auf die Instabilität des Verfahrens zurückführen, sondern stellt vielmehr eine grundsätzliche Problematik reaktiver Lastverteilung dar.

3.5.2 Bewertung der Netzperformanz

Dieser Abschnitt analysiert die Netzperformanz der beiden entwickelten Verfahren im eingeschwungenen und im transienten Zustand eines Lastverteilungsprozess. Die Güte der Netzperformanz im eingeschwungenen Zustand zeichnet sich durch das Erreichen bzw. das Annähern der optimalen Lastverteilung aus. Das wird für die beiden Verfahren in dem Abschnitt 3.5.2.1 diskutiert. Die Bewertung der Verfahren bezüglich der transienten Netzperformanz, welche sich durch eine kontinuierliche Verbesserung bzw. der Vermeidung von Performanz-Einbrüchen auszeichnet, ist Inhalt des Abschnitts 3.5.2.2. Auf die Netzperformanz bei dynamischem Eingangsverkehr wird schließlich in 3.5.2.3 eingegangen.

3.5.2.1 Netzperformanz im eingeschwungenen Zustand Mit der von beiden Verfahren verwendeten Zielfunktion der Minimierung der gesamten Linkkosten existieren eine oder mehrere optimale Lastverteilungen bei gegebener Eingangslast. Zum Erreichen der optimalen Lastverteilung kann es je nach Szenario notwendig sein, die Last eines oder mehrerer Eingangs-Ausgangsknotenpaare auf mehrere Pfade aufzuteilen.

Mit dem Mehrpfadansatz lässt sich die optimale Lastverteilung immer erreichen. Bei einer nicht-optimalen Lastverteilung existiert im Mehrpfadansatz mindestens ein LSP-Fächer, bei welchem Last auf einem LSP mit nicht minimaler erster Ableitung der Pfadkosten übertragen wird (siehe hierzu das Kuhn-Tucker-Theorem aus Gleichung 5). Der Mehrpfadansatz passt die Lastverteilung des LSP-Fächers an und reduziert somit die gesamten Linkkosten im Netz. Diesen Schritt wiederholt der Mehrpfadansatz solange, bis die gesamten Linkkosten dem Minimum entsprechen und die Last im Netz ausschließlich auf LSPs mit minimaler erster Ableitung der Pfadkosten übertragen wird.

Voraussetzung für das Erreichen der optimalen Lösung mit dem Mehrpfadansatz ist, dass zwischen jedem Eingangs-Ausgangsknotenpaar LSPs über alle möglichen Pfade eingerichtet sind. Das kann aber insbesondere bei großen Netzen eine sehr hohe Anzahl an LSPs bedeuten und je nach Szenario nicht praktikabel sein. Eine offene Frage ist daher, welche Netzperformanz der Mehrpfadansatz bei einer Begrenzung auf wenige LSPs pro Eingangs-Ausgangsknotenpaar erzielt. Eine quantitative Bewertung der Netzperformanz des Mehrpfadansatzes mit begrenzter Anzahl an LSPs pro LSP-Fächer präsentiert Kapitel 5 anhand ausgewählter Fallstudien.

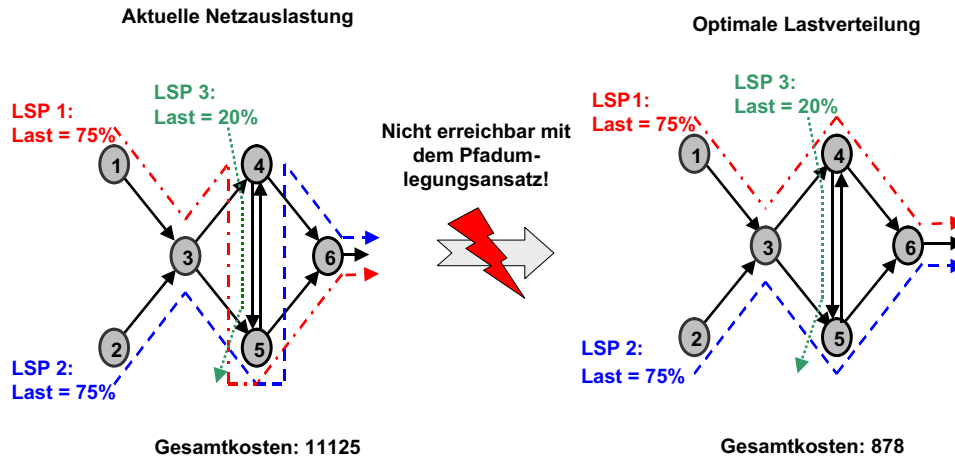


Abbildung 24: Verklemmungssituation im Pfadumlegungsansatz

Mit dem Pfadumlegungsansatz lässt sich offensichtlich die optimale Lastverteilung nicht immer erreichen, da nur ein LSP pro Eingangs-Ausgangsknotenpaar existiert. In der Folge bezieht sich der Pfadumlegungsansatz auf die optimale Lastverteilung mit der Nebenbedingung, dass nur ein LSP pro Eingangsknoten und Ausgangsknoten existiert. Auch hier zeigt sich, dass der Pfadumlegungsansatz nicht in allen Fällen die optimale Lösung erreicht.

Die Tatsache veranschaulicht die Abbildung 24. Links ist die aktuelle Netzauslastung dargestellt, bei der der Link zwischen Knoten 4 und 5 durch Überlagerung des LSP 1 und LSP 3 mit 95% ausgelastet ist. Die Gesamtkosten, welche hauptsächlich durch die Überlastung dieses Links entstehen, liegen bei 11125. Bei der rechts dargestellten optimalen Lösung betragen die gesamten Linkkosten 878. Diese lässt sich mit dem Pfadumlegungsansatz jedoch nicht erreichen, da dieser die LSPs sequentiell auswählt. Wählt der Pfadumlegungsansatz den LSP 1 aus, entspricht der neu berechnete Pfad dem alten Pfad. Der Pfad über die Knoten 1, 3, 4 und 6 ist ungünstiger, weil das auf dem Link zwischen Knoten 4 und 6 durch die Überlagerung von LSP 1 und LSP 2 eine Auslastung von 150% bedeutet. Das entspricht einem Kostenwert von ca. $1e07$. Gleiches gilt, wenn der Pfadumlegungsansatz den LSP 2 auswählt.

Als Fazit lässt sich aus diesen Betrachtungen für den Pfadumlegungsansatz ziehen, dass eine Abhängigkeit von dem jeweiligen Netz- und Auslastungsszenario vorliegt. Zur Bewertung dieser Abhängigkeit präsentiert Kapitel 5 eine quantitative Untersuchung des Pfadumlegungsansatzes anhand ausgewählter Fallstudien.

3.5.2.2 Netzperformanz im transienten Zustand Die Netzperformanz beschreibt die Güte eines reaktiven Lastverteilungs-Verfahrens während des Lastverteilungsprozesses. Dieser Abschnitt betrachtet die Netzperformanz im transienten Zustand für den sequentiellen und parallelen Durchführungsmodus.

Werden die beiden entwickelten Verfahren im sequentiellen Modus betrieben, ist die Einhaltung der Reihenfolge der LSP-Reroutings bzw. der Anpassung der Lastverteilungsgewichte der LSP-Fächer gewährleistet. Setzt man die fehlerfreie Funktionsweise des Netzes voraus, garantiert das die Vermeidung von unerwarteten und unerwünschten Netzauslastungen.

Der parallele Durchführungsmodus hingegen gibt keine Garantie der Einhaltung der bei der Berechnung festgelegten Reihenfolge während eines Lastverteilungsprozesses. Unterschiedliche Startzeiten der Lastverteilungsprozesse an den verteilten Eingangs-LSRs bzw. unterschiedliche Geschwindigkeiten bei der Lastverteilungsberechnung können Vertauschungen in der Reihenfolge bei der Durchführung der Lastverteilungsaktionen bewirken. Das kann zu einer kurzzeitigen Verringerung der Netzperformanz führen.

3.5.2.3 Netzperformanz bei dynamischem Eingangsverkehr Die Haupteinflussfaktoren des dynamischen Verkehrs, welche sich auf die Netzperformanz der reaktiven Verkehrslenkungsverfahren auswirken, ist der mittlere zeitliche Abstand der Verkehrsschwankungen sowie die Größe der Verkehrsschwankungen. Während die Größe einer Verkehrsschwankung keinen negativen Einfluss auf die Netzperformanz der Verfahren hat, kann eine zu geringe Dauer zwischen den Verkehrsschwankungen, welche in der Größenordnung der Dauer des Mess-Intervalls und darunter liegt, zu ungenauen oder zum Zeitpunkt der Lastumverteilung nicht mehr aktuellen Mess-Ergebnissen führen. Ungenaue bzw. veraltete Mess-Ergebnisse können wiederum Fehlentscheidungen bei der Lastumverteilung bedeuten und zu einer Verringerung der Netzperformanz führen.

Zur Verifizierung dieses Effekts, präsentiert Kapitel 5 eine quantitative Untersuchung der Netzperformanz in Abhängigkeit der Verkehrsdynamik.

3.5.3 Bewertung der Konvergenzdauer

Die Konvergenzdauer bezeichnet die Zeitspanne zwischen dem Auftreten und der durch eine Lastumverteilung erzielten Auflösung einer unerwarteten Überlastsituation im Netz. Dieser Abschnitt analysiert qualitativ die Konvergenzdauer der beiden Verfahren sowohl im sequentiellen als auch im parallelen Ausführungsmodus.

Die Konvergenzdauer setzt sich aus der Dauer der Überlasterkennung und der Dauer der Lastumverteilung zusammen. Die Dauer der Erkennung T_{erk} ist abhängig von der Größe des Mess-Intervalls. Je kürzer das Mess-Intervall ist, desto schneller erkennen die Verkehrslenkungsverfahren eine Überlastsituation. Das wiederum verringert die Konvergenzdauer der Verfahren. Dagegen bringt ein zu kurzes Mess-Intervall den unerwünschten Effekt mit sich, dass die Verfahren auf kurzfristige Verkehrsschwankungen reagieren. Das ist insbesondere dann unerwünscht, wenn die Dauer der Lastumverteilung größer als die Dauer zwischen den Verkehrsschwankungen ist, da die Berechnung der Lastumverteilung dann auf veralteten Messwerten beruht. Daher muss bei der Festlegung des Mess-Intervalls sowohl die Konvergenzdauer als auch die Dynamik der Verfahren berücksichtigt werden.

Bei der Betrachtung der Dauer der Lastumverteilung wird der sequentielle und der parallele Durchführungsmodus unterschieden. Im sequentiellen Modus hängt die Dauer der Lastumverteilung von der Anzahl der Lastverteilungsaktionen N_a pro Lastverteilungsprozess, von der Anzahl der Lastverteilungsprozesse N_p und von der Dauer einer einzelnen LSP-Lastverteilungsaktion ab. Die Gesamtdauer einer einzelnen LSP-Lastverteilungsaktion ergibt sich aus der Summe der Dauer des Austausches der LSP-Auslastungs-Informationen $T_{last-info}$, der Dauer zur Berechnung der Lastumverteilung T_{iv-cal} und im Pfadumlegungsansatz der Dauer des Verbindungsaufbaus $T_{lsp-sig}$. (siehe hierzu Abschnitt 3.3.4).

Die Konvergenzdauer des Pfadumlegungsansatzes im sequentiellen Modus T_{p-seq} beschreibt die Gleichung 39.

$$T_{p-seq} = T_{erk} + N_p \cdot (3 \cdot T_{last-info} + N_a \cdot (2 \cdot T_{last-info} + T_{lv-cal} + T_{lsp-sig})) \quad (39)$$

Die Konvergenzdauer des Mehrpfadansatzes im sequentiellen Modus T_{mp-seq} gibt Gleichung 40 wieder.

$$T_{mp-seq} = T_{erk} + N_p \cdot (3 \cdot T_{last-info} + N_a \cdot (2 \cdot T_{last-info} + T_{lv-cal})) \quad (40)$$

Mit zunehmender Anzahl an Lastverteilungsaktionen N_a nimmt die Konvergenzdauer im sequentiellen Betriebsmodus linear zu. Da eine Abhängigkeit der Anzahl der LSP-Lastverteilungsaktionen vom Netz- und Auslastungsszenario wahrscheinlich ist, sind starke Schwankungen der Konvergenzdauer zu erwarten.

Im parallelen Durchführungsmodus ist die Konvergenzdauer unabhängig von der Anzahl der tatsächlich durchgeführten Lastverteilungsaktionen eines einzelnen LSPs bzw. LSP-Fächers pro Lastverteilungsprozess, da die Ausführung der Lastumverteilungsaktionen parallel abläuft. Hier ergibt sich die Konvergenzdauer aus der Anzahl der durchzuführenden Lastverteilungsprozesse sowie aus der Dauer eines Lastverteilungsprozesses. Diese berechnet sich aus der Dauer zum Austausch der LSP-Auslastungs-Informationen $T_{last-info}$, der Dauer zur Berechnung aller LSP-Lastumverteilungen $T_{lv-par-cal}$ im parallelen Betriebsmodus und im Pfadumlegungsansatz der Dauer des Verbindungsaufbaus $T_{lsp-sig}$. Die Konvergenzdauer des Pfadumlegungsansatzes im parallelen Durchführungsmodus T_{p-par} beschreibt die Gleichung 41.

$$T_{p-par} = T_{erk} + N_p \cdot (5 \cdot T_{last-info} + T_{lv-par-cal} + T_{lsp-sig}) \quad (41)$$

Gleichung 42 zeigt die Konvergenzdauer des Mehrpfadansatzes im parallelen Durchführungsmodus T_{mp-par} .

$$T_{mp-par} = T_{erk} + N_p \cdot (5 \cdot T_{last-info} + T_{lv-par-cal}) \quad (42)$$

Durch die Parallelisierung der LSP-Lastverteilungsaktionen ist eine erhebliche Reduktion der Konvergenzdauer gegenüber dem sequentiellen Durchführungsmodus zu erwarten. Dennoch kann über die Konvergenzdauer der Verfahren im parallelen Betriebsmodus keine quantitative Aussage getroffen werden, da hierzu eine Abschätzung der Anzahl der notwendigen Lastverteilungsprozesse benötigt wird.

Um die Konvergenzdauer der beiden Verfahren im sequentiellen und parallelen Durchführungs-Modus zu quantifizieren, führt Kapitel 5 eine Analyse der Konvergenzdauer anhand ausgewählter Szenarien durch.

3.5.4 Bewertung der Komplexität der Lastverteilungsberechnung

Dieser Abschnitt analysiert im ersten Schritt die Komplexität der Lastverteilungsberechnung in Abhängigkeit von der Netzgröße. Das wird in Abschnitt 3.5.4.1 präsentiert. In einem zweiten Schritt vergleicht der Abschnitt den Mehrpfadansatz mit dem OpIATE-Verfahren (siehe Abschnitt 2.5.3.3). Diesen Vergleich präsentiert Abschnitt 3.5.4.2.

3.5.4.1 Komplexitäts-Analyse Die notwendigen Berechnungen zur Realisierung einer Lastumverteilung zeigt die folgende Aufzählung:

- Bestimmung der Link-Auslastungen aus den mittleren LSP-Auslastungen
- Bestimmung eines neuen Pfades für einen LSP im Pfadumlegungsansatz
- Bestimmung der neuen Verteilgewichte eines LSP-Fächers im Mehrpfadansatz

Bestimmung der Link-Auslastungen aus den mittleren LSP-Auslastungen: Die Bestimmung der Netzauslastung benötigt pro LSP und Link eine Rechenoperation. Die mittlere Anzahl der Rechenoperationen N_{op} entspricht der Multiplikation der Anzahl der LSPs N_{lsp} im Netz und dem mittleren Hop Count HC_{mean} (siehe Gleichung 43).

$$N_{op} = N_{lsp} \cdot HC_{mean} \quad (43)$$

Die Anzahl der LSPs N_{lsp} nimmt bei angenommener Vollvermaschung quadratisch mit der Anzahl der Netzknoten V zu. Der mittlere Hop Count HC_{mean} ist von der Anzahl der Netzknoten und Links und von der Vermaschung abhängig. Nimmt man den ungünstigsten Fall an und schaltet die Knoten in Serie, dann ist die Zunahme des mittleren Hop Counts HC_{mean} geringer als die Zunahme der Anzahl der Netzknoten. Daher ist die Abhängigkeit der Berechnungsdauer der Netzauslastung von der Netzgröße von geringerer Ordnung als $O(V^3)$. Das erlaubt eine obere Abschätzung der Berechnungsdauer in Abhängigkeit der Netzgröße.

Die Berechnungsdauer einer Floating Point Operation wird von einem Pentium 3 mit 200 MFLOPs abgeleitet. Bei einer Netzgröße von 100 Netzknoten, bei welcher zur Vollvermaschung 9900LSPs benötigt werden, und einem angenommenen mittleren Hop Count von 10 lässt sich als obere Grenze der Berechnungsdauer ein Wert von ca. $500\mu s$ abschätzen.

Als Fazit ergibt sich hieraus, dass zwar eine Abhängigkeit hoher Ordnung zwischen Berechnungsdauer und Netzgröße besteht, aber die tatsächliche Berechnungsdauer bei einem sehr großen Netz mit 100 Knoten gering ist.

Bestimmung eines neuen Pfades für einen LSP im Pfadumlegungsansatz: Die Berechnungsdauer eines Dijkstra-Algorithmus zur Bestimmung eines neuen Pfades ist nach [CLR90] maximal mit quadratischer Ordnung $O(V^2)$ abhängig von der Anzahl der Netzknoten V .

In [SG01] ist für zwei verschiedene Typen von Routern die Berechnungsdauer des Dijkstra-Algorithmus über der Anzahl der Netzknoten angegeben. Bei 100 Netzknoten benötigen beide Router ca. 25 ms.

Auch hier zeigt sich eine Abhängigkeit hoher Ordnung zwischen Berechnungsdauer und Netzgröße. Allerdings ist die tatsächliche Berechnungsdauer in einem großen 100Knoten-Netz auch hier eher gering.

Es ist allerdings zu beachten, dass in Abhängigkeit der Netzgröße die Anzahl der LSPs mit der Ordnung $O(V^2)$ zunimmt und somit voraussichtlich auch die Anzahl der Lastverteilungsaktionen pro Lastverteilungsprozess ansteigt. Da sich die gesamte Berechnungsdauer aus der Multiplikation der Dauer einer Shortest Path Berechnung und der Anzahl

der LSP-Lastverteilungsaktionen ergibt, kann das in Fällen extremer Schiefast zu großen Berechnungsdauern führen. Typische Überlast-Szenarien benötigen allerdings nur wenige LSP-Lastverteilungsaktionen verglichen zur gesamten Anzahl an LSPs im Netz.

Bestimmung der neuen Verteilgewichte eines LSP-Fächers im Mehrpfadansatz: Die Berechnungsdauer der Verteilgewichte eines LSP-Fächers nach dem Mehrpfadansatz ist abhängig von dem mittleren Hop Count der LSPs. Somit ist die Abhängigkeit von der Anzahl der Netzknoten von einer geringeren Ordnung als $O(V)$. Die Dauer der Berechnung neuer Verteilgewichte hängt zudem von der Anzahl der LSPs pro LSP-Fächer N_{lplf} , von der Anzahl der Rechenoperationen zur Bestimmung der ersten Ableitung eines Links N_{eal} und von der mittleren Anzahl an Rechenschritten S_{mean} des Gradienten-Projektions-Algorithmus zur Bestimmung der neuen Lastverteilungsgewichte ab. Gleichung 44 gibt die Anzahl der Rechenoperationen in Abhängigkeit der genannten Parameter.

$$N_{op} = N_{eal} \cdot S_{mean} \cdot HC_{mean} \cdot N_{lplf} \quad (44)$$

Die Berechnungsdauer einer Floating Point Operation wird ebenfalls von einem Pentium 3 mit 200 MFLOPs abgeleitet. Bei ca. 170 Rechenoperationen N_{eal} , einem angenommenen mittleren Hop Count von 10, drei LSPs pro LSP-Fächer und einer mittleren Anzahl von 100 Schritten des Mehrpfad-Algorithmus aus Abschnitt 3.4.3 bedeutet das eine reine Prozessorzeit von ca. $2.5ms$.

Das zeigt, dass die Bestimmung neuer Verteilgewichte eines LSP-Fächers einerseits nur gering von der Größe des Netzes abhängt und andererseits die Dauer der Berechnung neuer Verteilgewichte eines LSP-Fächers gering ist.

Wie bei der Bestimmung eines neuen Pfades im Pfadumlegungsansatz ist auch hier zu beachten, dass mit zunehmender Netzgröße und Anzahl der LSP-Fächer die Anzahl der LSP-Lastverteilungsaktionen steigt.

3.5.4.2 Vergleich mit dem Konkurrenz-Verfahren OpIATE Wie bereits in vorausgegangenem Abschnitt 3.5.4.1 erwähnt, führen die in dieser Arbeit entwickelten Verfahren nur für die LSPs bzw. LSP-Fächer eine Lastverteilungs-Berechnung durch, welche zuvor zur Lastumverteilung ausgewählt wurden. In typischen Überlastsituationen ist die Anzahl der durchgeführten Lastverteilungs-Berechnungen gering gegenüber der gesamten LSP-Anzahl.

Im Gegensatz dazu führt das OpIATE-Verfahren bei Erkennung einer Überlast die Berechnung neuer Verteilgewichte immer für jeden LSP-Fächer aus. Somit hängt das OpIATE-Verfahren mit quadratischer Ordnung $O(V^2)$ von der Größe des Netzes ab.

In dem genannten 100Knoten-Netz ergibt das bei geschalteten 9900 LSP-Fächern eine Berechnungsdauer von ca. 25 Sekunden, wenn für die Dauer zur Berechnung der Verteilgewichte eines LSP-Fächers der Wert von 2.5ms angenommen werden kann. Wenn, wie bei dem OpIATE-Verfahren vorgeschlagen, die Lastanpassung in mehreren Stufen durchgeführt wird, multipliziert sich die Berechnungsdauer mit der Anzahl der Stufen. Mit einer in [SP02] vorgeschlagenen Mindestanzahl von fünf Lastanpassungsstufen ergibt sich eine reine Berechnungsdauer von 125 Sekunden, die bei jeder Überlastsituation notwendig ist.

Die Auswertung der Berechnungskomplexität des OpIATE-Verfahrens zeigt die eingeschränkte Skalierbarkeit des Verfahrens. Der Vergleich mit der Berechnungskomplexität des OpIATE-Verfahrens zeigt, dass das in dieser Arbeit entwickelte Verfahren die Berechnungskomplexität erheblich reduziert.

3.6 Zusammenfassung wichtiger Aspekte des Verkehrslenkungsverfahrens

Dieser Abschnitt fasst die wesentlichen Aspekte des vorgestellten reaktiven verteilten Verkehrslenkungsverfahrens zusammen.

Das Verfahren basiert auf der MPLS Technologie und verwendet eine Vollvermaschung der Randknoten über LSPs. Da das Verfahren sowohl im Pfadumlegungsansatz als auch im Mehrpfadansatz betrieben werden kann, werden die Rand-Knotenpaare entweder über einen oder über mehrere LSPs verbunden.

Die Bestimmung der Netzauslastung erfolgt über die Messung der LSP-Auslastungen an den Eingangs-LSRs. Diese wird periodisch über Zeitintervalle in der Größenordnung von Sekunden bis Minuten durchgeführt. Über einen Flutungsmechanismus verteilen die Eingangs-LSRs die LSP-Auslastungs-Informationen an alle anderen Eingangs-LSRs im Netz. Basierend auf dem Kenntnis der Auslastungen und der Pfade aller LSPs bestimmen die Eingangs-LSRs die Netzauslastung. Da allen Eingangs-LSRs die gleichen Informationen vorliegen, berechnen sie eine identische Netzauslastung.

Ein Lastumverteilung wird durch das Verfahren nur dann realisiert, wenn eine Überlastung eines oder mehrerer Links vorliegt. Zur Feststellung von Überlastsituationen prüfen die Eingangs-LSRs nach jeder Aktualisierung der Auslastungs-Informationen und der darauf folgenden Neuberechnung der Netzauslastung, ob die Auslastung eines oder mehrere Links eine Überlastschwelle überschreitet.

Im Fall einer Überlastsituation führen die Eingangs-LSRs die Lastumverteilungsaktionen aus. Hierzu bestimmen diese über einen in diesem Kapitel definierten Algorithmus einen LSP. Da die Eingangs-LSR eine identische Sicht der Netzauslastung haben und den gleichen Algorithmus verwenden, wählen alle LSRs den selben LSP aus. Der Eingangs-LSR des ausgewählten LSPs führt die Umverteilung der Last schließlich aus.

Wird das Verfahren mit dem Pfadumlegungsansatz betrieben, berechnet der Eingangs-LSR einen neuen Pfad für den LSP. Der hierzu verwendete Routing-Algorithmus basiert auf dem Routing im Codex-Netz [BG87]. Wird der Mehrpfadansatz verwendet, legt der Eingangs-LSR die neuen Gewichte fest, mit denen die Last auf die LSPs zwischen zwei Randknoten verteilt wird. Der Algorithmus zur Bestimmung der Gewichte basiert auf dem Kuhn-Tucker-Theorem [BG87].

Zur Auflösung einer Überlastsituation führen die Eingangs-LSRs in der Regel mehrere LSP-Lastverteilungsaktionen aus. Im sequentiellen Durchführungsmodus des Verfahrens erfolgen die LSP-Lastverteilungsaktionen hintereinander. Erst nach erfolgreicher Durchführung einer LSP-Lastverteilungsaktion beginnen die Eingangs-LSRs mit der Bestimmung der nächsten. Dieser Modus hat den Vorteil, dass die festgelegte Reihenfolge der LSP-Lastverteilungsaktionen eingehalten wird und somit keine undefinierten Auslastungszustände entstehen. Im parallelen Durchführungsmodus werden die LSP-Lastverteilungsaktionen ausgeführt, ohne die Durchführung der vorherigen abzuwarten. Dadurch lässt

sich auf der einen Seite eine Verkürzung der Konvergenzdauer erzielen. Auf der anderen Seite ist die Einhaltung der Reihenfolge der LSP-Lastverteilungsaktionen nicht garantiert, so dass während der Ausführung der Lastumverteilung kurzzeitige Performanzeinbrüche denkbar sind.

Die in diesem Abschnitt durchgeführte Bewertung weist die Stabilität des Verfahrens nach. Ferner konnte die Reduktion der Berechnungskomplexität gegenüber dem Konkurrenzverfahren OpIATE gezeigt werden. Die qualitative Analyse der Netzperformanz des Verfahrens hat ergeben, dass das Verfahren bei Verwendung des Mehrpfadansatzes unabhängig vom Netz- und Auslastungsszenario die Performanz der optimalen Lastverteilung annähern kann. Fraglich ist allerdings die Anzahl der hierfür benötigten LSPs. Bei Verwendung des Pfadumlegungsansatzes hingegen ist von einer Abhängigkeit des Konvergenzverhaltens von dem Netz- und Auslastungsszenario auszugehen.

Die qualitative Analyse des Verfahrens bezüglich der Konvergenzdauer zeigt, dass bei Verwendung des parallelen Durchführungsmodus eine erhebliche Verkürzung dieser gegenüber dem sequentiellen Durchführungsmodus zu erwarten ist. Während bei dem sequentiellen Durchführungsmodus eine große Abhängigkeit vom Netz- und Auslastungsszenario zu erwarten ist, lässt der parallele Modus eine geringe Abhängigkeit vermuten.

Zur Untersuchung der Abhängigkeit des Verfahrens von dem Netz- und Auslastungsszenario und zur Quantifizierung des Konvergenzverhaltens ist eine simulative Untersuchung der Verfahren anhand ausgewählter Fallstudien erforderlich. Die Ergebnisse dieser Untersuchung präsentiert Kapitel 5. Die eingesetzte Simulations-Methode, welche eine Bewertung der Verfahren auch in großen Netzen erlaubt, wird zuvor in Kapitel 4 beschrieben.

4 Eine skalierbare ratenbasierte Simulations-Methode

Die Analyse der reaktiven Lastverteilungsverfahren in Kapitel 3 zeigt, dass das Konvergenzverhalten der Verfahren abhängig von dem jeweiligen Netz- und Auslastungsszenario ist. Um das Konvergenzverhalten anhand einiger ausgewählter Fallstudien quantitativ bewerten zu können, wird eine Methode zur Untersuchung der Verfahren benötigt.

Die Anforderungen an diese Untersuchungs-Methode lassen sich in den folgenden drei Punkten zusammenfassen. Erstens muss die Untersuchungs-Methode die Bewertung der reaktiven Lastverteilungsverfahren bei statischem und dynamischem Eingangsverkehr zulassen. Zweitens muss die Methode eine realistische Abschätzung der Performanz in IP-Netzen ermöglichen. Hierbei sollte die Methode eine Performanz-Analyse von TCP-Verbindungen erlauben, welche den Hauptanteil des IP-Verkehrs ausmachen. Die dritte Anforderung an die Untersuchungs-Methode ist die Bewertung der reaktiven Lastverteilungsverfahren in Netzen einer Größe, welche typisch für die Netze der IP-Netzbetreiber ist. Diese Anforderung wird in der Folge als Skalierbarkeit der Untersuchungs-Methode beschrieben.

Auf Grund der Komplexität der reaktiven Lastverteilungsverfahren ist eine analytische Betrachtung zur Bewertung der Verfahren nicht realisierbar. Daher konzentriert sich dieses Kapitel auf Untersuchungs-Methoden, welche auf der Performanz-Simulation von IP-Netzen basieren.

Zur Auswahl einer geeigneten Methode zur Performanz-Simulation von IP-Netzen wird in Abschnitt 4.1 zunächst der Stand der Technik analysiert. In dem Abschnitt werden zwei Methoden zur Simulation von IP-Netzen präsentiert und hinsichtlich der definierten Anforderungen bewertet. Die erste Methode basiert auf der paketbasierten Modellierung des IP-Verkehrs. Diese Methode wird typischerweise zur Simulation von IP-Netzen eingesetzt. Die zweite Methode basiert auf einer Modellierung des IP-Verkehrs auf Ratenebene.

Die Analyse der Methoden zeigt, dass beide die gestellten Anforderungen zur Untersuchung der reaktiven Lastverteilungsverfahren nicht erfüllen. Beide Methoden erlauben zwar eine genaue Analyse der Performanz, lassen sich jedoch auf Grund der zu großen Simulationskomplexität lediglich für die Untersuchung kleiner Testnetze verwenden.

Aus diesem Grund wurde im Rahmen dieser Arbeit eine neuartige Methode zur Simulation von IP-Netzen entwickelt, welche der Abschnitt 4.2 beschreibt. Diese Methode ist eine modifizierte Variante der ratenbasierten Simulations-Methode. Auf Grund einer vereinfachten Modellierung des Verkehrs lässt sich mit dieser Methode eine erhebliche Reduzierung der Simulationskomplexität erzielen, welche die Untersuchung von realistischen Netzszenerarien ermöglicht. Die Analyse der modifizierten ratenbasierten Simulations-Methode in Abschnitt 4.2 zeigt, dass diese den Anforderungen zur Untersuchung von reaktiven Lastverteilungsverfahren entspricht.

Abschnitt 4.3 fasst schließlich die wesentlichen Erkenntnisse des Kapitels zusammen.

4.1 Methoden zur Simulation von IP-Netzen: Stand der Technik

Dieser Abschnitt beschreibt die beiden bisher eingesetzten Methoden zur Performanz-Simulation von IP-Netzen und analysiert diese hinsichtlich der Anforderungen zur Untersuchung von reaktiven verteilten Lastverteilungsverfahren. Die Performanz-Simulation

von IP-Netzen, welche auf der paketbasierten Modellierung des IP-Verkehrs beruht, beschreibt und analysiert der Abschnitt 4.1.2. Abschnitt 4.1.3 geht auf die Simulations-Methode mit ratenbasierter Verkehrsmodellierung ein. Die Ergebnisse der Analyse der beiden Simulations-Methoden fasst Abschnitt 4.1.4 zusammen.

Beiden Methoden basieren auf der ereignisgesteuerten Simulationstechnik, welche in dem folgenden Abschnitt 4.1.1 näher erläutert wird.

4.1.1 Grundlagen zur ereignisgesteuerten Simulationstechnik

Die ereignisgesteuerte Simulationstechnik steuert den Ablauf einer Simulation über zeitdiskrete Ereignisse. Ein typischer Anwendungsfall der ereignisgesteuerten Simulationstechnik ist die Simulation von Paketnetzen.

Die Komponenten der ereignisgesteuerten Simulation sind in Abbildung 25 dargestellt. Diese setzen sich aus mindestens einem Ereignis-Handler sowie genau einem zentralen Scheduler zusammen. Die Ereignis-Handler und der zentrale Scheduler tauschen Ereignisse miteinander aus.

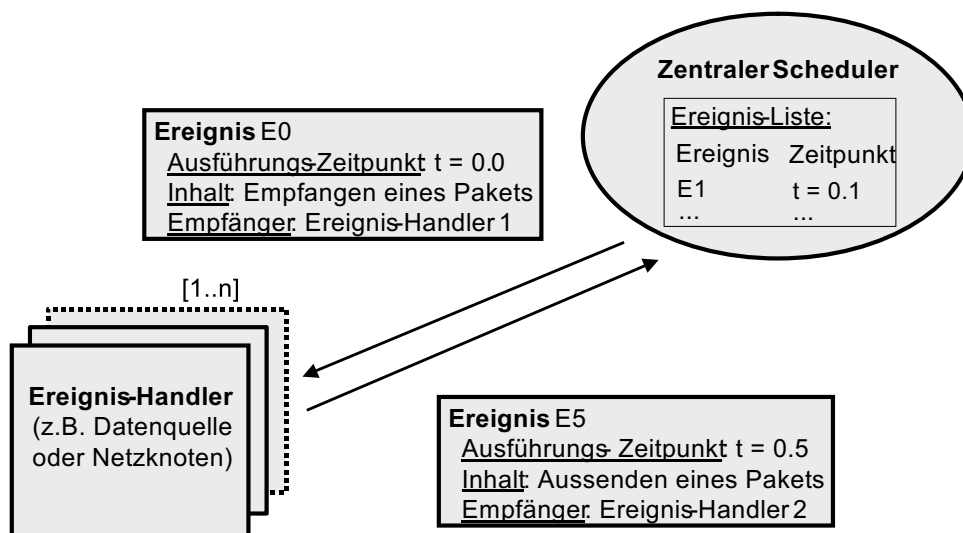


Abbildung 25: Komponenten der ereignisgesteuerten Simulation

Ein Ereignis-Handler bildet eine Instanz des simulierten Szenarios ab. Bei der Simulation eines Netzes kann ein Ereignis-Handler z.B. eine Datenquelle, ein Puffer oder ein Netzknoten sein. Ein Ereignis-Handler generiert Ereignisse, welche eine in der Zukunft auszuführende Aktion beschreiben. Für einen Ereignis-Handler, der eine Datenquelle abbildet, kann das z.B. das Aussenden eines Datenpakets sein. Das generierte Ereignis übergibt der Ereignis-Handler an den zentralen Scheduler. Der zentrale Scheduler fügt das neue Ereignis in eine vom ihm verwaltete zentrale Ereignisliste. Um die Einhaltung der kausalen Zusammenhänge zwischen den Ereignissen zu garantieren, trägt der zentrale Scheduler die Ereignisse in der Reihenfolge der Ausführungszeitpunkte in die Liste ein.

Der zentrale Scheduler steuert den Ablauf der Simulation. Dafür nimmt dieser das Ereignis mit dem nächsten Ausführungszeitpunkt aus der Ereignisliste und sendet das Ereignis an die entsprechende Ereignis-Handler-Instanz. Der empfangende Ereignis-Handler führt die durch das Ereignis ausgelöste Aktion aus. Bei der Ausführung der Aktion kann ein

Ereignis-Handler neue Ereignisse generieren, welche der zentralen Scheduler wiederum in die Ereignisliste einträgt.

Die aktuelle Simulationszeit in einer ereignisgesteuerten Simulation entspricht dem Ausführungszeitpunkt des aktuell bearbeiteten Ereignisses. Die Dauer der Ausführung eines Ereignisses ist gleich null.

4.1.2 Simulations-Methode mit paketbasierter Verkehrsmodellierung

Dieser Abschnitt erläutert die grundlegenden Aspekte der paketbasierten Simulations-Methode und präsentiert die Analyse der Methode hinsichtlich der Anforderungen zur Untersuchung von reaktiven Lastverteilungsverfahren. Hierzu beschreibt der Abschnitt 4.1.2.1 die Aspekte zur Modellierung der Netzkomponenten. Die beiden folgenden Abschnitte 4.1.2.2 und 4.1.2.3 bewerten die Genauigkeit und die Skalierbarkeit der paketbasierten Simulations-Methode.

4.1.2.1 Modellierungsaspekte der Netzkomponenten Das Hauptmerkmal dieser Simulations-Methode ist, dass jedes im Netz übertragene Paket in der Simulation abgebildet wird. Somit kann das Simulationsmodell beliebig genau an das reale Netzszenario angepasst werden. Die zu modellierenden Komponenten eines Netzes sind die Datenquellen, die Netzknoten sowie die Links zwischen den Netzknoten.

Ein Netzknoten setzt sich aus einer Switching-Matrix und aus Puffern zusammen. Die Switching-Matrix hat die Aufgabe, die Daten von den Eingängen des Netzknotens an die entsprechenden Ausgänge weiterzuleiten. Die Datenpakete werden von den Puffern in den Netzknoten zwischengespeichert. Die Switching-Matrix zeichnet sich durch ihre Übertragungskapazität aus. Die Modellierungsaspekte eines Puffers betreffen die Pufferstrategie (z.B. FIFO) und die Größe des Puffers.

Eine Datenquelle wird durch die Verkehrsparameter beschrieben, mit welcher diese den Verkehr in das Netz sendet. Bei der Übertragung von TCP-Verkehr muss das Modell der Datenquelle zudem die Eigenschaften des TCP berücksichtigen.

Ein modellierter Link in der paketbasierten Simulationsumgebung bildet die Paketverzögerung nach, welche sich in realen Netzen aus der Länge des Links x_{link} sowie der Signalausbreitungsgeschwindigkeit v_{sp} ergibt (siehe Gleichung 45).

$$d_{link} = \frac{x_{link}}{v_{signal}} \quad (45)$$

4.1.2.2 Analyse der Genauigkeit Die paketbasierte Simulations-Methode erlaubt eine exakte Modellierung eines IP-Netzes. Die Genauigkeit der Bestimmung der Performanzparameter, welche mit der paketbasierten Simulation erzielt wird, wird durch die Genauigkeit der Modellierung des IP-Netzes sowie des Eingangsverkehrs vorgegeben.

4.1.2.3 Analyse der Skalierbarkeit Die Abbildung jedes einzelnen Pakets in der Simulation bedeutet, dass für jede Übertragung eines Pakets zwischen zwei Netzkomponenten ein Ereignis in der Simulation generiert wird. Mit zunehmender Größe des Netzes

und zunehmender Übertragungskapazität der Links steigt somit die Anzahl der erzeugten Ereignisse, die in der Simulation verarbeitet werden müssen. Dadurch müssen in der Simulation mehr Ereignisse verarbeitet werden, was einen Anstieg der Simulationsdauer bewirkt. Eine zunehmende Anzahl an Ereignissen bedeutet in der Regel auch, dass die Länge der Ereignisliste im Scheduler anwächst und somit der Einsortiervorgang neuer Ereignisse in die Liste länger dauert. Das bewirkt eine weitere Zunahme der Simulationsdauer. Somit ist fraglich, bis zu welcher Netzgröße die paketbasierte Simulations-Methode eingesetzt werden kann.

Um die Skalierbarkeit der paketbasierten Simulations-Methode auf einem Standard-PC zu quantifizieren, wurde mit einem selbst entwickelten paketbasierten Simulator [S97] eine Simulation eines Netzes mit der Topologie aus Abbildung 50 durchgeführt. Das Netz besteht aus 14 Knoten und aus 48 bidirektionalen Links, welche jeweils eine Übertragungskapazität von $155 \frac{MBit}{s}$ haben. 90 Datenquellen erzeugen den Datenverkehr, welcher zwischen den 10 Randknoten ausgetauscht wird. Die mittlere Netzauslastung entspricht 40%. Diese ergibt sich aus dem Quotienten der gesamten Eingangslast, welche mit der mittleren Anzahl an verwendeten Links pro Verbindung multipliziert wird, und der Summe der Linkkapazitäten des Netzes.

Das Ergebnis der Untersuchung ist, dass ein Pentium4-Prozessor mit einer Taktfrequenz von 1.7GHz für die Simulation einer Sekunde Netzbetrieb 267 Sekunden benötigt. Da reaktive Lastverteilungsverfahren typischerweise auf Lastschwankungen reagieren, die im Abstand von mehreren Minuten aufeinander folgen, ist die Simulation der Verfahren über die Dauer einer oder mehrerer Stunden notwendig. Mit der zulässigen Annahme, dass die Simulationsdauer linear mit der simulierten Netzbetriebszeit ansteigt, dauert die Simulation einer Stunde Netzbetriebszeit in diesem Szenario 267 Stunden.

Diese Untersuchung zeigt, dass die paketbasierte Simulations-Methode für die Untersuchung reaktiver Lastverteilungsverfahren nicht die notwendige Skalierbarkeit aufweist. Diese Dauer ist wesentlich zu lang, um eine Analyse reaktiver Lastverteilungsverfahren praktikabel durchzuführen. Zudem ist eine Untersuchung weitaus größerer Netze wünschenswert.

4.1.3 Simulations-Methode mit ratenbasierter Verkehrsmodellierung

Wie in Abbildung 26 dargestellt, modelliert die ratenbasierte Simulations-Methode aufeinanderfolgende Datenpakete abschnittsweise durch konstante Raten. Bei Verwendung der ereignisgesteuerten Simulation wird jede Übertragung einer Ratenänderung als ein Ereignis modelliert. Je nach der Anzahl der Datenpakete, die durch einen Zeitabschnitt mit konstanter Rate zusammengefasst werden, lässt sich die Anzahl der Ereignisse gegenüber der paketbasierten Simulation reduzieren.

Die Modellierung des Verkehrs auf Ratenebene erfordert eine entsprechende Anpassung der Netzkomponenten. Daher beschreiben die folgenden zwei Abschnitte 4.1.3.1 und 4.1.3.2 Aspekte der Anpassung des Puffers und der Quelle an die ratenbasierte Verkehrsmodellierung. Da die Modellierung eines Links sowie einer Switching-Matrix keine relevanten Unterschiede zur Modellierung auf Paketebene aufweist, wird auf diese nicht näher eingegangen. Der nachfolgende Abschnitt 4.1.3.3 erläutert, wie die paketbezogenen Performanzparameter in einer ratenbasierten Simulationsumgebung bestimmt werden. Die abschließenden Abschnitte 4.1.3.4 und 4.1.3.5 analysieren die Genauigkeit sowie die Skalierbarkeit der ratenbasierten Simulations-Methode.

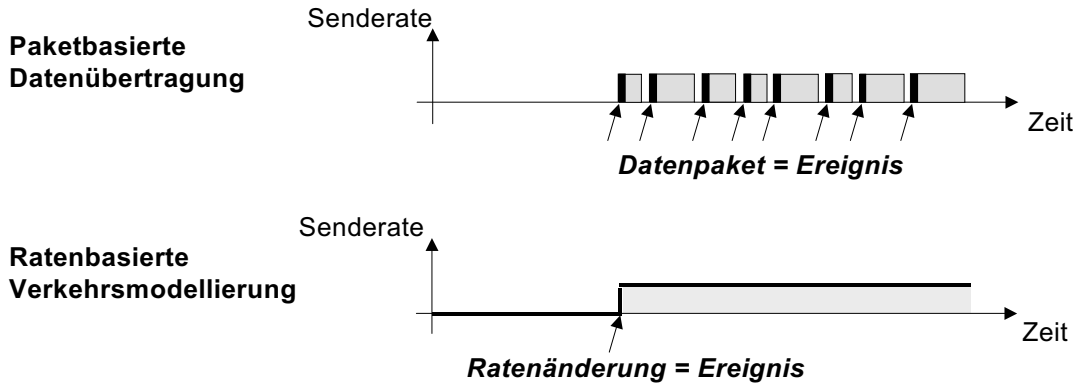


Abbildung 26: Modellierung des paketbasierten Datenverkehrs auf Ratenebene

4.1.3.1 Modellierung eines Puffers Dieser Abschnitt beschreibt die ratenbasierte Modellierung eines Paket-Puffers mit First In First Out FIFO Puffer-Strategie.

Abbildung 27 stellt die Kenngrößen des modellierten FIFO-Puffers dar. Mit der Eingangsrate I treffen die Daten an dem Puffer ein und mit der Ausgangsrate O sendet der Puffer die Daten wieder aus. Die Ausgangsrate O ist begrenzt durch die maximale Ausgangsrate O_{max} , welche der Übertragungskapazität am Pufferausgang entspricht. Der Parameter C_{max} definiert die Puffergröße. Mit Hilfe der Kenngrößen der Pufferfüllrate Q und des aktuellen Pufferfüllstands C wird das dynamische Verhalten des Pufferfüllstands modelliert. Zur Modellierung des Paketverlusts werden die Verlustrate L und der Datenverlust CL verwendet. Werden mehrere Verkehrsströme über einen Puffer übertragen, teilen sich diese die Puffer- und Ausgangskapazität proportional auf.

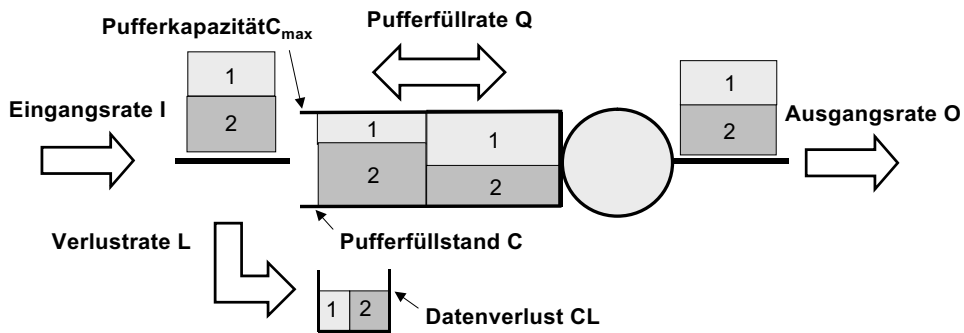


Abbildung 27: Ratenbasierte Modellierung des Pufferverhaltens

In den folgenden Gleichungen 46 bis 51 werden die Zustände des ratenbasierten Puffers beschrieben. Ist der Puffer leer und ist die Eingangsrate I kleiner als die maximale Ausgangsrate O_{max} , entspricht die Ausgangsrate O der Eingangsrate. In diesem Fall sind die Pufferfüllrate und die Verlustrate gleich null. Diesen Sachverhalt stellt Gleichung 46 dar.

$$I = O \wedge Q = 0 \wedge L = 0 \quad \text{für } C = 0 \wedge I \leq O_{max} \quad (46)$$

Ist die Eingangsrate I hingegen größer als die maximale Ausgangsrate O_{max} , so wird die Ausgangsrate O auf die maximale Ausgangsrate O_{max} gesetzt und die Pufferfüllrate entspricht der Differenz zwischen der Eingangsrate I und der maximalen Ausgangsrate O_{max} (siehe Gleichung 47).

$$Q = I - O \wedge O = O_{max} \wedge L = 0 \quad \text{für } C = 0 \wedge I > O_{max} \quad (47)$$

Wie Gleichung 48 zeigt, gilt das gleiche, wenn der Puffer nicht leer ist. Ist in diesem Fall die Eingangsrate I größer oder gleich der maximalen Ausgangsrate O_{max} , so ist die Pufferfüllrate Q größer bzw. gleich null. Ist die Eingangsrate I hingegen kleiner als die maximale Ausgangsrate O_{max} , so ist die Pufferfüllrate Q negativ.

$$Q = I - O \wedge O = O_{max} \wedge L = 0 \quad \text{für } 0 < C < C_{max} \quad (48)$$

Den dritten Fall, in dem die Pufferfüllrate Q der Differenz zwischen der Eingangsrate I und der Ausgangsrate O entspricht und die Ausgangsrate O gleich der maximalen Ausgangsrate O_{max} ist, zeigt Gleichung 49. Das entspricht dem Fall, dass der Puffer voll ist und die Eingangsrate I kleiner oder gleich der maximalen Ausgangsrate O_{max} ist. In diesem Fall ist die Pufferfüllrate kleiner oder gleich null.

$$Q = I - O \wedge O = O_{max} \wedge L = 0 \quad \text{für } C = C_{max} \wedge I \leq O_{max} \quad (49)$$

Ist hingegen bei einem vollen Puffer die Eingangsrate I größer als die Ausgangsrate O_{max} , so wird die Pufferfüllrate zu null gesetzt und die Verlustrate entspricht der Differenz der Eingangsrate I und der Ausgangsrate O (siehe Gleichung 50).

$$L = I - O \wedge Q = 0 \wedge O = O_{max} \quad \text{für } C = C_{max} \wedge I > O_{max} \quad (50)$$

Gleichung 51 fasst die oben genannten Gleichungen zu einer Gleichgewichtsbedingung zusammen, welche zu jedem Zeitpunkt an einem Puffer gelten muss.

$$I = Q + L + O \quad (51)$$

Der Pufferfüllstand C zu einem Zeitpunkt t_{n+1} lässt sich nach der folgenden Gleichung 52 aus dem aktuellen Pufferfüllstand C_{t_n} und der Pufferfüllrate Q_{t_n} bestimmen, wenn zwischen den Zeitpunkten t_n und t_{n+1} keine Ratenänderungen auftreten.

$$C(t_{n+1}) = C(t_n) + Q(t_n) \cdot (t_{n+1} - t_n) \quad (52)$$

In der gleichen Weise wird die verlorene Datenmenge CL bestimmt, was in Gleichung 53 gezeigt ist.

$$CL(t_{n+1}) = CL(t_n) + L \cdot (t_{n+1} - t_n) \quad (53)$$

Eine Neuberechnung der oben genannten Kenngrößen erfolgt bei jeder Änderung der Raten an dem Puffer. Abgesehen von der Änderung der Eingangsrate können Ratenänderungen noch durch zwei weitere Ereignisse hervorgerufen werden. Diese werden in der Folge beschrieben.

Das eine Ereignis tritt in dem Moment ein, wenn der Pufferfüllstand die maximale Pufferkapazität bzw. den Wert null erreicht. Im ersten Fall wird die Pufferfüllrate Q von dem Wert $I - O$ auf den Wert 0 und die Verlustrate L entsprechend von 0 auf $I - O$ gesetzt. Der Zeitpunkt des Erreichens des maximalen Pufferfüllstands lässt sich mit Gleichung

54 bestimmen, welche sich direkt aus Gleichung 52 ableitet. Im zweiten Fall muss die Pufferfüllrate von dem Wert $I - O$ auf den Wert 0 gesetzt werden und die Ausgangsrate O gleich der Eingangsrate I gesetzt werden. Der Zeitpunkt des Leerlaufens des Puffers errechnet sich in analoger Weise zum ersten Fall.

$$t_{n+1} = t_{C=C_{max}} = t_n + \frac{(C_{max} - C(t_n))}{Q(t_n)} \quad (54)$$

Das andere Ereignis, welches eine Neuberechnung der Pufferkenngrößen erfordert, wird als Propagations-Ereignis bezeichnet. Ein Propagations-Ereignis tritt auf, wenn sich der Datenverkehr eines Puffers aus mehreren Verkehrsströmen zusammensetzt. Eine Änderung der Eingangsrate eines Verkehrsstroms bedeutet eine Neuaufteilung der Puffer- und Ausgangskapazität auf die Ausgangsraten der Verkehrsströme. Bei einem nicht leeren Puffer wirken sich die Änderung der Eingangsrate mit einer Verzögerung auf die Ausgangsrate aus. Der Zeitpunkt, bei dem sich die Änderungen der Raten auf die Ausgangsrate auswirken, wird durch ein Propagations-Ereignis modelliert. Der Zeitpunkt t_{n+1} des Auftretens eines Propagations-Ereignisses nach einer Änderung der Eingangsrate zum Zeitpunkt t_n errechnet sich aus Gleichung 55.

$$t_{n+1} = t_{prop} = t_n + \frac{C(t_n)}{O_{max}} \quad (55)$$

4.1.3.2 Modellierung einer Quelle Die ratenbasierte Modellierung des Quellenverkehrs ist der Gegenstand dieses Abschnitts. Das ratenbasierte Quellenmodell legt fest, welche aufeinander folgenden Pakete zusammengefasst und durch eine Sendrate abgebildet werden. Zudem definiert es die Modellierung der TCP-Eigenschaften.

Dieser Abschnitt beschreibt drei mögliche Varianten der ratenbasierten Quellenmodellierung. Diese sind in Abbildung 28 dargestellt.

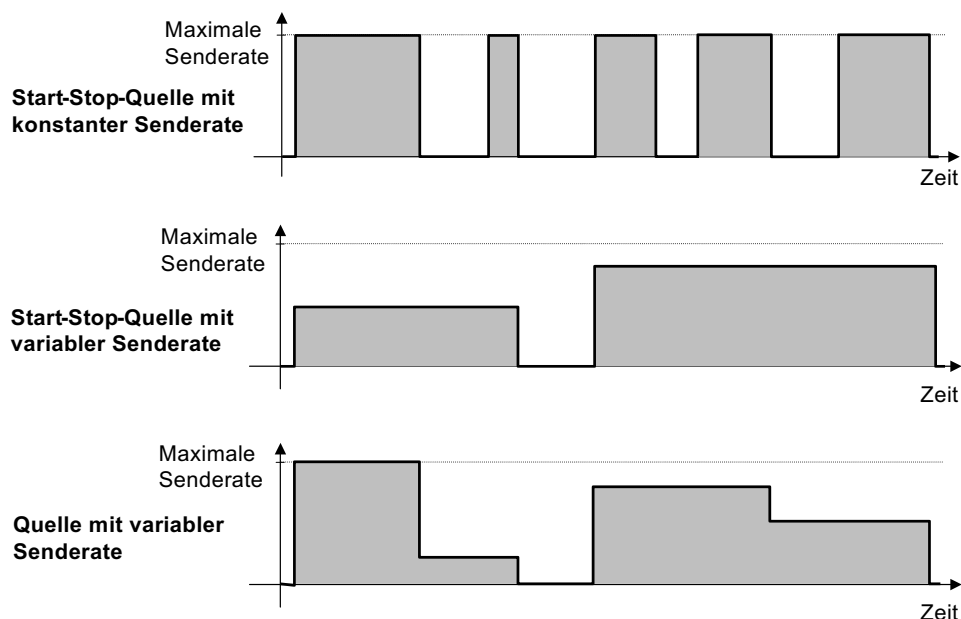


Abbildung 28: Ratenbasierte Modellierung des Quellenverkehrs

Die Start-Stop-Quelle mit konstanter Senderate unterscheidet zwei Zustände. Entweder sendet die Quelle mit der maximalen Senderate oder die Senderate der Quelle ist gleich null. Dieses Quellenmodell eignet sich für die Modellierung von Datenverkehr, bei dem die Quelle Paket-Bursts mit direkt aufeinander folgenden Datenpaketen überträgt. Bei diesem Quellentyp entspricht die Senderate der ratenbasierten Quelle der Senderate der paketbasierten Quelle.

Die Start-Stop-Quelle mit variabler Senderate wird ebenfalls für Datenverkehr verwendet, bei dem die Datenpakete in Paket-Bursts gesendet werden. Allerdings berücksichtigt dieses Quellenmodell den Fall, dass die Datenpakete innerhalb eines Bursts nicht direkt aufeinander folgen und die Zeit zwischen zwei Datenpaketen größer null sein darf. Der Zeitabstand zwischen zwei Datenpaketen wird in diesem Quellenmodell auf ein Maximum Δt_{max} begrenzt. Die Senderate des ratenbasierten Quellenmodells entspricht der mittleren Senderate eines Paket-Bursts und berechnet sich aus der Datenmenge und der Sendedauer eines Bursts.

Die beiden zuvor genannten Quellenmodelle sind Sonderfälle des dritten Quellenmodells. Das allgemeine Quellenmodell mit variabler Senderate ist so definiert, dass aufeinander folgende Datenpakete zu Abschnitten mit konstanter Rate zusammengefasst werden. Diese allgemeine Definition macht keine Vorgaben darüber, wie die Zuordnung eines Datenpakets zu einem Zeitabschnitt mit konstanter Rate erfolgt. Hier sind verschiedene Möglichkeiten denkbar, auf welche im Rahmen dieser Arbeit nicht näher eingegangen wird. Die Senderate des ratenbasierten Quellenmodells über die jeweiligen Zeitabschnitte entspricht auch hier der mittleren Senderate.

Zur Modellierung der Eigenschaften von TCP in dem Quellenmodell sind zwei unterschiedliche Varianten denkbar. In der ersten Variante modelliert eine Quelle den Verkehr einer TCP-Verbindung. In der zweiten Variante erzeugt eine Quelle einen aggregierten Verkehrsstrom, welcher dem Verkehr mehrerer TCP-Verbindungen entspricht.

4.1.3.3 Bestimmung der Performanzparameter Um die Performanz in IP-Netzen bewerten zu können, müssen die Performanzparameter des IP-Verkehrs mit Hilfe der Simulations-Methode bestimmt werden. Hierzu zählen die Paketverlustwahrscheinlichkeit, die Paketverzögerung sowie der Durchsatz einer TCP-Verbindung. Dieser Abschnitt geht darauf ein, wie die paketbasierten Performanzparameter aus dem ratenbasierten Simulationsmodell abgeleitet werden.

Die Anzahl der Paketverluste N_{pv} an einem Puffer lässt sich in einer ratenbasierten Simulationsumgebung direkt aus dem Datenverlust CL sowie der mittleren Paketlänge b_{mpl} der im Netz übertragenen Datenpakete ableiten (siehe Gleichung 56).

$$N_{pv} = \frac{CL_i}{b_{mpl}} \quad (56)$$

Die Paketverzögerung eines Verkehrsstroms lässt sich in der ratenbasierten Simulationsumgebung indirekt aus der Eingangsrate I und Ausgangsrate O sowie der Paketlängenverteilung eines Verkehrsstroms berechnen. Der Ansatz wurde in [BT00] beschrieben.

Abbildung 29 stellt das Vorgehen zur Bestimmung der Paketverzögerung dar. Mit der Kenntnis der Paketlängenverteilung werden aus der Senderate der Quelle die Zeitpunkte

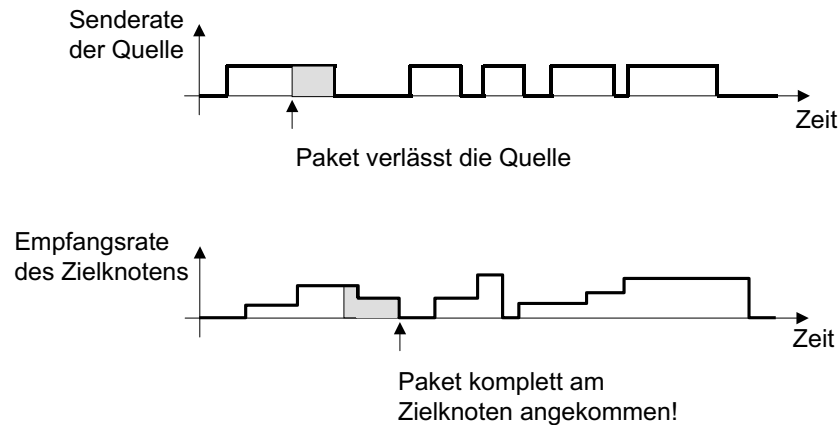


Abbildung 29: Berechnung der Paketverzögerung aus der Senderate und der Empfangsrate

ermittelt, zu denen die Datenpakete die Quelle verlassen. Aus der Empfangsrate am Zielknoten werden die Zeitpunkte ermittelt, zu denen die Datenmenge der einzelnen Pakete den Zielknoten erreicht. Die Differenz zwischen dem Aussende-Zeitpunkt an der Quelle und dem Empfangs-Zeitpunkt an der Senke entspricht der Verzögerung eines Pakets durch das Netz. Die Einschränkung dieser Methode ist, dass diese nur im verlustfreien Fall anwendbar ist. Wenn Daten verloren gehen, können die Empfangs-Zeitpunkte nicht mehr eindeutig den Paketen zugeordnet werden. In diesem Fall ist eine Erweiterung des Modells notwendig, welches zusätzlich die Paketverluste berücksichtigt. Diese Erweiterung ist nicht Gegenstand dieser Arbeit.

Der Durchsatz einer TCP-Verbindung lässt sich aus der Empfangsrate am Zielknoten ableiten.

4.1.3.4 Analyse der Genauigkeit Dieser Abschnitt bewertet die Genauigkeit, mit der die ratenbasierte Simulations-Methode die Performanzparameter in einem Paketnetz bestimmt. Die in dieser Analyse berücksichtigten Performanzparameter sind der Paketverlust, die Paketverzögerung sowie der TCP-Durchsatz. Dieser Abschnitt zeigt den Einfluss des Quellenmodells sowie der Puffergröße auf die Genauigkeit der Performanzparameter. Zur Evaluierung der Genauigkeit der ratenbasierten Simulations-Methode wird diese mit der paketbasierten Simulations-Methode verglichen.

Um den Einfluss des Quellenmodells sowie der Puffergröße auf die Genauigkeit der Performanzparameter zu bestimmen, werden in der Folge zwei Untersuchungen vorgestellt. In den Untersuchungen werden die ratenbasierte und paketbasierte Simulations-Methode verglichen. Als ratenbasierte Simulationsumgebung wurde der in [B97] vorgestellte ratenbasierte Simulator verwendet. Zur paketbasierten Simulation wird der bereits in Abschnitt 4.1.2.3 erwähnte Simulator verwendet. Beide Untersuchungen basieren auf einem Szenario, welches in Abbildung 30 dargestellt ist.

Der Einfluss der Puffergröße auf die Genauigkeit der ratenbasierten Simulation wurde in [BB03] untersucht. In dem gewählten Szenario wird der Verkehr von 64 Start/Stop-Quellen mit konstanter Senderate erzeugt. Jede Quelle sendet mit einer maximalen Senderate $R_{i \max}$ von $10 \frac{MBit}{s}$ Paket-Bursts mit einer mittleren Länge von 100 Paketen an den Puffer. Die Phasen, in denen die Quellen nicht senden, sind so gewählt, dass sich für jede Quelle eine mittlere Senderate von $1 \frac{MBit}{s}$ ergibt. Sowohl die Länge der Bursts als auch

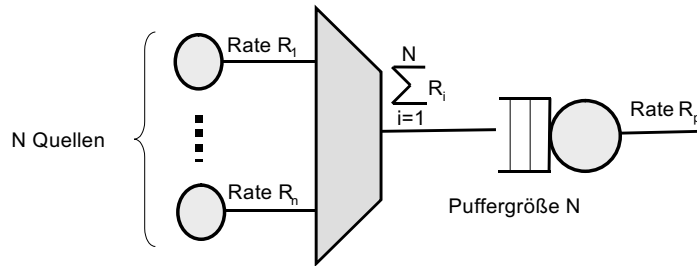


Abbildung 30: NetzszENARIO zur Genauigkeits-Analyse der ratenbasierten Simulation

die Dauer zwischen zwei Bursts unterliegt einer negativ exponentiellen Verteilung. Für die Ausgangsrate R_p des Puffers wurde der Wert $155 \frac{MBit}{s}$ gewählt. Zur Untersuchung des Einflusses der Puffergröße wurde diese zwischen einem und 80 Paketen variiert. Der Vergleich der Paketverlustwahrscheinlichkeit in Abhängigkeit der Puffergröße zwischen der paketbasierten und der ratenbasierten Simulation ist in Abbildung 31 dargestellt.

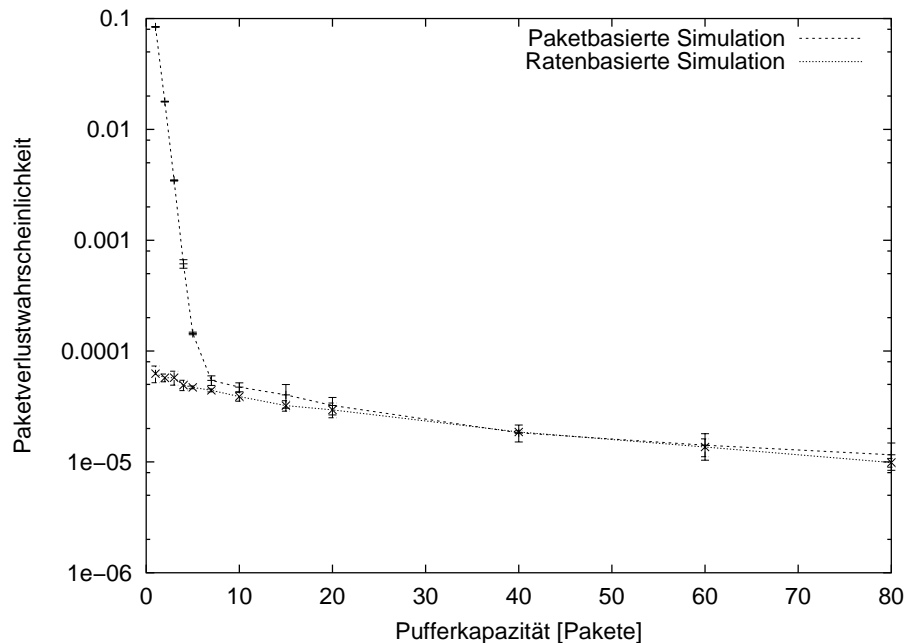


Abbildung 31: Einfluss der Puffergröße auf die Paketverlustwahrscheinlichkeit

Es ist zu erkennen, dass bei kleinen Puffergrößen die ratenbasierte Simulations-Methode die Paketverlustwahrscheinlichkeit der paketbasierten Variante deutlich unterschätzt. Dieser Effekt entsteht dadurch, dass Pakete an einem paketbasierten Puffer verloren gehen, wenn die Anzahl der gleichzeitig ankommenden Pakete größer als die maximale Speicherkapazität des Puffers ist. Dieser Fall tritt an einem paketbasierten Puffer auch dann auf, wenn die aktuelle Eingangsrate kleiner als die Ausgangsrate des Puffers ist. Die ratenbasierte Simulations-Methode misst hingegen nur dann Datenverlust, wenn die aktuelle Eingangsrate größer als die Ausgangsrate des Puffers ist. Mit steigender Puffergröße verringert sich der Einfluss dieses Effekts. Ab einer Pufferkapazität von 20 Paketen und größer ist der Effekt in dem beschriebenen Szenario nicht mehr messbar. Der Einfluss der Puffergröße auf die Genauigkeit der ratenbasierten Simulation wurde in einer vergleichbaren Untersuchung in [P95] vorgestellt.

In der Folge wird die Analyse des Einflusses des Quellenmodells auf die Genauigkeit der

ratenbasierten Simulations-Methode vorgestellt. In der in [BT00] beschriebenen Untersuchung wurde das StartStop-Quellenmodell mit konstanter Senderate (siehe Abschnitt 4.1.3.2) verwendet. Die Untersuchung basiert auf dem Szenario aus Abbildung 30, wobei vier Quellen mit einer maximalen Senderate von $155 \frac{MBit}{s}$ Paket-Bursts mit einer mittleren Länge von 100 Paketen an den Puffer senden. Der Abstand zwischen den Bursts ist so gewählt, dass die mittlere Senderate der Quellen 20% der maximalen Senderate entsprechen. Der Puffer hat eine Ausgangskapazität von $155 \frac{MBit}{s}$. Sowohl die Länge eines Paket-Bursts als auch die Dauer zwischen den Paket-Bursts sind negativ exponentiell verteilt. Zur Bestimmung der Paketverlustwahrscheinlichkeit wurde die Puffergröße auf 80 Datenpakete gesetzt. Bei der Bestimmung der Paketverzögerung wurden verlustfreie Puffer verwendet.

Das Ergebnis der Untersuchung zeigt, dass die ratenbasierte Simulation unter den genannten Voraussetzungen sowohl die Paketverlustwahrscheinlichkeit als auch die Paketverzögerung mit einer hohen Genauigkeit bestimmt. Das Ergebnis der ratenbasierten Simulation wurde durch den Vergleich mit einer paketbasierten Simulation verifiziert.

In der zweiten Untersuchung wird das StartStop-Quellenmodell mit variabler Senderate verwendet (siehe Abschnitt 4.1.3.2). Dieser Untersuchung liegt die Annahme zu Grunde, dass die Datenpakete eines Bursts nicht direkt hintereinander folgen, sondern mit einem bestimmten Abstand zueinander gesendet werden. In dieser Untersuchung wird die Genauigkeit der ratenbasierten Simulations-Methode in Abhängigkeit zum mittleren Abstand zwischen den Datenpaketen bestimmt. Dabei werden zwei Fälle unterschieden. In dem ersten Fall ist der Abstand zwischen den Paketen innerhalb eines Bursts konstant. Im zweiten Fall ist der Abstand gleichverteilt zwischen dem Wert 0 und der doppelten Größe des mittleren Abstandes.

Die Untersuchung basiert auf dem gleichen NetzszENARIO, welches in der vorherigen Untersuchung mit dem Quellenmodell mit konstanter Senderate eingesetzt wurde. Der einzige Unterschied ist die Puffergröße, welche auf den Wert von 20 Datenpaketen gesetzt wurde.

Die Tabellen 3 und 4 stellen die Ergebnisse der gemessenen Paketverlustwahrscheinlichkeit sowie des mittleren Pufferfüllstands in Abhängigkeit des mittleren Paketabstandes unter Angabe des Konfidenz-Intervalls dar. Der mittlere Paketabstand ist relativ zur mittleren Größe der Pakete angegeben.

Mittl. Paketabstand [Pakete]	Ratenbasierte Simulation	Paketbasierte Simulation mit festem Paketabstand	Paketbasierte Simulation mit variierendem Paketabstand
0.001	$0.2334 \pm 8.0 \cdot 10^{-05}$	$0.2336 \pm 6.0 \cdot 10^{-05}$	$0.2346 \pm 6.0 \cdot 10^{-05}$
0.01	$0.2324 \pm 3.6 \cdot 10^{-05}$	$0.2336 \pm 3.5 \cdot 10^{-05}$	$0.2336 \pm 3.5 \cdot 10^{-05}$
0.1	$0.2225 \pm 7.2 \cdot 10^{-05}$	$0.2238 \pm 6.0 \cdot 10^{-05}$	$0.2236 \pm 4.6 \cdot 10^{-05}$
1.0	$0.1092 \pm 6.0 \cdot 10^{-05}$	$0.1105 \pm 3.1 \cdot 10^{-05}$	$0.1140 \pm 3.7 \cdot 10^{-05}$
2.0	$0.0381 \pm 1.7 \cdot 10^{-05}$	$0.0395 \pm 2.0 \cdot 10^{-05}$	$0.0435 \pm 2.0 \cdot 10^{-05}$
3.0	0.0	0.0	$0.00012 \pm 4.7 \cdot 10^{-07}$

Tabelle 3: Paketverlustwahrscheinlichkeit in Abhängigkeit des mittleren Paketabstandes

Bei der Analyse der Genauigkeit der Paketverlustwahrscheinlichkeit zeigt sich, dass die ratenbasierte Simulations-Methode diese nur bei geringen mittleren Abständen zwischen den

Paketen in allen Fällen annähert. Ansonsten unterschätzt die ratenbasierte Simulation die Paketverlustwahrscheinlichkeit an dem Puffer. Für den Fall mit einem mittleren Abstand von drei Paketlängen zwischen zwei Paketen wird mit der ratenbasierten Simulations-Methode kein Paketverlust mehr gemessen. In diesem Fall entspricht die Senderate einer Quelle einem Viertel der maximalen Senderate, so dass in dem untersuchten Szenario die Summe der Senderaten der vier Quellen nicht größer als die Ausgangskapazität des Puffers ist. In diesem Fall bleibt der Puffer konstant leer. Bei der paketbasierten Simulation kommt es hingegen durch den variierenden Paketabstand dazu, dass die Eingangsrate größer als die Ausgangskapazität des Puffers ist und Pakete im Puffer zwischengespeichert werden. In diesen Phasen tritt auch Paketverlust auf, wie die Ergebnisse aus Tabelle 3 belegen.

Mittl. Paket-abstand [<i>Pakete</i>]	Ratenbasierte Simulation [<i>Bytes</i>]	Paketbasierte Simulation mit festem Paketabstand [<i>Bytes</i>]	Paketbasierte Simulation mit variierendem Paketabstand [<i>Bytes</i>]
0.001	389.2 ± 0.3	385.4 ± 0.4	385.4 ± 0.4
0.01	386.3 ± 0.2	382.9 ± 0.2	383.0 ± 0.2
0.1	363.9 ± 0.2	359.8 ± 0.3	360.7 ± 0.3
1.0	358.6 ± 0.2	357.1 ± 0.3	363.0 ± 0.3
2.0	286.7 ± 0.1	291.9 ± 0.2	307.4 ± 0.2
3.0	0.0	30.8 ± 0.004	116.3 ± 0.2

Tabelle 4: Mittlerer Pufferfüllstand in Abhängigkeit des mittleren Paketabstandes

Lediglich für den Sonderfall, dass die Pakete in einem Paket-Burst die gleichen Abstände zueinander haben, nähert die ratenbasierte Simulations-Methode die Paketverlustwahrscheinlichkeit der paketbasierten Simulation an. Dieser Fall tritt in der Praxis aber meistens nicht auf.

Zur Bewertung der Genauigkeit der Paketverzögerung, die mit der ratenbasierten Simulations-Methode erzielt werden kann, lässt sich analog zum Paketverlust argumentieren. Bei zu großem mittleren Abstand zwischen den Paketen bleibt der Puffer bei der ratenbasierten Simulations-Methode konstant leer und es tritt keine Paketverzögerung durch Pufferung auf. Das unterschätzt die tatsächliche Paketverzögerung an einem Puffer in einem Paketnetz, wenn die Paketabstände zwischen den Paketen variieren.

Da die Bestimmung der Paketverlustwahrscheinlichkeit sowie der Paketverzögerung die Voraussetzung zur Bestimmung des TCP-Durchsatzes sind, gelten die zuvor getätigten Aussagen auch für die Bewertung der Genauigkeit des TCP-Durchsatzes.

Als Fazit lässt sich aus diesen Untersuchungen schließen, dass die ratenbasierte Simulations-Methode nur unter Einhaltung der folgenden beiden Bedingungen eine ausreichende Genauigkeit bei der Bestimmung der Performanzparameter erzielt. Die erste Bedingung fordert eine ausreichende Größe der Puffer. Die zweite Bedingung fordert, dass nur direkt bzw. mit geringem Abstand aufeinander folgende Datenpakete zu Abschnitten mit konstanter Rate zusammengefasst werden dürfen. Somit eignen sich für die ratenbasierte Simulations-Methode nur das StartStop-Quellenmodell mit konstanter Senderate bzw. variabler Senderate, wobei bei Verwendung des zweiten Quellentyps ein geringer Abstand

zwischen den Paketen gewährleistet sein muss. Der Einsatz des Quellenmodells mit variabler Senderate ist nicht sinnvoll, da dieses auf der Mittelung der Senderate über Zeitabschnitte basiert.

4.1.3.5 Analyse der Skalierbarkeit Dieser Abschnitt vergleicht die ratenbasierte Simulations-Methode mit der paketbasierten Simulations-Methode bezüglich der Simulationsdauer. Da die Simulationsdauer der ratenbasierten Simulation von der Länge der Paket-Bursts sowie von dem Auslastungs- und NetzszENARIO abhängt, wird der Einfluss dieser Parameter berücksichtigt.

Im ersten Schritt vergleicht die folgende Untersuchung die Simulationsdauer der ratenbasierten und paketbasierten Simulations-Methode anhand des in Abschnitt 4.1.2 beschriebenen Szenarios.

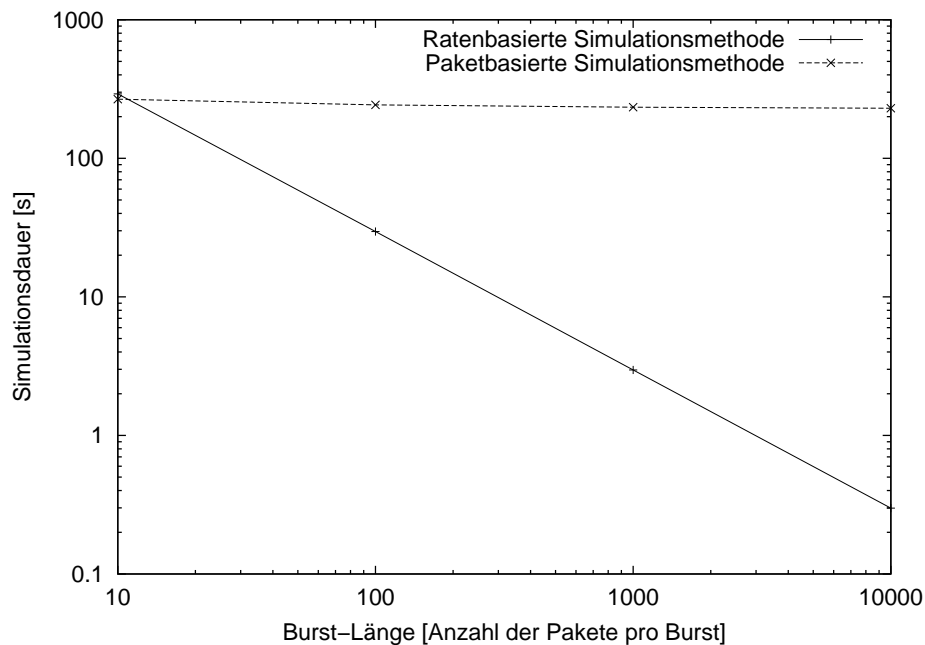


Abbildung 32: Vergleich der Simulationsdauer in Abhängigkeit der Burst-Länge

Die gemessenen Simulationsdauern der ratenbasierten und der paketbasierten Simulations-Methode in Abhängigkeit der Burst-Länge zeigen Abbildung 32. Der Vergleich der Simulations-Methoden zeigt, dass diese bei einer mittleren Burst-Länge von ca. 10 Paketen die gleiche Simulationsdauer haben. Mit steigender mittlerer Burst-Länge nimmt die Simulationsgeschwindigkeit der ratenbasierten Simulations-Methode gegenüber der paketbasierten Simulations-Methode zu.

Es ist zu erkennen, dass die Simulationsdauer $T_{RB\ BL}$ der ratenbasierten Simulations-Methode vom Kehrwert der mittleren Burst-Länge BL abhängt. Diesen Zusammenhang stellt Gleichung 57 dar. Die in der Gleichung angegebene Konstante $T_{BL=1}$ entspricht der Simulationsdauer der ratenbasierten Simulation bei einer Burst-Länge von einem Paket pro Burst. Die paketbasierte Simulations-Methode zeigt wie erwartet keine Abhängigkeit von der Burst-Länge.

$$T_{RB\ BL} = \frac{T_{BL=1}}{BL} \quad (57)$$

Um eine Analyse der Performanz von reaktiven Lastverteilungsverfahren in diesem Netz über eine bzw. mehrere Netzbetriebsstunden zu ermöglichen, ist somit eine Burst-Länge in der Größenordnung von 100 bis 1000 Paketen erforderlich.

Das verwendete ratenbasierte StartStop-Quellenmodell bildet einen Paket-Burst durch ein Ereignis am Beginn und ein Ereignis am Ende des Paket-Bursts ab. Daraus lässt sich schließen, dass bei einer mittleren Burst-Länge von zwei Paketen die ratenbasierte Simulations-Methode die gleiche Anzahl an Ereignissen zur Modellierung des Verkehrs benötigt wie die paketbasierte Simulations-Methode. Diese modelliert jedes Paket mit einem Ereignis. Bedingt durch diese Tatsache ist bei einer ersten Betrachtung zu erwarten, dass die Simulationsdauern der beiden Methoden bei einer mittleren Burst-Länge von zwei Paketen gleich sind. Dass dies nicht der Fall ist und die ratenbasierte Simulations-Methode erst bei einer mittleren Burst-Länge von 10 Paketen die Simulationsdauer der paketbasierten Methode erreicht, lässt sich auf die folgenden drei Punkte zurückführen.

Erstens erfordert die Verarbeitung des Verkehrs auf Ratenebene an den einzelnen Netzelementen eine komplexere Berechnung als bei der paketbasierten Modellierung. Für jede Änderung einer Rate muss ein Netzelement den eigenen Zustand neu berechnen. Zweitens erzeugt der ratenbasierte Puffer zusätzliche Ereignisse, um Änderungen des Pufferzustands zu modellieren. Diese Ereignisse wie z.B. ein Puffer-Überlauf beschreibt Abschnitt 4.1.3.1. Drittens kann eine Ratenänderung eines Verkehrsstroms Ratenänderungen anderer Verkehrsströme nach sich ziehen. Das geschieht durch die Aufteilung der Ausgangskapazität eines Puffers auf die vorhandenen Verkehrsströme. Dieser Sachverhalt ist ebenfalls in Abschnitt 4.1.3.1 erläutert. Der Einfluss dieser drei Punkte auf die Simulationsgeschwindigkeit hängt von dem jeweiligen Netz- und Auslastungsszenario ab. Dies wird in den beiden folgenden Untersuchungen analysiert.

In der ersten der beiden Untersuchungen wird die Steigerung der Simulationsgeschwindigkeit gegenüber der paketbasierten Simulation in Abhängigkeit der Anzahl der Quellen für das NetzszENARIO aus Abbildung 30 bewertet. Dabei werden die Szenarien mit einer Quelle, mit vier Quellen und mit 64 Quellen untersucht. In den drei Szenarien erzeugen die Quellen jeweils einer mittleren Pufferauslastung von 80%. Der Puffer ist verlustfrei und hat eine Ausgangsrate von $155 \frac{MBit}{s}$. Die Quellen senden Paket-Bursts mit einer Rate von $155 \frac{MBit}{s}$.

Als Ergebnis der Untersuchung zeigt Abbildung 33 die Steigerung der Simulationsgeschwindigkeit der ratenbasierten Methode gegenüber der paketbasierten Methode in Abhängigkeit der Burst-Länge. Die Ergebnisse der Untersuchung werden mit dem optimalen Steigerungsfaktor verglichen. Der optimale Steigerungsfaktor ergibt sich bei der Vernachlässigung der genannten drei Einfluss-Faktoren.

Zur Bewertung des Einflusses der Komplexität der ratenbasierten Simulations-Methode wird das Szenario mit einem Puffer und einer Quelle betrachtet. Im diesem erzeugt der Puffer keine zusätzlichen Ereignisse, da die maximale Senderate der Quelle nicht größer als die Ausgangskapazität des Puffers ist und somit keine Pufferung auftritt. Aus Abbildung 33 ist zu erkennen, dass der Steigerungsfaktor der Simulationsgeschwindigkeit in diesem Szenario bedingt durch die komplexere Verarbeitung der ratenbasierten Methode der Hälfte des optimalen Steigerungsfaktors entspricht.

Durch das Hinzunehmen weiterer Quellen erzeugt der Puffer zusätzliche Ereignisse. Die Anzahl der Propagationserereignisse sowie der nachfolgenden Ratenänderungs-Ereignisse steigt mit zunehmender Anzahl an Quellen an. Abbildung 33 zeigt den Einfluss dieser

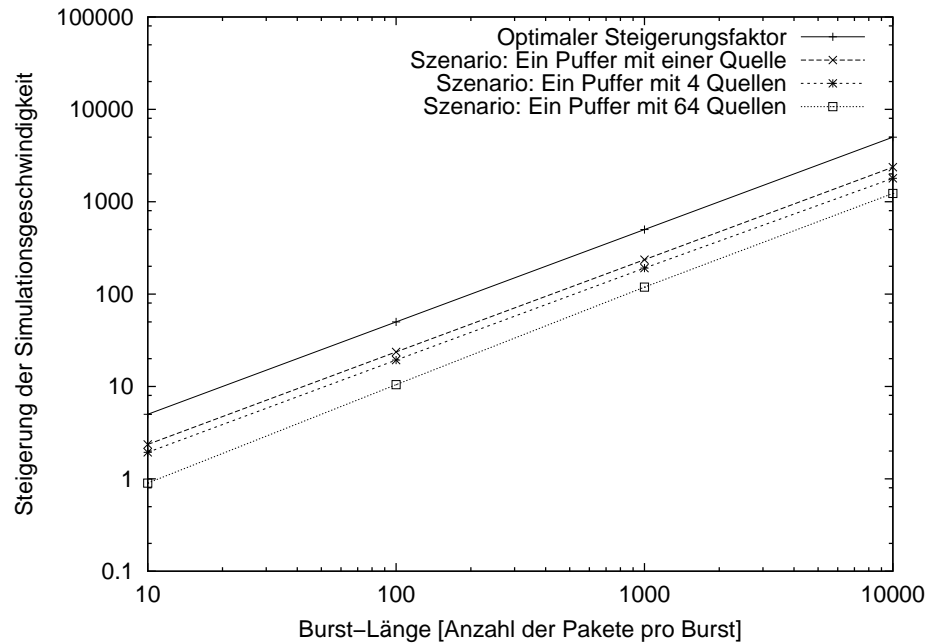


Abbildung 33: Simulationsgeschwindigkeit in Abhängigkeit der Quellen-Anzahl

zusätzlichen Ereignisse auf den Steigerungsfaktor der Simulationsgeschwindigkeit. Der Steigerungsfaktor in dem Szenario mit 64 Quellen und einem Puffer entspricht lediglich 20% des optimalen Steigerungsfaktors. Hier ist zu erkennen, dass der Steigerungsfaktor der Simulationsgeschwindigkeit mit steigender Anzahl an Verkehrsströmen pro Puffer abnimmt. Das weist eine Abhängigkeit des Steigerungsfaktors von der Größe des Netzes nach.

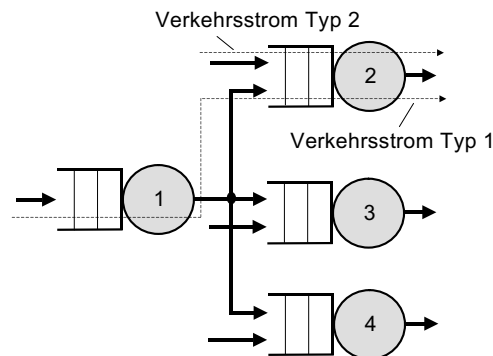


Abbildung 34: Verschaltung von vier Puffern in einer Baumstruktur

In einer zweiten Untersuchung wird der Einfluss der Netztopologie und des Routings der Verkehrsströme auf die Steigerung der Simulationsgeschwindigkeit analysiert. Hierzu wird der Steigerungsfaktor der Simulationsgeschwindigkeit in drei verschiedenen Netzszenarien bewertet. Das erste Netzszenario setzt sich aus vier voneinander unabhängigen Puffern zusammen. Jeder Puffer ist mit vier Quellen verbunden. In einem zweiten Szenario sind die vier Puffer in einer Baumstruktur verschaltet. Dieses Netzszenario zeigt Abbildung 34. Jeder der vier Puffer ist mit drei Quellen verbunden. Die Verkehrsströme, die von den drei Quellen des Puffers 1 erzeugt werden, werden zudem über einen der drei anderen Puffer übertragen (Verkehrsstrom Typ 1). Die Verkehrsströme der Quellen der anderen Puffer werden ausschließlich über einen Puffer übertragen (Verkehrsstrom Typ 2).

Das dritte NetzszENARIO ist in Abbildung 35 dargestellt. In diesem Szenario sind vier Puffer in einer Ringstruktur verschaltet. Jeder Puffer ist mit drei Quellen verbunden. Ein Verkehrsstrom einer der drei Quellen wird über zwei Puffer übertragen (Verkehrsstrom Typ 1). Die übrigen Verkehrsströme werden jeweils über einen Puffer übertragen (Verkehrsstrom Typ 2).

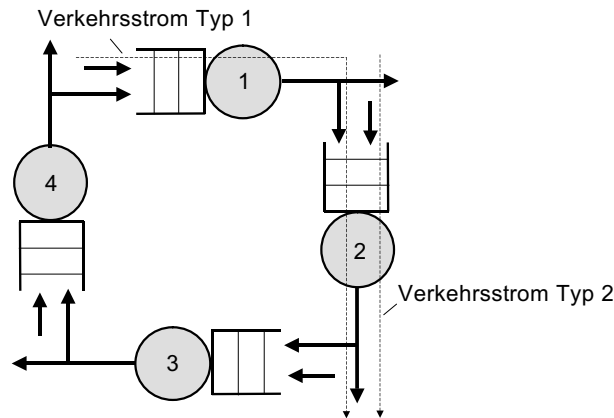


Abbildung 35: Verschaltung von vier Puffern in einer Ringstruktur

Die Ausgangskapazität des verlustfreien Puffers sowie die Senderate der Quellen ist wie in den Szenarien zuvor auf $155 \frac{\text{MBit}}{\text{s}}$ gesetzt. Die mittlere Senderate der Quellen ist so gewählt, dass die Eingangsrate an den Puffern im Mittel 80% der Ausgangsrate entspricht.

Das Ergebnis der Untersuchung des Einflusses verschiedener NetzszENARIEN auf die Simulationsgeschwindigkeit der ratenbasierten Simulations-Methode im Vergleich zur paketbasierten Simulations-Methode zeigt Abbildung 36. Es ist zu erkennen, dass das NetzszENARIO einen Einfluss auf die Simulationsdauer hat. Der Steigerungsfaktor in dem NetzszENARIO mit Ringstruktur ist ca. 80% von dem Steigerungsfaktor in dem Szenario mit vier unabhängigen Puffern. Dieser Effekt ist darin begründet, dass sich Ratenänderungen in den verschiedenen NetzszENARIEN in einer anderen Weise fortpflanzen.

In dem NetzszENARIO mit der Baumstruktur kann eine Ratenänderung eines Verkehrsstroms am Eingang des Puffers 1 Ratenänderungen der übrigen Verkehrsströme erzeugen. Dies wiederum erzeugt Ratenänderungen an den übrigen Puffern.

In dem NetzszENARIO mit der Ringstruktur kann eine Ratenänderung am Eingang eines Puffers Ratenänderungen an den folgenden Puffern und schließlich wieder an dem ersten Puffer auslösen. Dieser Fall tritt auf, wenn die gesamte Eingangsrate an jedem Puffer größer als die Ausgangskapazität des Puffers ist. Das Kreisen einer Ratenänderung wird nur dann unterbrochen, wenn bei einem der Puffer diese Gegebenheit nicht mehr vorliegt. Um das Kreisen der Ratenänderungen künstlich zu unterbrechen, kann die Tatsache ausgenutzt werden, dass die Änderungen der Rate immer kleiner werden. Mit der Definition einer Schwellwerts, der eine minimale Größe der Ratenänderung definiert, werden Ratenänderungen unterhalb dieses Schwellwerts nicht weitergeleitet.

Als Fazit lässt sich aus der vorgestellten Untersuchung der Skalierbarkeit der ratenbasierten Simulations-Methode ableiten, dass die Skalierbarkeit bedingt durch die Abhängigkeit von der Burst-Länge sowie dem Netz- und Auslastungsszenario für jeden Fall einzeln geprüft werden muss. Als Richtwert für ein Netz mit 14 Knoten sollte die mittlere Burst-Länge in der Größenordnung von 100 bis 1000 Paketen liegen, um eine ausreichende

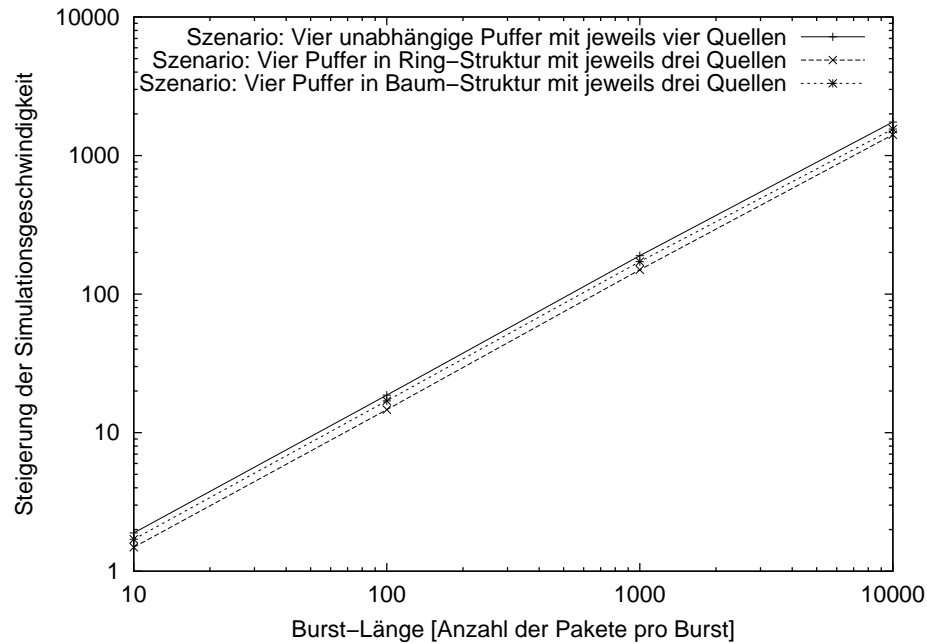


Abbildung 36: Simulationsgeschwindigkeit in Abhängigkeit des Netzszenarios

Skalierbarkeit zu erzielen. Für größere Netze ist bedingt durch die Abhängigkeit von der Netzgröße eine größere Burst-Länge erforderlich.

4.1.4 Zusammenfassende Bewertung bekannter Simulations-Methoden

Der Abschnitt 4.1 hat zwei Methoden zur Simulation von Paketnetzen vorgestellt. Die eine Methode ist die klassische paketbasierte Simulations-Methode, welche im überwiegenden Teil der Performanz-Untersuchungen in Paketnetzen eingesetzt wird. Die paketbasierte Simulations-Methode bildet jedes im Netz übertragene Datenpaket in der Simulation ab. Das lässt auf der einen Seite zwar eine sehr genaue Modellierung des Netzszenarios zu, hat aber auf der anderen Seite einen sehr hohen Simulationsaufwand zur Folge. Die Analyse der Simulationsdauer hat gezeigt, dass sich die Methode nicht in Netzen mit einer Größe anwenden lässt, welche zur Untersuchung von reaktiven Lastverteilungsverfahren notwendig ist.

Die zweite vorgestellte Simulations-Methode beruht auf der Modellierung des Datenverkehrs auf Ratenebene. Diese ratenbasierte Simulations-Methode fasst aufeinander folgende Datenpakete zu Abschnitten mit konstanter Rate zusammen. Diese Simulations-Methode bildet die Änderungen einer Rate in der Simulation ab und berechnet nach jeder Rateänderung den Zustand im Netz neu. Durch die Zusammenfassung mehrerer Pakete in Abschnitte konstanter Rate kann der Simulationsaufwand gegenüber der paketbasierten Simulations-Methode reduziert werden. Je größer die mittlere Anzahl an Paketen pro Zeitabschnitt ist, desto größer ist die Steigerung der Simulationsgeschwindigkeit gegenüber der paketbasierten Simulations-Methode.

Die quantitative Analyse der Simulationsgeschwindigkeit der ratenbasierten Simulations-Methode hat gezeigt, dass für die Untersuchung eines Netzes mit 14 Knoten eine mittlere Burst-Länge in der Größenordnung von 100 bis 1000 Paketen erforderlich ist. Durch die Abhängigkeit der Simulationsgeschwindigkeit von der Netzgröße ist für größere Netze ein höherer Mittelwert der Burst-Länge notwendig.

Die Analyse der Genauigkeit der ratenbasierten Simulations-Methode hat ergeben, dass diese die Performanzwerte Paketverlust und Paketverzögerung eines Paketnetzes dann annähert, wenn die modellierte Senderate der tatsächlichen Senderate entspricht. Das ist der Fall, wenn die Pakete eines Bursts unmittelbar aufeinander folgen. Für gemittelte Senderaten liefert die ratenbasierte Simulations-Methode keine relevante Aussage bezüglich der Performanz.

Die Analyse des Verkehrs in IP-Netzen in [TMW97] zeigt, dass die übertragenen IP-Flüsse eine mittlere Größe von 15 Paketen haben. Werden diese unter Verwendung von TCP über das Netz übertragen, erfolgt die Übertragung der Pakete nicht direkt aufeinander. Das ist bedingt durch den Fenstermechanismus des TCP (siehe [S94]). Die sich daraus ergebende mittlere Burst-Länge liegt deutlich unter der Burst-Länge, die die ratenbasierte Simulations-Methode für die Untersuchung von reaktiven Verkehrslenkungsverfahren in Netzen relevanter Größe benötigt.

Als Fazit lässt sich aus der Betrachtung der paketbasierten und ratenbasierten Simulations-Methoden schließen, dass sich beide Methoden nicht für die Untersuchung der reaktiven Lastverteilungsverfahren in größeren IP-Netzen eignen.

4.2 Modifizierte ratenbasierte Simulations-Methode

Die Analyse der beiden vorherigen Simulations-Methoden hat gezeigt, dass sich eine exakte Bestimmung der Netzperformanz nicht mit der skalierbaren Untersuchung von reaktiven Lastverteilungsverfahren in realistischen Netzszenarien vereinbaren lässt. Aus diesem Grund wurde eine Simulations-Methode entwickelt, welche eine Abschätzung der Netzperformanz erlaubt, wenn das Netz sowie der Eingangsverkehr bestimmte Voraussetzungen erfüllen. Die Simulations-Methode erlaubt auf diese Weise den Vergleich verschiedener Lastverteilungsverfahren. Dieser Abschnitt beschreibt diese Simulations-Methode.

Die Bewertung der ratenbasierten Simulations-Methode hat gezeigt, dass sich mit der Zusammenfassung aufeinanderfolgender Pakete zu Abschnitten mit konstanter Rate die Simulationsgeschwindigkeit gegenüber der paketbasierten Simulations-Methode erheblich steigern lässt. Die Steigerung hängt dabei von der Anzahl der Pakete pro Abschnitt mit konstanter Rate ab.

Der Nachteil der in Abschnitt 4.1.3 vorgestellten ratenbasierten Simulations-Methode ist jedoch, dass diese nur dann eine ausreichende Genauigkeit erzielt, wenn Paket-Bursts mit direkt aufeinander folgenden Paketen zu Abschnitten mit konstanter Rate zusammengefasst werden. Solche Paket-Bursts mit größerer Anzahl an Paketen treten in IP-Netzen aber nicht auf.

Um eine Abschätzung der Performanzparameter in IP-Netzen relevanter Größe unter Verwendung der ratenbasierten Simulations-Methode zu ermöglichen, muss die Modellierung von Paketfolgen beliebiger Länge durch Abschnitte mit konstanter Rate zulässig sein. Das bedeutet, dass eine ratenbasierte Simulations-Methode benötigt wird, welche eine Abschätzung der Performanz in IP-Netzen auf der Basis gemittelter Senderaten erlaubt. Das schließt insbesondere auch die Abschätzung der Performanz des TCP-Verkehrs mit ein.

Dieser Abschnitt stellt die Realisierung einer modifizierten ratenbasierten Simulations-Methode vor, welche die geforderten Bedingungen einhält. Diese Simulations-Methode ba-

siert auf der Modellierung und Performanzabschätzung von aggregierten Verkehrsströmen, welche sich aus mehreren Einzel-Verbindungen zusammensetzen.

Die beiden folgenden Abschnitte 4.2.1 und 4.2.2 gehen zunächst auf die Modellierungsaspekte der Netzkomponenten Puffer und Quelle ein. Die Modellierung der übrigen Netzkomponenten ist zur Beschreibung und Bewertung der modifizierten ratenbasierten Simulations-Methode nicht relevant. Der Abschnitt 4.2.3 beschreibt die Abschätzung der Performanz-Parameter. Die anschließenden Abschnitte 4.2.4 und 4.2.5 bewerten die Genauigkeit sowie die Skalierbarkeit der modifizierten ratenbasierten Simulations-Methode.

4.2.1 Modellierung eines Puffers

Wie in Abbildung 37 dargestellt, wird ein Puffer in der modifizierten ratenbasierten Simulations-Methode durch ein Wartesystem mit endlicher Anzahl an Warteplätzen modelliert. Ein solches Wartesystem kann wie in der Abbildung angegeben z.B. ein M/M/1/S-Wartesystem mit S Warteplätzen sein. Das ausgewählte Wartesystem sollte an die Verteilung der Paketankünfte sowie an die Paketlängenverteilung angepasst werden.

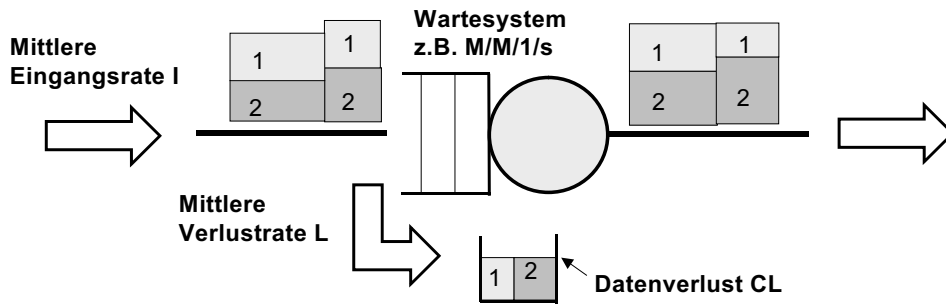


Abbildung 37: Puffer-Modellierung der modifizierten ratenbasierten Simulations-Methode

Ändert sich die mittlere Eingangsrate I_m an einem Puffer, wird mit Hilfe des Wartesystems eine neue mittlere Verlustrate L_m sowie eine neue mittlere Ausgangsrate O_m berechnet. Die Berechnung der Verlustrate L_m basiert auf der Berechnung der Paketverlustwahrscheinlichkeit P_{vw} , welche für das M/M/1/S-Wartesystem in Gleichung 58 gegeben ist. Gleichung 59 zeigt die Berechnung der Verlustrate.

$$P_{vw} = \frac{\left(\frac{I_m}{O_{max}}\right)^{S+1} \cdot \left(1 - \frac{I_m}{O_{max}}\right)}{1 - \left(\frac{I_m}{O_{max}}\right)^{S+2}} \quad (58)$$

$$L_m = P_{vw} \cdot I_m \quad (59)$$

Die Berechnung der mittleren Ausgangsrate O_m ergibt sich aus der Differenz der mittleren Eingangsrate I_m und der mittleren Verlustrate L_m (siehe Gleichung 60).

$$O_m = I_m - L_m \quad (60)$$

Die Neuberechnung der mittleren Verlustrate und der mittleren Ausgangsrate erfolgt direkt bei der Ankunft der Ratenänderung der mittleren Eingangsrate. Das bedeutet, dass

sich die Ankunft einer Ratenänderung ohne Zeitverzögerung am Ausgang des Puffers auswirkt. Der Übergang von einer mittleren Senderate auf die folgende mittlere Senderate wird vernachlässigt. Das bedeutet zudem, dass in der modifizierten ratenbasierten Simulations-Methode keine Propagations-Ereignisse erzeugt werden.

Werden mehrere Verkehrsströme über einen Puffer übertragen, werden die mittlere Verlustrate sowie die mittlere Ausgangsrate proportional auf die Verkehrsströme aufgeteilt. Die proportionale Aufteilung der Pufferkapazität auf die Verkehrsströme bedeutet, dass bei jeder Ratenänderung eines Verkehrsstroms die Raten sämtlicher Verkehrsströme des Puffers angepasst werden müssen. Dieser unerwünschte Effekt kann insbesondere in großen Netzen eine erhebliche Reduktion der Simulationsgeschwindigkeit bedeuten, da sich eine Ratenänderung eines Verkehrsstroms über das gesamte Netz auf sämtliche Verkehrsströme fortpflanzt (siehe hierzu den Abschnitt 4.1.3.5).

Um den Einfluss dieses Effekts einzuschränken, werden die Ausgangsraten in der vorgestellten Realisierung nur dann angepasst, wenn die Summe der Eingangsraten die maximale Ausgangskapazität überschreitet. Der Einfluss dieser Vereinfachung auf die Genauigkeit der Simulations-Methode wird in dem Abschnitt 4.2.4 beschrieben.

4.2.2 Modellierung einer Quelle

Dieser Abschnitt beschreibt die Modellierung einer Quelle in der modifizierten ratenbasierten Simulationsumgebung. Insbesondere geht dieser Abschnitt auf die Modellierung der Eigenschaften des TCP-Verkehrs ein.

Wie bereits erwähnt, basiert die modifizierte ratenbasierte Simulationsumgebung auf der Abbildung von Paketströmen auf mittlere Senderaten. Ein Paketstrom entspricht dem Verkehr, der durch die Überlagerung mehrerer einzelner Verbindungen entsteht. Hierfür wird ein Quellenmodell benötigt, welches variable Raten erzeugt, die über Zeitabschnitte variabler Länge konstant sind. Diesen Anforderungen entspricht das in Abschnitt 4.1.3.2 beschriebene allgemeine Quellenmodell mit variabler Senderate.

Ein weiterer wichtiger Aspekt des Quellenmodells ist die Berücksichtigung der Eigenschaften von TCP-Verkehr. Das Quellenmodell der modifizierten ratenbasierten Simulations-Methode bildet das Verhalten von aggregierten Verkehrsströmen im eingeschwungenen Zustand ab. Aggregierte Verkehrsströme setzen sich aus mehreren TCP-Verbindungen zusammen.

Zur Erläuterung dieser Modellierung zeigt Abbildung 38 zunächst das prinzipielle Verhalten einer UDP-Quelle und einer TCP-Quelle.

Eine UDP-Quelle sendet Daten in das Netz, ohne den Zustand des Netzes zu berücksichtigen. Die Senderate wird ausschließlich durch die Quelle vorgegeben.

Eine TCP-Quelle hingegen passt die Senderate an den Zustand des Netzes an. Hierfür ist die Flusskontrolle des TCP verantwortlich. Die Flusskontrolle wird durch die TCP-Mechanismen SlowStart, Timeout, Congestion Avoidance, FastRetransmit sowie FastRecovery vorgegeben. Diese Mechanismen sind in [S94] beschrieben.

Bei der Modellierung des Verhaltens von TCP im eingeschwungenen Zustand wird in der Literatur davon ausgegangen, dass sich die einzelnen TCP-Flüsse die Übertragungskapazität an einem Puffer zu gleichen Anteilen aufteilen (siehe hierzu auch [RBP00]). Ist

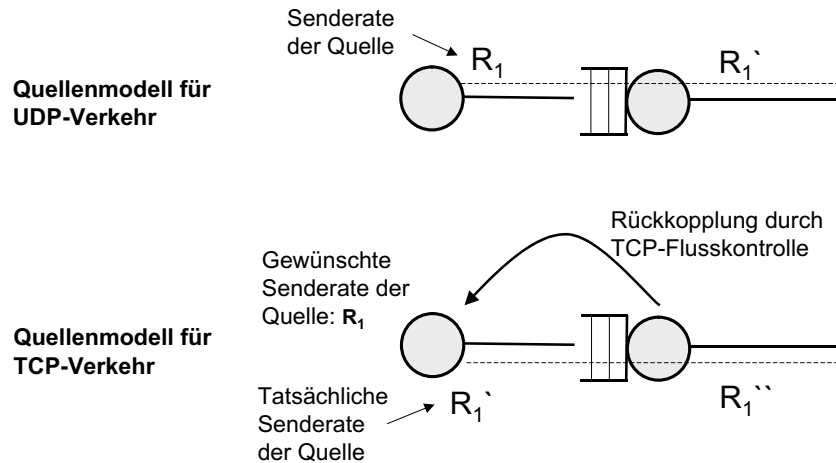


Abbildung 38: Vergleich zwischen dem UDP- und dem TCP-Verkehrsmodell

die Summe der maximalen Senderaten der einzelnen TCP-Flüsse kleiner als die Übertragungskapazität des Puffers, so senden alle TCP-Quellen mit der maximalen Senderate. Ist hingegen die Summe der maximalen Senderaten der einzelnen TCP-Flüsse größer als die Ausgangskapazität des Puffers, so passen sich die Senderaten der TCP-Quellen entsprechend an.

Bei Betrachtung von aggregierten Verkehrsströmen, welche sich aus mehreren TCP-Verbindungen zusammensetzen, geht das verwendete Quellenmodell davon aus, dass diese sich die Ausgangskapazität an einem Puffer proportional aufteilen. Diesem Vorgehen liegt die zulässige Annahme zu Grunde, dass sich die mittlere Senderate sowie die Anzahl der TCP-Verbindungen eines aggregierten Verkehrsstroms proportional zueinander verhalten.

Bei netzweiter Anpassung der Senderaten der TCP-Quellen muss der Fall berücksichtigt werden, dass mehrere Puffer im Netz überlastet sind. Eine Anpassung der Senderaten der TCP-Quellen an einem Puffer beeinflusst wiederum die Auslastungen anderer Puffer und somit die Senderaten der über diesen Puffer übertragenen TCP-Flüsse. Aus diesem Grund müssen in dem verwendeten Ansatz die Anpassungen der Senderaten der einzelnen TCP-Quellen koordiniert werden.

Hierbei geht der Ansatz der Quellenmodellierung davon aus, dass die Senderate eines TCP-Flusses im wesentlichen durch den am stärksten überlasteten Puffer auf dem Pfad des TCP-Flusses beeinflusst wird. Das entspricht der Annahme, die in dem vergleichbaren Ansatz von Ma, Steenkiste und Zhang in [MSZ96] getroffen wird. Der Algorithmus zur netzweiten koordinierten Anpassung der Senderaten der Quellen mit aggregiertem TCP-Verkehr läuft wie folgt ab:

1. Zu Beginn liegen dem Algorithmus die gewünschten Senderaten aller TCP-Quellen ohne Ratenanpassung vor.
2. Es wird der Puffer ausgewählt, der die relative größte Eingangslast $\rho = \frac{\sum L_i}{C_{in_k}}$ hat.
3. Ist die relative Auslastung $\rho \leq 1$, dann endet der Algorithmus.
4. Gilt für die relative Auslastung des ausgewählten Puffers $\rho > 1$, dann passt der Algorithmus die Senderaten der Quellen L_i , die über diesen Puffer geroutet sind, proportional an: $L_{ineu} = \frac{\sum L_i}{\rho}$.

5. Die angepassten Senderaten der TCP-Quellen mit Ratenanpassung L_{ineu} werden von den Ausgangskapazitäten C_{link} der übrigen Puffer subtrahiert.
6. Die TCP-Quellen mit angepasster Senderate werden aus der Liste der anzupassenden TCP-Quellen entfernt. Ein einmal angepasste Senderate einer TCP-Quelle wird bei einer koordinierten Anpassung kein zweites Mal angepasst.
7. Der ausgewählte Puffer mit der zuvor größten relativen Auslastung wird aus der Liste der Puffer entfernt. Anschließend springt der Algorithmus zurück zu Schritt 2.

4.2.3 Bestimmung der Performanzparameter

Zu den in diesem Abschnitt betrachteten Performanzparametern zählen der Paketverlust, die Paketverzögerung sowie der Durchsatz von TCP-Verbindungen.

Die mittlere Paketverlustwahrscheinlichkeit eines Puffers errechnet sich mit Hilfe des verwendeten Wartesystems. Im Fall des M/M/1/S-Wartesystems berechnet sich die Paketverlustwahrscheinlichkeit an einem Puffer mit der Gleichung 58. Die Paketverlustwahrscheinlichkeit $P_{vw\ pfad}$ auf einem Pfad R durch das Netz berechnet sich aus den Paketverlustwahrscheinlichkeiten der einzelnen Puffer $P_{vw\ i}$. Die Formel zur Berechnung stellt Gleichung 61 dar.

$$P_{vw\ pfad} = 1 - \prod_{i \in R} (1 - P_{vw\ i}) \quad (61)$$

Wie die mittlere Paketverlustwahrscheinlichkeit berechnet sich die Paketverzögerung an einem Puffer ebenfalls mit Hilfe des verwendeten Wartesystems. Die mittleren Paketverzögerung T_m an einem Puffer setzt sich aus der mittleren Wartedauer sowie der mittleren Aussendedauer zusammen. Im Fall eines M/M/1/S-Wartesystems ergibt sich die mittlere Wartedauer an einem Puffer aus der mittleren Warteschlangenlänge Ω (siehe Gleichung 62) und der Ausgangskapazität O_{max} . Die mittlere Aussendedauer ergibt sich aus der mittleren Paketlänge b_{mpl} sowie der Ausgangskapazität O_{max} . Die Berechnung der mittleren Paketverzögerung D_{puffer} zeigt Gleichung 63.

$$\Omega = \frac{\left(\frac{I_m}{O_{max}}\right)^2 \cdot \left(1 - \left(\frac{I_m}{O_{max}}\right)^S\right)}{1 - \left(\frac{I_m}{O_{max}}\right)^{S+2}} \cdot \left(\frac{1}{1 - \left(\frac{I_m}{O_{max}}\right)} - \frac{S \cdot \left(\frac{I_m}{O_{max}}\right)^S}{1 - \left(\frac{I_m}{O_{max}}\right)^S}\right) \quad (62)$$

$$D_{puffer} = \frac{\Omega}{O_{max}} + \frac{b_{mpl}}{O_{max}} \quad (63)$$

Die Paketverzögerung der Datenpakete eines Verkehrsstroms durch das Netz berechnet sich aus der Summe der mittleren Verzögerungen an den einzelnen Puffern sowie aus den Verzögerungen auf den Links im Netz. Wie in Gleichung 64 dargestellt, ergibt sich die Verzögerung auf einem Link D_{link} aus der Länge $x_{länge}$ sowie der Signalausbreitungsgeschwindigkeit v_{sp} des Links. Die Gleichung 65 gibt die Berechnungsvorschrift der mittleren Pfadverzögerung an.

$$D_{link} = \frac{x_{länge}}{v_{sp}} \quad (64)$$

$$D_{pfad} = \sum_{i \in R} D_{puffer\ i} + \sum_{i \in L} D_{link\ i} \quad (65)$$

Zur Evaluierung der Performanz des TCP-Verkehrs bestimmt die modifizierte ratenbasierte Simulations-Methode den mittleren Durchsatz einer TCP-Verbindung im eingeschwungenen Zustand. Die betrachtete TCP-Verbindung startet und endet dabei jeweils an einem Randknoten des simulierten Netzes. Dieser Betrachtungsweise liegt die Tatsache zu Grunde, dass der Durchsatz einer TCP-Verbindung von dem gesamten Pfad abhängt und der Einfluss durch andere nicht simulierte Teile des Pfades (z.B. Übertragung über ein Zugangsnetz) nicht bekannt ist.

Das verwendete Modell zur Bestimmung des TCP-Durchsatzes von Mathis et al. [MSM97] basiert auf der Kenntnis der Umlaufzeit RTT der Pakete zwischen Quelle und Senke sowie der Paketverlustwahrscheinlichkeit $P_{vw\ pfad}$ auf dem Pfad. Es geht von der Annahme einer konstanten Paketumlaufzeit sowie in konstanten Abständen auftretenden Paketverlusten aus. Ferner beschränkt es sich auf die Modellierung des Congestion Avoidance Algorithmus (siehe hierzu auch [S94]). Durch Vergleich mit Messungen und Simulationen weisen Mathis et al. nach, dass das Modell unter vielen Bedingungen eine gute Vorhersage des TCP-Durchsatzes erzielt. Die Formel zur Berechnung der mittleren Durchsatzes einer TCP-Verbindung zeigt Gleichung 66.

$$T_{tcp} = \frac{1}{RTT \cdot \sqrt{\frac{2 \cdot P_{vw\ pfad}}{3}}} \quad (66)$$

Es ist durchaus möglich, andere Modelle zur Bestimmung des TCP-Durchsatzes zu verwenden. Hier bietet sich unter anderem das Modell von Padhye et al. [PFT98] an, welches das Modell von Mathis et al. um die Berücksichtigung des TCP-Timeout-Mechanismus erweitert.

4.2.4 Analyse der Genauigkeit

Grundsätzlich lässt eine Simulation nur dann eine exakte Performanz-Analyse eines IP-Netzes zu, wenn diese auf einer genauen Modellierung der Netzkomponenten und des Verkehrs beruht. Die Analyse des Stands der Technik in Abschnitt 4.1 hat jedoch gezeigt, dass eine solche genaue Modellierung keine Untersuchung der reaktiven Verkehrslenkungsverfahren in Netzen relevanter Größe erlaubt. Um die reaktiven Verkehrslenkungsverfahren dennoch in realistischen Netzzenarien bewerten zu können, wurde die modifizierte ratenbasierte Simulations-Methode unter einer gelockerten Anforderung bezüglich der Simulationsgenauigkeit entwickelt. Die Anforderung an die Genauigkeit ist, dass die Simulations-Methode eine realistische Abschätzung der Netzperformanz erlauben muss, wenn das Netz- und Auslastungsszenario bestimmte Randbedingungen erfüllt. Dieser Abschnitt beschreibt die Randbedingungen.

Eine vereinfachende Annahme der modifizierten ratenbasierten Simulations-Methode ist die Modellierung des Verhaltens eines Puffers durch ein M/M/1/S-Wartesystem. Diese Vereinfachung setzt eine negative exponentielle Verteilung der Paketankünfte sowie der Paketlängenverteilung voraus.

Das Prinzip der Simulations-Methode basiert auf einer vereinfachten Betrachtung, welche die Unabhängigkeit der Puffer im Netz voraussetzt. Hierdurch wird die Korrelation

des Verkehrs vernachlässigt, welche sich durch die Übertragung des Verkehrs über mehrere Puffer ergibt. Nach der Kleinrockschen Unabhängigkeits-Approximation ist die Vernachlässigung der Korrelation dann zulässig, wenn ein Netz ausreichend vermascht ist und sich durch die Vermischung mehrerer korrelierter Verkehrsströme an einem Puffer näherungsweise wieder eine negative exponentielle Verteilung ergibt [BG87].

Das verwendete Modell zur Bestimmung des TCP-Durchsatzes beruht auf einer vereinfachenden Betrachtung des Verhaltens von TCP-Verbindungen. Die Zulässigkeit dieser Betrachtungsweise werden durch die in [MSM97] durchgeführten Untersuchungen für viele Fälle bestätigt.

Eine weitere Vereinfachung der Simulations-Methode ist die Vernachlässigung der Übergangsphase zwischen zwei mittleren Übertragungsraten. Diese Vereinfachung ist dann zulässig, wenn der mittlere Abstand zwischen zwei Ratenänderungen groß gegenüber der Dauer der Übergangsphase ist.

Die Vernachlässigung des Paketverlusts bei der Bestimmung der Ausgangsrate an nicht überlasteten Puffern führt zu einer Überschätzung der Übertragungsraten im Netz. Der Einfluss dieser vereinfachenden Betrachtung ist unbedeutend, wenn der dadurch entstehende zusätzliche Paketverlust sowie die zusätzliche Paketververzögerung an den nachfolgenden Puffern vernachlässigbar ist. Diese Annahme trifft in den meisten Fällen zu.

4.2.5 Analyse der Skalierbarkeit

Die modifizierte ratenbasierte Simulations-Methode weist eine vergleichbare Struktur zur der in Abschnitt 4.1.2 beschriebenen ratenbasierten Simulations-Methode auf. Aus diesem Grund beschreibt dieser Abschnitt einen quantitativen Vergleich zwischen der Simulationsgeschwindigkeit beider Methoden.

Die Unterschiede zwischen beiden Methoden bezogen auf die Simulationskomplexität sind gering. Die Simulationsgeschwindigkeit der modifizierten ratenbasierten Simulation hängt ebenfalls von der Anzahl der Pakete pro Zeitabschnitt mit konstanter Rate ab.

Zur Evaluierung der Simulationsgeschwindigkeit der modifizierten ratenbasierten Simulations-Methode wird ein selbst entwickelter Simulator verwendet. Dieser wird mit dem in Abschnitt 4.1.3.5 verwendeten ratenbasierten Simulator verglichen. Zum Vergleich der beiden Ansätze wird das in Abschnitt 4.1.2.3 beschriebene Szenario verwendet.

Das Ergebnis des Vergleichs beider ratenbasierter Simulations-Methoden über den mittleren zeitlichen Abstand zwischen zwei Ratenänderungen zeigt Abbildung 39. Es ist zu erkennen, dass beide Simulations-Methoden unter gleichen Bedingungen eine vergleichbare Simulationsgeschwindigkeit haben. Durch die vereinfachte Modellierung des Puffers ist die modifizierte ratenbasierte Methode geringfügig schneller.

Aus der in diesem Abschnitt vorgestellten Analyse der modifizierten ratenbasierten Variante zeigt sich, dass diese unter der Annahme der gleichen Voraussetzungen eine vergleichbare Simulationsgeschwindigkeit zu der in Abschnitt 4.1.3 beschriebenen ratenbasierten Methode hat.

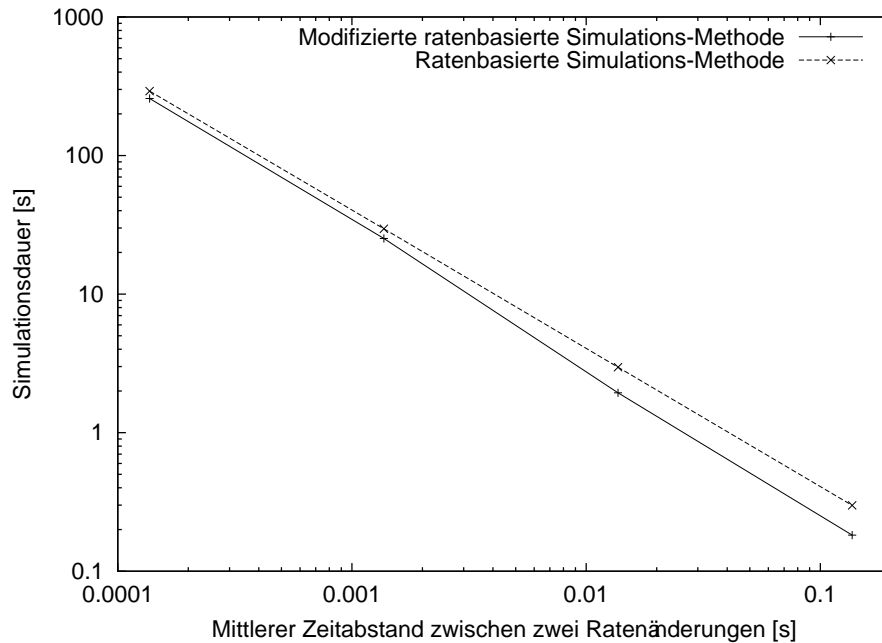


Abbildung 39: Vergleich der Simulationsdauer der ratenbasierten Simulations-Methoden

4.3 Zusammenfassung wichtiger Aspekte der Simulations-Methodik

Dieser Abschnitt fasst die Erkenntnisse des Kapitels 4 über geeignete Methoden zur Untersuchung von reaktiven Verkehrslenkungsverfahren zusammen. Das Kapitel beschränkte sich auf die Betrachtung von Simulations-Methoden.

Die im ersten Teil des Kapitels durchgeführte Analyse des Stands der Technik beschreibt und bewertet zwei bekannte Simulations-Methoden. Die eine Methode basiert auf der paketbasierten Modellierung des Datenverkehrs. Diese Methode ist der Standard zur Analyse der Performanz in Paketnetzen. Die zweite Methode verwendet eine Modellierung des Datenverkehrs auf Ratenebene.

Das Kapitel bewertet die beiden Methoden hinsichtlich der Anforderungen, welche sich bei der Analyse von reaktiven Verkehrslenkungsverfahren stellen. Zum einen soll die Simulations-Methode eine Untersuchung der reaktiven Verkehrslenkungsverfahren in realistischen Netzszenarien ermöglichen, welche der Größe und Topologie typischer IP-Netze entspricht. Zum anderen stellt sich die Anforderung an die Simulations-Methode, eine realistische Abschätzung der Performanz eines IP-Netzes unter Berücksichtigung der Eigenschaften von TCP-Verkehr zu ermöglichen.

Die Analyse der paketbasierten Simulations-Methode zeigt, dass sich mit dieser Methode zwar die maximale Genauigkeit bei der Bewertung der Performanz in Paketnetzen erzielen lässt, jedoch die Größe der zu untersuchenden Netzszenarien eingeschränkt ist. Für die Analyse der reaktiven Verkehrslenkungsverfahren in realistischen Netzszenarien eignet sich diese Methode nicht.

Die untersuchte ratenbasierte Simulations-Methode modelliert aufeinander folgende Datenpakete mit einer Rate, die über einen Zeitabschnitt variabler Länge konstant ist. Die Bewertung der Genauigkeit dieser Methode hat gezeigt, dass diese bei der Modellierung direkt aufeinanderfolgender Pakete zu einer Rate eine realistische Abschätzung der Performanz erlaubt. Die Untersuchung der reaktiven Verkehrslenkungsverfahren in realisti-

schen Netzen mit dieser Methode ist dann möglich, wenn die mittlere Anzahl der Pakete pro Burst ausreichend groß ist. Die quantitative Analyse der Simulationsgeschwindigkeit hat ergeben, dass zur Untersuchung eines 14-Knoten-Netzes die mittlere Burst-Länge in der Größenordnung von 100 bis 1000 Paketen liegen muss. Für größere Netze muss dieser Wert entsprechend größer sein. Untersuchungen des IP-Verkehrs zeigen jedoch, dass Paket-Bursts solcher Größe in IP-Netzen nicht auftreten. Somit eignet sich auch diese Methode nicht zur Untersuchung der reaktiven Verkehrslenkungsverfahren.

Aus diesem Grund beschreibt der zweite Teil dieses Kapitels eine Simulations-Methode, welche im Rahmen dieser Arbeit entwickelt wurde. Diese basiert ebenfalls auf der ratenbasierten Modellierung des Datenverkehrs. Im Gegensatz zu der anderen ratenbasierten Simulations-Methode modelliert diese Methode den Datenverkehr im Netz durch mittlere Senderaten. Dies erlaubt die Zusammenfassung der Übertragung beliebig langer Paketfolgen zu einer mittleren Senderate, sofern die Paketankünfte und die Paketlängenverteilung innerhalb der Paketfolge den angenommenen Verteilungen an den Puffern im Netz entsprechen. Die in dieser Arbeit vorgeschlagene Variante modelliert einen Puffer durch ein $M/M/1/S$ Wartesystem. Ferner muss das untersuchte Netzscenario die unabhängige Betrachtung der einzelnen Puffer im Netz erlauben. Unter Voraussetzung der Gültigkeit dieser Annahmen für ein Netz- und Auslastungsszenario lässt sich bedingt durch die Zusammenfassung beliebig langer Paketfolgen zu einer mittleren Senderate eine erhebliche Steigerung der Simulationsgeschwindigkeit erzielen. Das erlaubt die Bewertung der reaktiven Verkehrslenkungsverfahren, in welchen das Netz- und Auslastungsszenario die notwendigen Voraussetzungen zum Einsatz der Simulations-Methode erfüllt.

5 Performanz-Analyse anhand ausgewählter Fallstudien

Der Inhalt dieses Kapitels ist die quantitative Analyse der im Rahmen dieser Arbeit entwickelten reaktiven verteilten Verkehrslenkungsverfahren. Hierzu wird eine Simulations-Umgebung verwendet, welche auf der in Kapitel 4 beschriebenen modifizierten ratenbasierten Simulations-Methode beruht. Die in der Folge vorgestellte quantitative Analyse der Verfahren setzt mehrere Schwerpunkte.

Der erste Schwerpunkt ist die Bewertung der Netzperformanz, welche sich nach der Durchführung der Lastverteilung durch die reaktiven Verkehrslenkungsverfahren ergibt. Daher präsentiert der in Abschnitt 5.1 vorgestellte Teil der Untersuchungen die Netzperformanz anhand von Fallstudien für vier verschiedene Netzszenarien. Zur Evaluierung des Verfahrens wird dieses mit der optimalen Lastverteilung und mit dem Shortest Path Routing verglichen.

Die Untersuchung der Konvergenzdauer der reaktiven verteilten Verkehrslenkungsverfahren ist der zweite Schwerpunkt der quantitativen Analyse. Hierbei werden zwei Aspekte berücksichtigt. Einerseits sind die reaktiven Verkehrslenkungsverfahren vor dem Hintergrund entwickelt worden, die geringe Konvergenzgeschwindigkeit des Konkurrenz-Verfahrens MATE zu verbessern. Diesen Punkt gilt es im Rahmen der Untersuchung nachzuweisen. Andererseits gilt es, die Abhängigkeit der Konvergenzdauer von dem Netz- und Auslastungsszenario quantitativ zu analysieren. Abschnitt 5.2 präsentiert die Ergebnisse der Analyse der Konvergenzdauer, welche anhand von zwei Netzszenarien unterschiedlicher Größe durchgeführt werden.

Der dritte Schwerpunkt der quantitativen Analyse setzt sich mit der reaktiven verteilten Verkehrslenkungsverfahren bei dynamischem Eingangsverkehr auseinander. Hierbei gilt es, das Konvergenzverhalten des Verfahrens bei dynamischem Verkehr zu analysieren und mit den in Abschnitt 2.4.2 beschriebenen proaktiven Verkehrslenkungsverfahren zu vergleichen. Diese Untersuchung werden in Abschnitt 5.3 erläutert.

Die Ergebnisse der quantitativen Analyse der Verfahren werden am Ende des Kapitels in Abschnitt 5.4 zusammengefasst.

5.1 Analyse der Netzperformanz

Dieser Abschnitt bewertet die Netzperformanz, welche sich mit dem in dieser Arbeit vorgestellten reaktiven Verkehrslenkungsverfahren erzielen lässt und vergleicht diese mit der optimalen Lastverteilung und dem Shortest Path Routing.

Zur Evaluierung der Netzperformanz wird bei diesen Untersuchungen davon ausgegangen, dass ausschließlich TCP-Verkehr im Netz übertragen wird. Um den Einfluss der Verkehrsdynamik auf die Netzperformanz zunächst auszuschließen, erfolgen die Untersuchungen in diesem Abschnitt mit konstantem Eingangsverkehr. Zur Bestimmung des konstanten Eingangsverkehrs wird pro Eingangs-Ausgangsknoten-Paar eine Rate unter Verwendung einer Normalverteilung bestimmt. Der Mittelwert und die Varianz der Normalverteilung ist Szenario-abhängig.

Zur Bewertung der reaktiven Verkehrslenkungsverfahren wird einerseits die maximale Link-Auslastung im Netz und andererseits der minimale Durchsatz einer TCP-Verbindung

im Netz bestimmt. Der angegebene Wert des TCP Durchsatzes wird ausschließlich durch das Netz vorgegeben. Die maximale Senderate einer TCP-Quelle wird nicht berücksichtigt. Die dargestellten Ergebnisse entsprechen dem gemittelten Werten über 100 Simulationläufen.

Zur Bewertung der Verfahren werden die reaktiven Verkehrslenkungsverfahren mit der optimalen Lastverteilung verglichen. Zur Bestimmung der optimalen Lastverteilung wird im Rahmen dieser Arbeit ein lineares mathematisches Programm definiert, welches in Abschnitt 5.1.1 beschrieben wird. Die anschließenden Abschnitte 5.1.2 bis 5.1.5 präsentieren die Ergebnisse der Untersuchungen in vier Netzszenarien.

5.1.1 Definition der optimalen Lastverteilung

Zur Bestimmung der optimalen Lastverteilung wird ein lineares mathematisches Programm verwendet, welches die Lastverteilung mit minimalen Gesamtkosten berechnet. Die Formulierung des Programms erläutert dieser Abschnitt.

Um die Berechnung der optimalen Lastverteilung mit der in Abschnitt 3.4.1 beschriebenen Kostenfunktion durch ein lineares mathematisches Programm zu realisieren, muss zunächst eine Linearisierung der Kostenfunktion erfolgen. Hierzu werden einzelne Punkte der Kostenfunktion mit Geraden verbunden, so dass eine stückweise lineare Kostenfunktion entsteht. Die Auswahl der Punkte der Kostenfunktion ist in Abbildung 40 dargestellt. Die linearisierte Kostenfunktion startet bei dem Wert null. Die in Abbildung 40 dargestellte Schrittweite gibt den Abstand der Punkte zueinander in den einzelnen Auslastungsbereichen an. Die kleiner werdende Schrittweite bei höherer Auslastung erlaubt eine genauere Abbildung der Linkkosten in den Bereichen, in denen die Kostenfunktion stärker ansteigt.

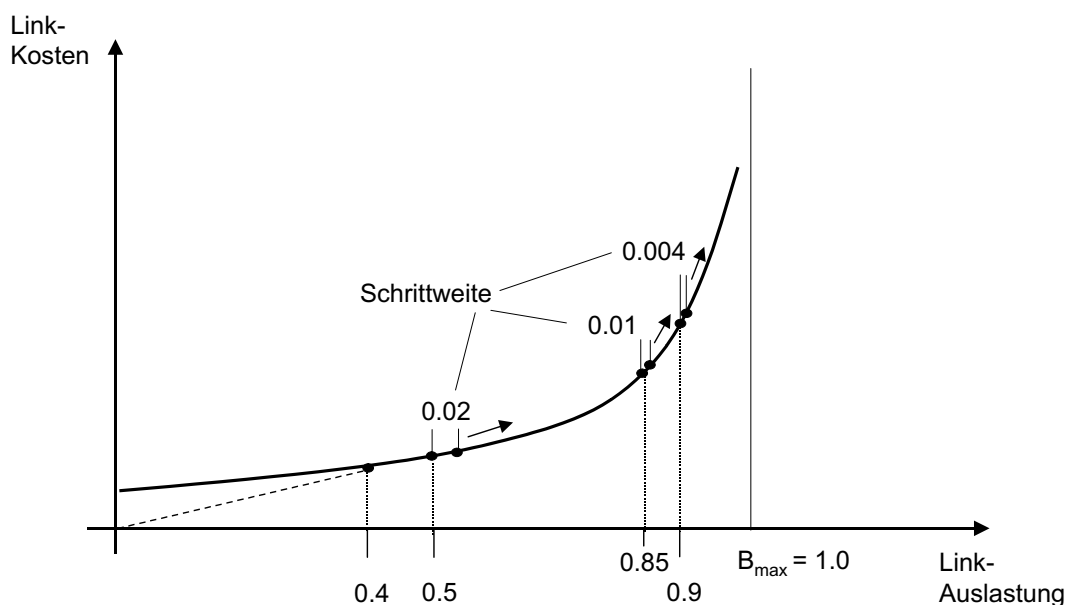


Abbildung 40: Linearisierung der Linkkosten

Zur Beschreibung des Lastverteilungsproblems als lineares mathematisches Programm werden die Menge der Knoten N und die Menge der Links E definiert. L_e entspricht der aktuellen Auslastung und $L_{e \max}$ der Übertragungskapazität des Links. $L_{lsp}(x, y)$ ist die

Last des Verkehrsstroms zwischen dem Eingangsknoten x und dem Ausgangsknoten y . $L_{lsp}(x, y, i, j)$ entspricht dem Anteil der Last der Verkehrsbeziehung zwischen Eingangsknoten x und Ausgangsknoten y auf dem Link zwischen Knoten i und j .

Die Zielfunktion des linearen Programms, welche die folgende Gleichung 67 zeigt, ist die Minimierung der gesamten Linkkosten. Das entspricht der Zielfunktion des definierten Verkehrslenkungsverfahrens.

$$\min(\sum_{e \in E} C(L_e)) \quad (67)$$

Die Gleichungen 68 bis 72 definieren die Nebenbedingungen des linearen Programms. Nebenbedingungen 68 und 69 geben vor, dass die Auslastung der Links nicht negativ und nicht größer als die maximale Linkkapazität sein darf. Die weiteren Nebenbedingungen 70 bis 72 definieren, dass die Summe der an einem Knoten ankommenden Last einer Verkehrsbeziehung gleich der Summe der abgehenden Last der Verkehrsbeziehung sein muss. Während die Nebenbedingungen 70 und 71 diese Anforderung für den Eingangsknoten und den Ausgangsknoten der Verkehrsbeziehung festlegen, beschreibt Nebenbedingung 72 diese Anforderung für die übrigen Knoten.

$$L_e \leq L_e \text{ max} \quad (68)$$

$$L_e \geq 0.0 \quad (69)$$

$$L_{lsp}(x, y) = \sum_{k \in N} L_{lsp}(x, y, x, k) \quad (70)$$

$$L_{lsp}(x, y) = \sum_{k \in N} L_{lsp}(x, y, k, y) \quad (71)$$

$$\sum_{i \in N} L_{lsp}(x, y, i, j) = \sum_{k \in N} L_{lsp}(x, y, j, k) \quad (72)$$

Zur Bestimmung der Paketverlustwahrscheinlichkeit sowie der Paketverzögerung eines LSPs wird die Summe der gewichteten Paketverluste bzw. Paketverzögerungen auf den Links berechnet, über welche der LSP übertragen wird. Die Wichtung ergibt sich aus dem Verhältnis der gesamten LSP-Last zu dem Anteil der LSP-Last, welcher über den jeweiligen Link übertragen wird.

$$P_{vw \text{ p}fad} = \sum_{i \in E} P_{vw \ i} \cdot \frac{L_{lsp}(x, y, i, j)}{L_{lsp}(x, y)} \quad (73)$$

$$D_{v \text{ p}fad} = \sum_{i \in E} D_{v \ i} \cdot \frac{L_{lsp}(x, y, i, j)}{L_{lsp}(x, y)} \quad (74)$$

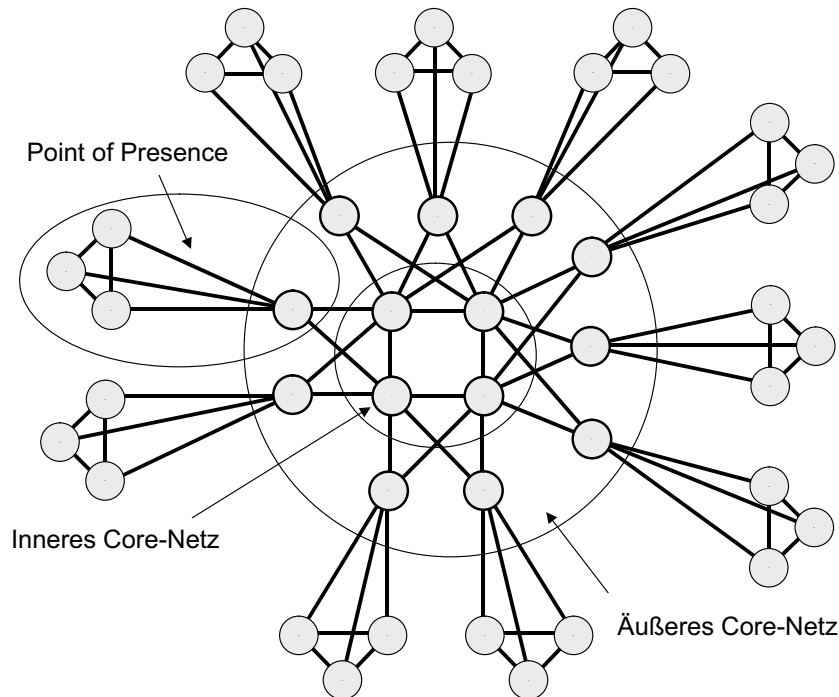


Abbildung 41: Topologie des 44Knoten-Netztes

5.1.2 Untersuchung eines 44Knoten-Netztes

Dieser Abschnitt präsentiert die Ergebnisse der Untersuchung der reaktiven Verkehrslenkungsverfahren an dem 44Knoten-Netz, welches in Abbildung 41 dargestellt ist.

Das Netz besteht aus 44 LSRs und 84 Links. Die Eingangs- bzw. Ausgangs-LSRs des Netztes sind die 30 LSRs, welche nicht zum Core-Netz gehören. Zur einfachen Vollvermaschung dieser LSRs sind im Pfadumlegungsansatz 870 LSPs notwendig. Im Mehrpfadansatz werden bei zwei LSPs pro LSP-Fächer 1740 LSPs und bei drei LSPs pro LSP-Fächer 2610 LSPs geschaltet. Die Links haben eine Übertragungskapazität von $2.5 \frac{G\text{Bit}}{s}$. Die Länge der Links zwischen den LSRs eines POPs ist 10 Meter und die Länge aller weiteren Links ist 200km.

In der Folge werden die Untersuchungen des Netztes in drei verschiedenen Szenarien vorgestellt. Im ersten Szenario verteilt sich die Last gleichmäßig auf alle LSPs bzw. LSP-Fächer. Die initiale Verteilung der Last vor dem Start der Lastverteilung durch die reaktiven Verfahren entspricht dem des Shortest Path Routings. Die Ergebnisse präsentiert Abschnitt 5.1.2.1. Im zweiten Szenario wird die Lastverteilung dahingehend variiert, dass sich die Last nur auf wenige LSPs verteilt. Hiermit wird der Einfluss der Last-Granularität der LSPs auf die Lastverteilung untersucht. Die initiale Verteilung der Last erfolgt wie im ersten Szenario durch das Shortest Path Routing. Die Ergebnisse der Untersuchung fasst Abschnitt 5.1.2.2 zusammen. Das dritte Szenario setzt sich mit der Analyse des Einflusses der initialen Netzauslastung auseinander. Hiermit wird der Einfluss der initialen Lastverteilung auf die Performanz des reaktiven Verkehrslenkungsverfahrens untersucht. Abschnitt 5.1.2.3 präsentiert die Ergebnisse dieser Untersuchung.

5.1.2.1 Auswertung des Gleichlast-Szenarios Das Gleichlast-Szenario zeichnet sich dadurch aus, dass die Varianz und der Mittelwert aller normal-verteilten Quellen pro

Untersuchung gleich sind. Während der Mittelwert in den verschiedenen Untersuchungen zwischen $30 \frac{MBit}{s}$ und $50 \frac{MBit}{s}$ variiert, entspricht die Varianz der Quellenlast $120 \frac{MBit}{s}$.

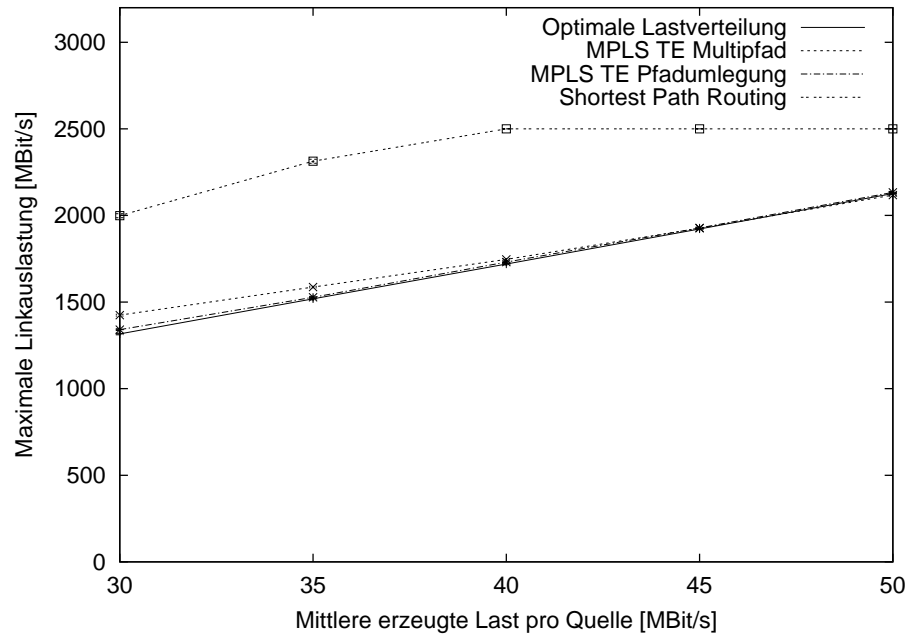


Abbildung 42: Maximale Linkauslastung im Gleichlast-Szenario

Die Abbildungen 42 und 43 stellen die maximale Linkauslastung sowie den minimalen Durchsatz einer TCP Verbindung im Netz in Abhängigkeit zur mittleren erzeugten Last pro Quelle dar. Die reaktiven Verkehrlenkungsverfahren nähern sowohl bei Verwendung des Mehrpfadansatzes, welcher zwei LSPs pro LSP-Fächer verwendet, als auch bei Verwendung des Pfadumlegungsansatzes die Netzperformanz der optimalen Lastverteilung sehr gut an. Insbesondere bei den höheren Netzauslastungen mit einer mittleren Quellenlast von $45 \frac{MBit}{s}$ bzw. $50 \frac{MBit}{s}$ entspricht die maximale Linkauslastung den Werten, die mit der optimalen Lastverteilung erzielt werden. Der minimale Durchsatz einer TCP Verbindung im Netz der optimalen Lastverteilung wird von dem Mehrpfadansatz zu 90% und von dem Pfadumlegungsansatz zu 85% erreicht.

Verglichen mit dem Shortest Path Routing erzielen die reaktiven Verkehrlenkungsverfahren eine erhebliche Steigerung der Performanz. Unter Verwendung der reaktiven Verkehrlenkungsverfahren lässt sich die Eingangslast um 50% steigern, ohne eine schlechtere Netzperformanz als das Shortest Path Routing zu erzielen.

5.1.2.2 Auswertung des Szenarios Lastgranularität Das zweite Szenario bewertet den Einfluss der Granularität, mit der sich die gesamte Netzlast auf die LSPs verteilt. Hierzu wird die gesamte Eingangslast auf 10% der LSPs im Netz verteilt. Um dies zu erreichen, trägt nur ein LSP bzw. LSP-Fächer die gesamte Last, welche zwischen zwei POPs ausgetauscht wird. Die mittleren Auslastungen der übrigen LSPs haben einen vernachlässigbaren geringen Wert. Bei 10 POPs in dem 44Knoten-Netz (siehe hierzu Abbildung 41) teilt sich Gesamtlast im Netz auf 90 LSPs bzw. LSP-Fächer auf. Der Mittelwert der Quellenlast dieser LSPs wird in den Untersuchungen zwischen $250 \frac{MBit}{s}$ und $450 \frac{MBit}{s}$ variiert. Die Varianz der Quellenlast entspricht $120 \frac{MBit}{s}$.

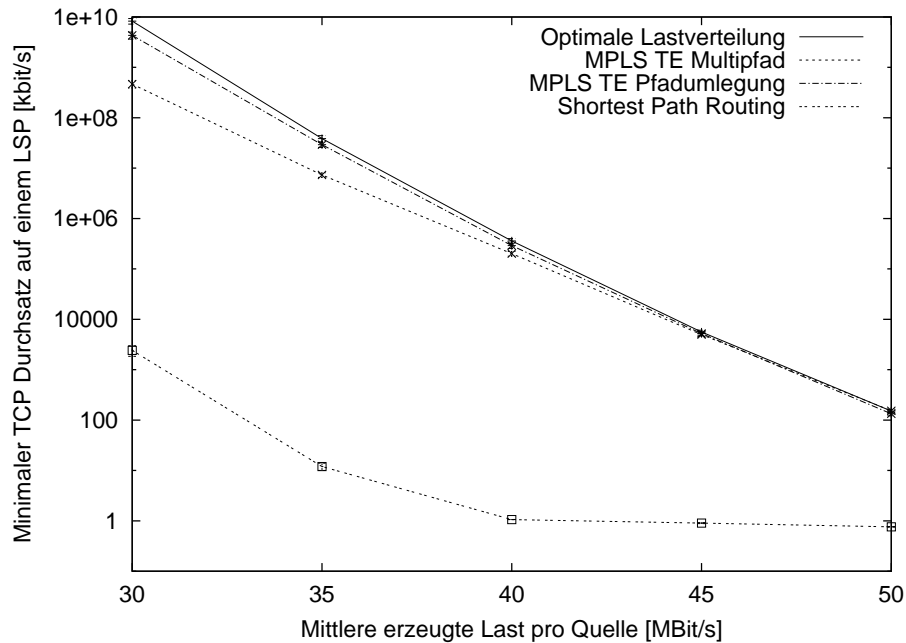


Abbildung 43: Minimaler TCP Durchsatz im Gleichlast-Szenario

Die Abbildungen 44 und 45 zeigen die maximale Linkauslastung sowie den minimalen Durchsatz einer TCP Verbindung für das veränderte Auslastungsszenario. Das reaktive Verkehrslenkungsverfahren nach dem Mehrpfadansatz, welcher in diesem Szenario drei LSPs pro LSP-Fächer verwendet, zeigt eine zum ersten Szenario vergleichbare Annäherung an die optimale Lastverteilung. Während bei den hohen Netzauslastungen die erzielte maximale Linkauslastung dem Wert der optimalen Verteilung entspricht, erreicht das Verfahren nach dem Mehrpfadansatz für den minimalen Durchsatz einer TCP Verbindung im Netz 90% des Wertes der optimalen Verteilung.

Das reaktive Verkehrslenkungsverfahren nach dem Pfadumlegungsansatz erzielt in diesem Szenario hingegen eine Netzperformanz, welche deutlich von der optimalen Lastverteilung abweicht. Der minimale Durchsatz einer TCP Verbindung im Netz entspricht ca. 10% des Wertes der optimalen Lastverteilung und die erzielte maximale Linkauslastung liegt um $100 \frac{MBit}{s}$ bis $200 \frac{MBit}{s}$ über der optimalen Lastverteilung. In dieser Untersuchung zeigt sich der Einfluss der Granularität der Lastverteilung auf die Performanz des reaktiven Verkehrslenkungsverfahrens nach dem Pfadumlegungsansatz.

Verglichen mit dem Shortest Path Routing verbessert das reaktive Verkehrslenkungsverfahren nach beiden Ansätzen die Netzperformanz erheblich. Während in dem untersuchten Auslastungsbereich das reaktive Verkehrslenkungsverfahren nach dem Pfadumlegungsansatz eine um 80% höhere Eingangslast bei gleicher Netzperformanz erlaubt, ist der Wert für den Mehrpfadansatz sogar noch höher.

5.1.2.3 Auswertung des Szenarios *Initiale Lastverteilung* Dieser Abschnitt untersucht den Einfluss der initialen Lastverteilung auf die Performanz der reaktiven Verkehrslenkungsverfahrens. Diese Untersuchung basiert auf dem Gleichlast-Szenario aus Abschnitt 5.1.2.1, in welchem die Netzlast gleichmäßig auf die LSPs aufgeteilt ist.

Um die initiale Lastverteilung für den Pfadumlegungsansatz zu variieren, wird pro Link des 44Knoten-Netzes eine zufällige imaginäre Link-Auslastung zwischen dem Wert 0 und

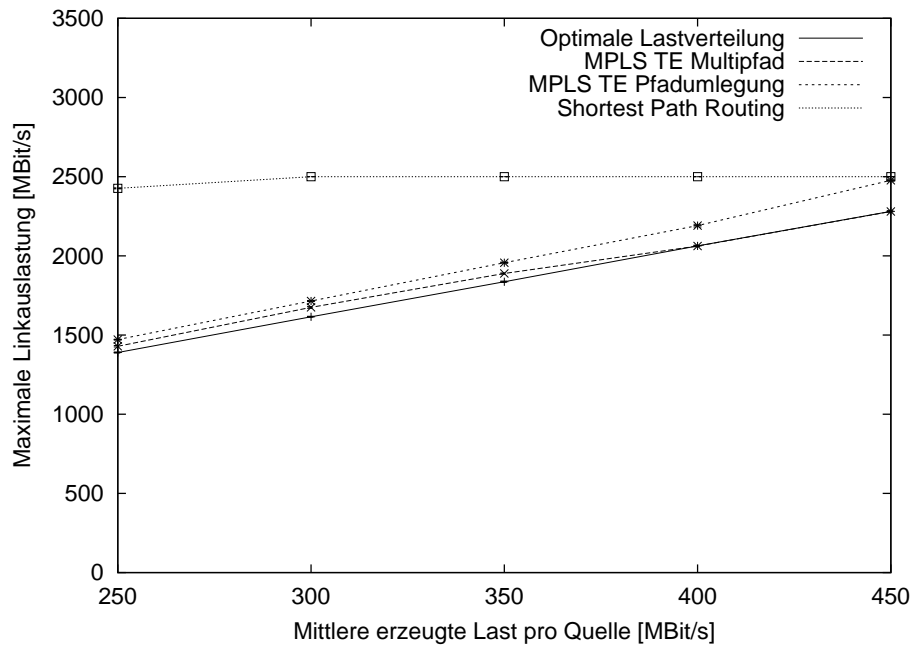


Abbildung 44: Maximale Linkauslastung bei grober Lastgranularität

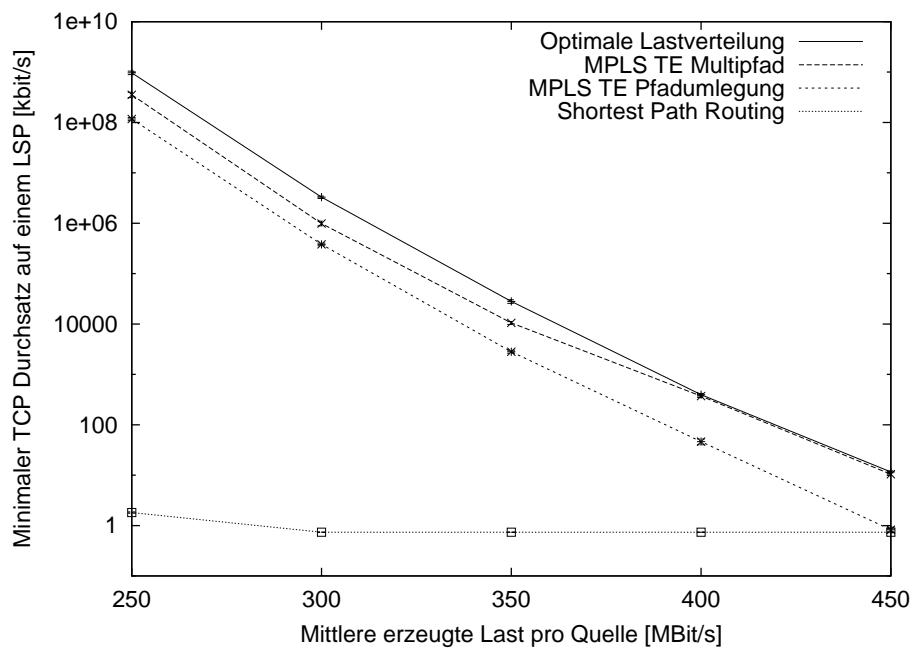


Abbildung 45: Minimaler TCP Durchsatzes bei grober Lastgranularität

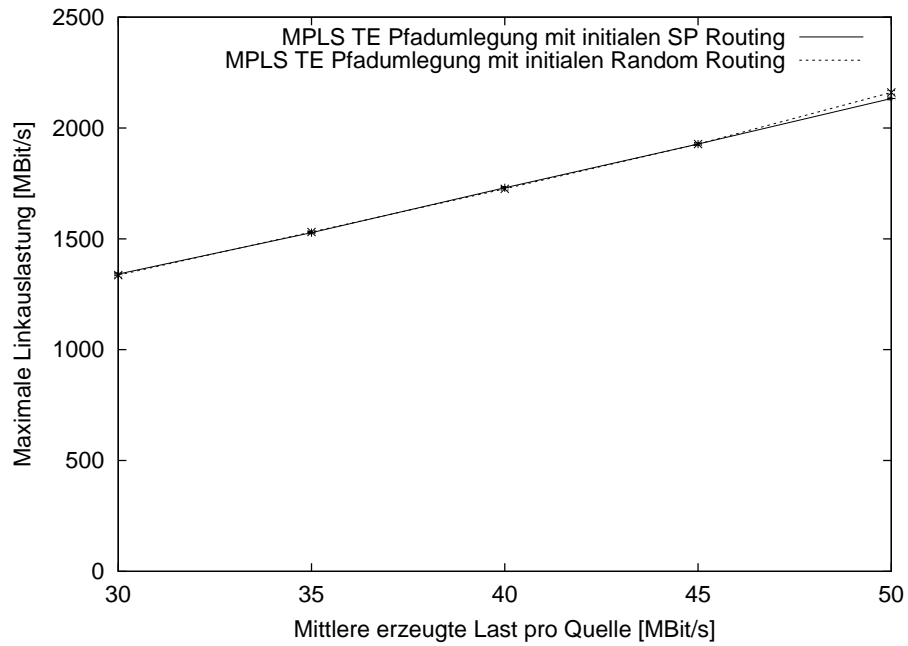


Abbildung 46: Maximale Linkauslastung des Pfadumlegungsansatzes bei variierender initialer Lastverteilung

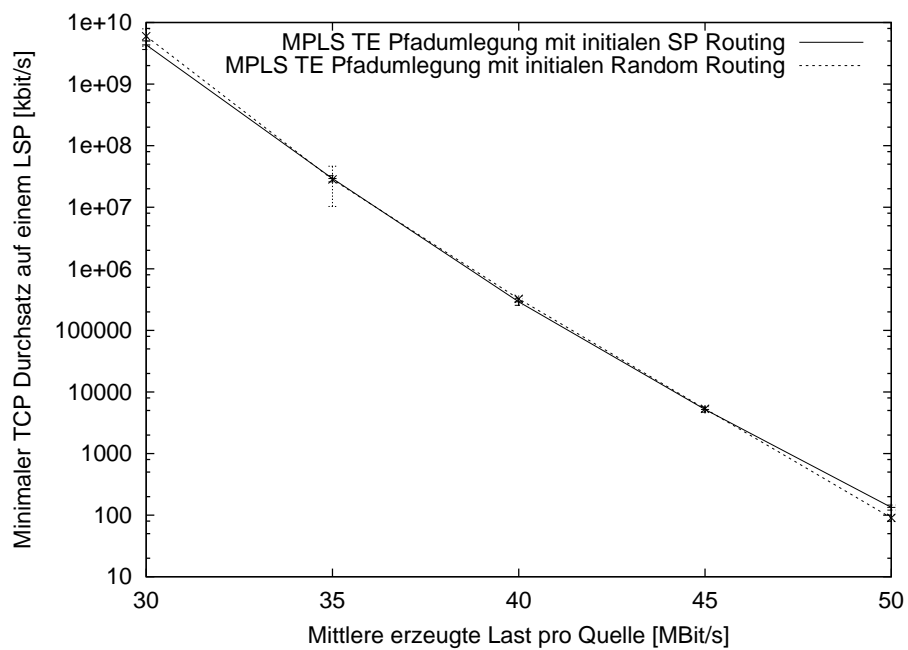


Abbildung 47: Minimaler TCP Durchsatz des Pfadumlegungsansatzes bei variierender initialer Lastverteilung

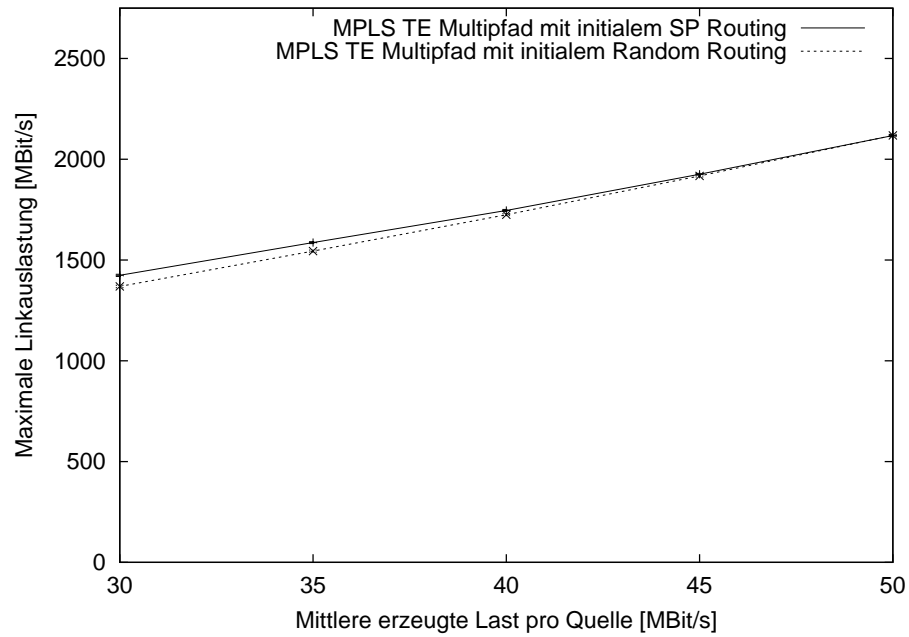


Abbildung 48: Maximale Linkauslastung des Multipfadansatzes bei variierender initialer Lastverteilung

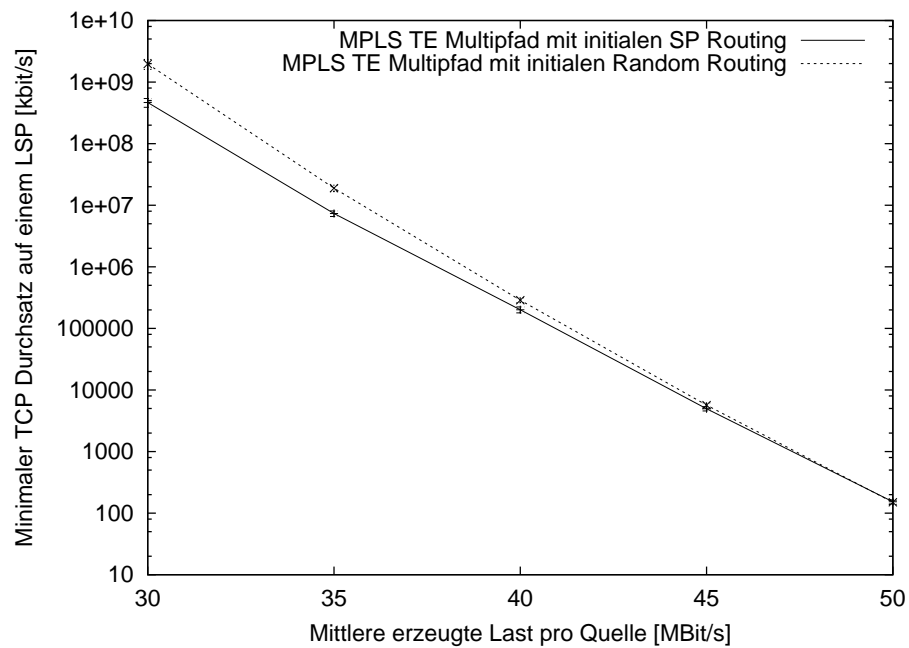


Abbildung 49: Minimaler TCP Durchsatz des Multipfadansatzes bei variierender initialer Lastverteilung

der maximalen Übertragungskapazität des Links angenommen. Diese Werte werden mit einer gleichverteilten Zufallszahl generiert. Basierend auf der zufällig generierten imaginären Netzauslastung werden die LSPs mit dem in Abschnitt 3.4.2 beschriebenen Pfadumlegungsalgorithmus durch das Netz geroutet.

Für den Mehrpfadansatz wird die initiale Lastverteilung über eine zufällig generierte Aufteilung des Verkehrs auf die LSPs eines LSP-Fächers erzeugt. Die zufällige Aufteilung wird ebenfalls mit einer gleichverteilten Zufallszahl generiert.

Die Ergebnisse der maximalen Linkauslastung und des minimalen Durchsatzes einer TCP Verbindung im Netz für den Pfadumlegungsansatz zeigen die Abbildungen 46 und 47. Es ist zu erkennen, dass die initiale Lastverteilung für den Pfadumlegungsansatz in diesem Szenario keinen nennenswerten Einfluss auf die Performanz nimmt. Wie aus den Abbildungen 48 und 49 zu erkennen ist, zeigen sich für den Mehrpfadansatz die gleichen Ergebnisse. Auch hier hat die initiale Lastverteilung keinen nennenswerten Einfluss auf die Performanz.

5.1.3 Untersuchung eines 14Knoten-Netzes

Dieser Abschnitt evaluiert die Performanz der reaktiven Verkehrlenkungsverfahren in einem 14Knoten-Netz. Die Topologie dieses Netzes, welches in Abbildung 50 dargestellt ist, entspricht dem Core-Netz des 44Knoten-Netzes.

Das 14Knoten-Netz besteht aus 14 LSRs und 48 Links, wobei die 10 äußeren LSRs des Netzes die Eingangs- bzw. Ausgangs-LSRs sind. Für eine einfache Vollvermaschung dieser LSRs im Pfadumlegungsansatz werden 90 LSPs benötigt. Im Mehrpfadansatz werden bei zwei LSPs pro LSP-Fächer 180 LSPs geschaltet. Die Länge der Links entspricht 200km.

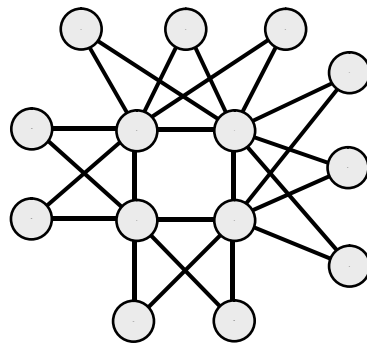


Abbildung 50: Topologie des 14Knoten-Netzes

Das 14Knoten-Netz wird für zwei verschiedene Szenarien untersucht. Im ersten Szenario verteilt sich die Last gleichmäßig auf die LSPs und es wird vom einem initialen Shortest Path Routing ausgegangen. Die Ergebnisse der Untersuchung präsentiert Abschnitt 5.1.3.1. In der in Abschnitt 5.1.3.2 vorgestellten Untersuchung des zweiten Szenarios wird der Einfluss der initialen Lastverteilung evaluiert.

5.1.3.1 Auswertung des Gleichlast-Szenarios Wie bei der Untersuchung des Gleichlast-Szenarios des 44Knoten-Netzes teilt sich Netzlast gleichmäßig auf die LSPs bzw. LSP-Fächer auf. Der Mittelwert der über eine Normal-Verteilung erzeugten Quellenlast

variiert je nach Untersuchung zwischen $300 \frac{MBit}{s}$ und $450 \frac{MBit}{s}$. Die Varianz entspricht in allen Szenarien dem Wert $1200 \frac{MBit}{s}$.

Die Abbildungen 51 und 52 zeigen die maximale Linkauslastung sowie den minimalen Durchsatz einer TCP Verbindung im Netz. Die Ergebnisse der Untersuchungen zeigen ein vergleichbares Verhalten zu der Untersuchung des 44Knoten-Netzes mit veränderter Lastgranularität aus Abschnitt 5.1.2.2. Die Netzperformanz, welche das reaktive Verkehrslenkungsverfahren nach dem Pfadumlegungsansatz erzielt, weicht deutlich von der optimalen Lastverteilung ab. Der minimale Durchsatz einer TCP Verbindung entspricht lediglich 3% des Wertes der optimalen Lösung. Die maximale Linkauslastung ist um ca. $200 \frac{MBit}{s}$ größer als bei der optimalen Linkauslastung. Auch hier zeigt sich der Einfluss der Granularität auf die Performanz des Pfadumlegungsansatzes.

Der Mehrpfadansatz hingegen erzielt sowohl für die maximale Linkauslastung als auch den minimalen TCP Durchsatz Werte, die mit der optimalen Lastverteilung vergleichbar sind.

Gegenüber dem Shortest Path Routing lässt sich der Eingangsverkehr mit dem Mehrpfadansatz um ca. 35% erhöhen, ohne die Netzperformanz zu verschlechtern. Mit dem Pfadumlegungsansatz lässt sich hingegen lediglich eine Erhöhung des Eingangsverkehrs um ca. 25% erzielen.

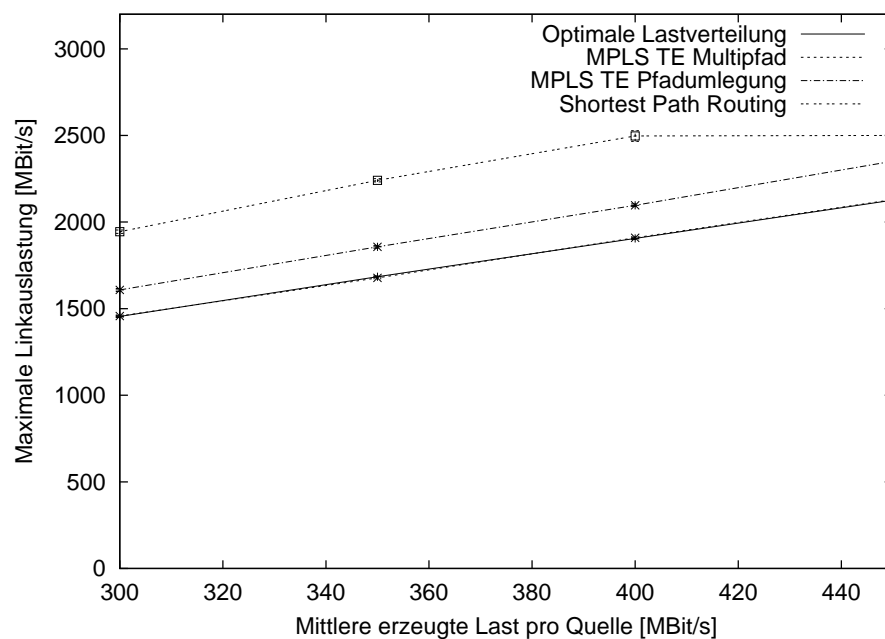


Abbildung 51: Maximale Linkauslastung im Gleichlast-Szenario

5.1.3.2 Auswertung des Szenarios *Initiale Lastverteilung* Die Durchführung der Untersuchung erfolgt analog zu der in Abschnitt 5.1.2.3 beschriebenen Untersuchung am 44Knoten-Netz. Das verwendete Lastszenario entspricht dem Gleichlast-Szenario aus dem zuvor beschriebenen Abschnitt 5.1.3.1.

Die Ergebnisse der Untersuchung für den Pfadumlegungsansatz sind in den Abbildungen 53 und 54 dargestellt. Hier ist bei den hohen Netzauslastungen eine deutliche Verschlechterung der Performanz bedingt durch die veränderte initiale Lastverteilung zu erkennen.

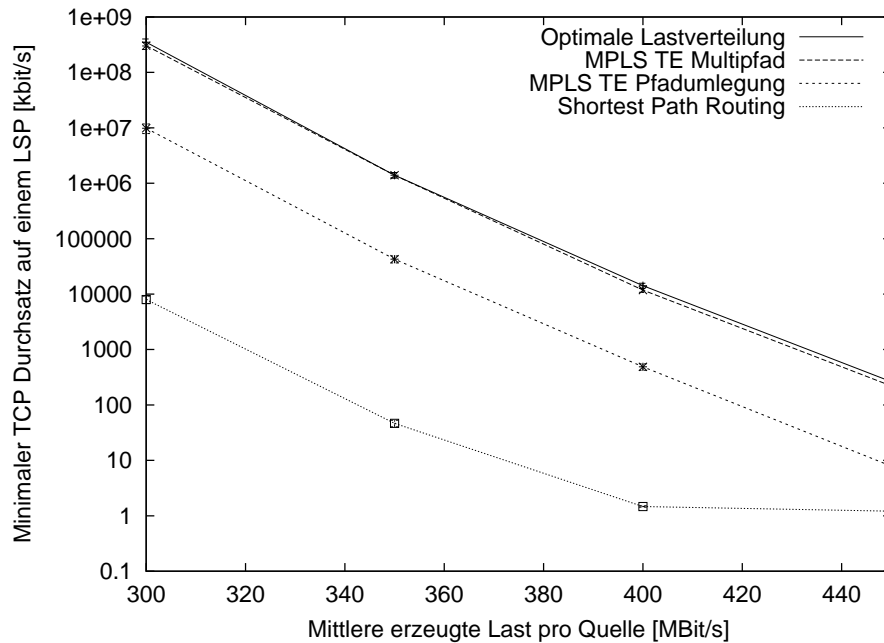


Abbildung 52: Minimaler TCP Durchsatz im Gleichlast-Szenario

Der minimale TCP Durchsatz reduziert sich auf ca. 10% des Wertes, welcher bei initialem Shortest Path Routing erzielt wird. Die maximale Linkauslastung ist um bis zu $200 \frac{MBit}{s}$ höher. Dieser Effekt lässt sich durch die in Abschnitt 3.5.2 erläuterten Verklemmungssituationen erklären, welche bei dem Pfadumlegungsansatz auftreten können. Diese treten insbesondere in Netzen mit geringer Anzahl an LSPs auf, da hier nur wenige Möglichkeiten zur Umverteilung der Last bestehen.

Die Performanz des Mehrpfadansatzes zeigt hingegen keine Abhängigkeit von der initialen Lastverteilung. Dies geht aus den Abbildungen 55 und 56 hervor, welche die Ergebnisse der Untersuchungen für den Mehrpfadansatz darstellen.

5.1.4 Untersuchung eines 45Knoten-Netzes

Zur Evaluierung der Performanz-Analyse der reaktiven Verkehrslenkungsverfahren präsentiert dieser Abschnitt die Ergebnisse der Untersuchungen an dem in Abbildung 57 dargestellten 45Knoten-Netz.

Das 45Knoten-Netz ist ein typisches amerikanisches Netz, welches aus 45 LSRs und 81 Links besteht. In diesem ist jeder LSR ein Eingangs- bzw. Ausgangs-LSRs, wodurch im Pfadumlegungsansatz bei einer einfachen Vollvermaschung 1980 LSPs und bei dem Mehrpfadansatz mit zwei LSPs pro LSP-Fächer 3960 LSPs geschaltet werden. Das entspricht ca. der doppelten Anzahl an LSPs gegenüber dem 44Knoten-Netz. Die Links haben eine Übertragungskapazität von $2.5 \frac{GBit}{s}$ und die Länge der Links entsprechen den Abständen der Städte, in denen die LSRs platziert sind.

In dem in diesem Abschnitt untersuchten Gleichlast-Szenario teilt sich die Last gleichmäßig auf die LSPs im Netz auf. Die mittlere Senderate der Quellen variiert je nach Untersuchung zwischen $15 \frac{MBit}{s}$ und $25 \frac{MBit}{s}$. Die Varianz der Quellenlast ist auf den Wert von $60 \frac{MBit}{s}$ gesetzt.

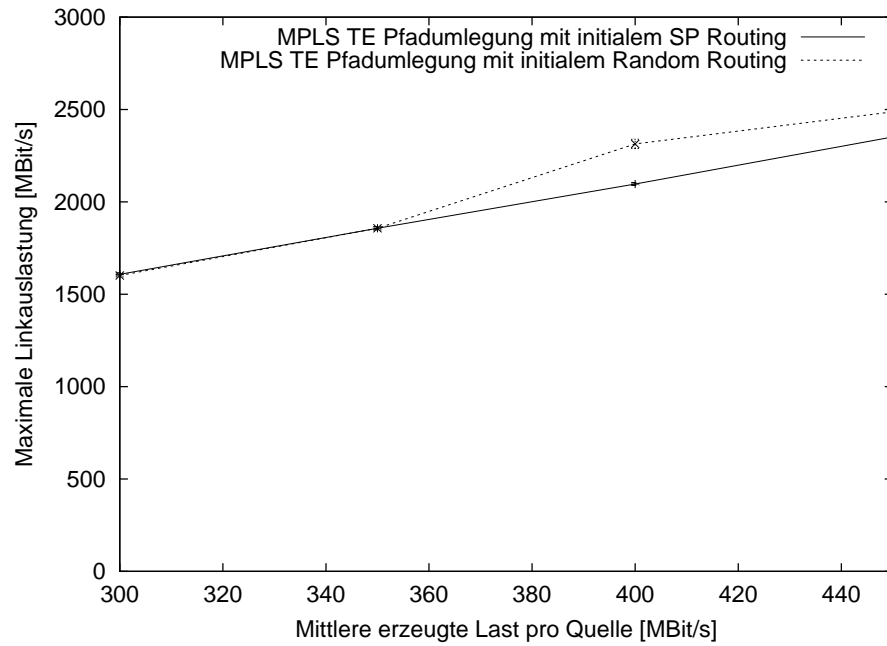


Abbildung 53: Maximale Linkauslastung des Pfadumlegungsansatzes bei variierender initialer Lastverteilung

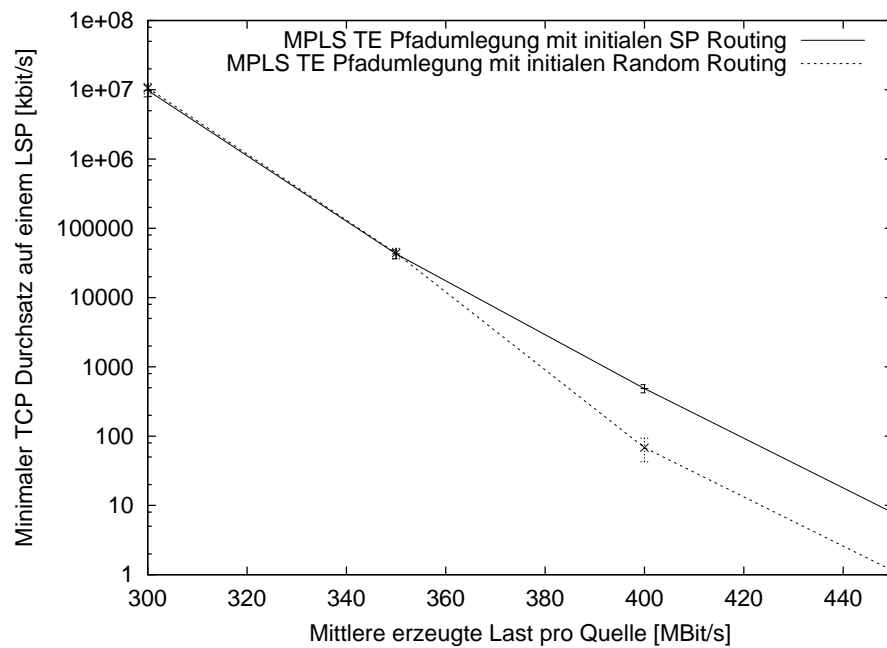


Abbildung 54: Minimaler TCP Durchsatz des Pfadumlegungsansatzes bei variierender initialer Lastverteilung

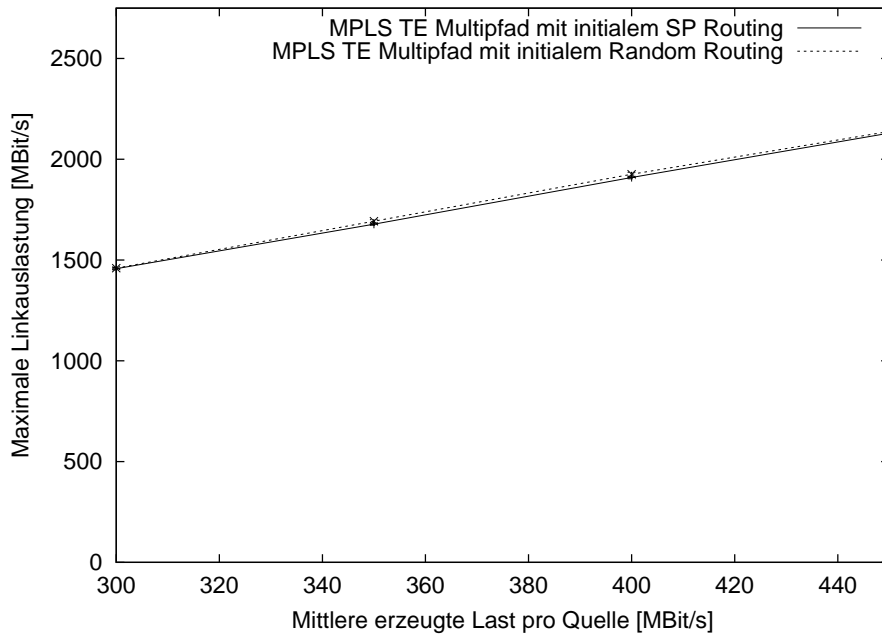


Abbildung 55: Maximale Linkauslastung des Multipfadansatzes bei variierender initialer Lastverteilung

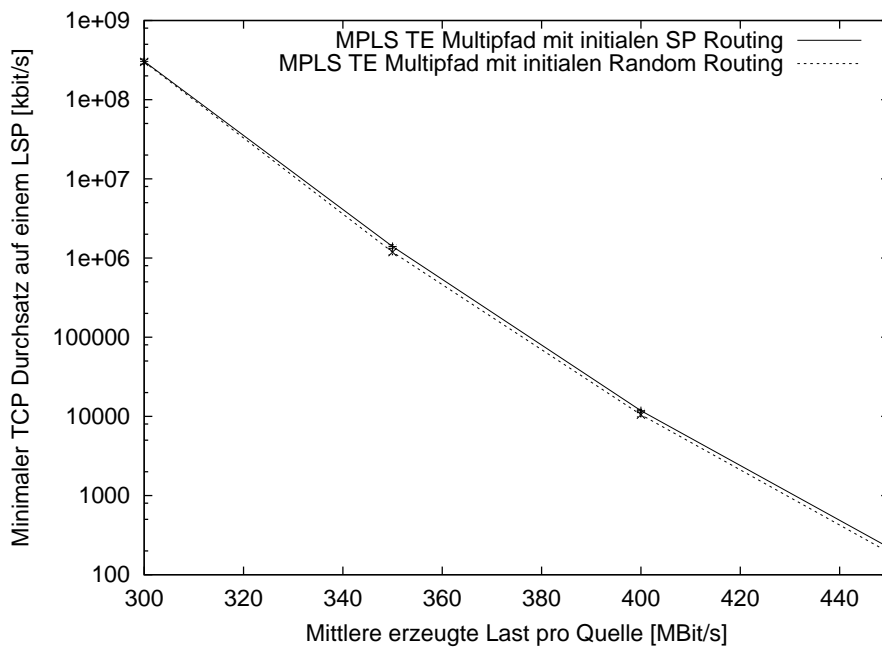


Abbildung 56: Minimaler TCP Durchsatz des Multipfadansatzes bei variierender initialer Lastverteilung

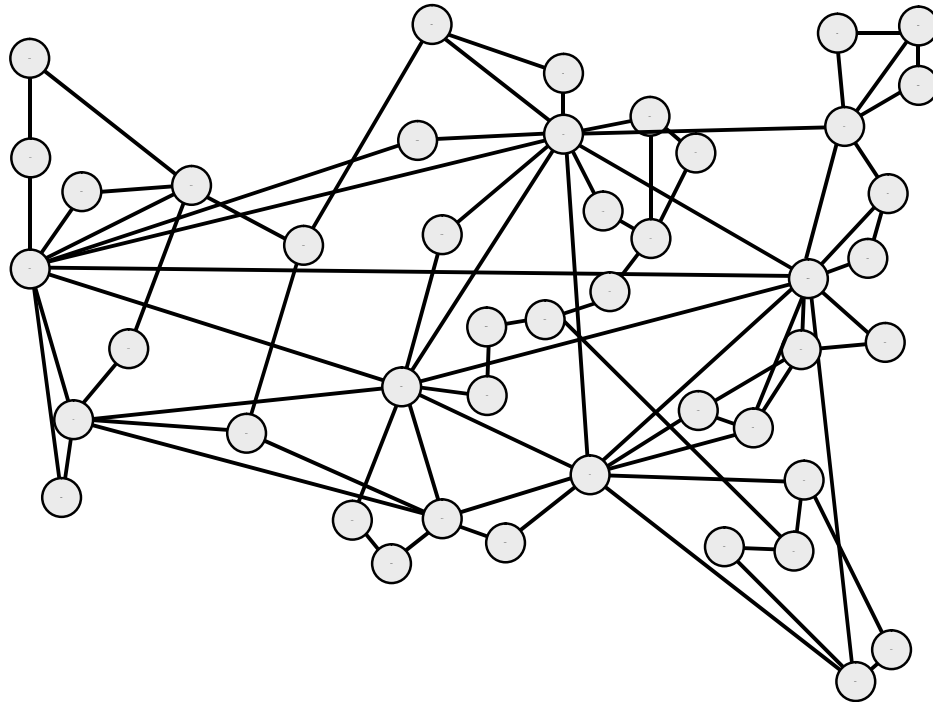


Abbildung 57: Topologie des 45Knoten-Netzes

Aus den in den Abbildungen 58 und 59 dargestellten Ergebnissen lässt sich erkennen, dass sowohl der Pfadumlegungsansatz als auch der Mehrpfadansatz die Performanzwerte der optimalen Lastverteilung annähern. Der minimale Durchsatz einer TCP Verbindung im Mehrpfadansatz entspricht mindestens 50% und der Wert des Pfadumlegungsansatzes mindestens 70% des Wertes der optimalen Lastverteilung. Die Performanz des Mehrpfadansatzes ist in diesem Szenario schlechter als die des Pfadumlegungsansatzes, da die Anzahl von 2 LSPs pro LSP-Fächer die Performanz in diesem Szenario begrenzt.

Gegenüber dem Shortest Path Routing lässt sich bei diesem Szenario unter Verwendung des Verfahrens die Eingangslast um über 70% erhöhen ohne die Netzperformanz zu verschlechtern.

5.1.5 Untersuchung eines 16Knoten-Netzes

Mit der Untersuchung der reaktiven Verkehrslenkungsverfahren am 16Knoten-Netz stellt dieser Abschnitt eine weitere Evaluierung der Performanz der reaktiven Verkehrslenkungsverfahren vor.

Das in Abbildung 60 dargestellte norditalienische 16Knoten-Netz besteht aus 16 LSRs und 36 Links. Alle LSRs dienen als Eingangs- bzw. Ausgangs-LSR, was bei einer Vollvermaschung aller LSRs im Pfadumlegungsansatz 240 LSPs und im Mehrpfadansatz mit zwei LSPs pro LSP-Fächer 480 LSPs ergibt. Die Länge der Links entspricht den Entfernungen der Städte, in denen die LSRs platziert sind.

In dem untersuchten Gleichlast-Szenario verteilt sich die Eingangslast wie in den zuvor beschriebenen Szenarien gleichmäßig über die LSPs bzw. LSP-Fächer. Der Mittelwert der Quellenlast wird in den Untersuchungen zwischen $100 \frac{MBit}{s}$ und $200 \frac{MBit}{s}$ variiert. Die Varianz der erzeugten Quellenlast ist in allen Untersuchungen $300 \frac{MBit}{s}$.

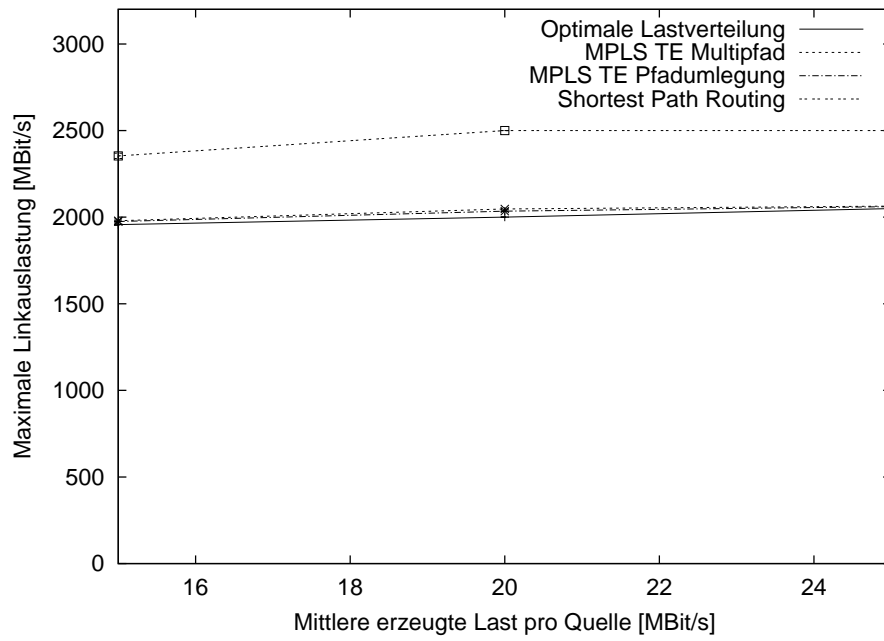


Abbildung 58: Maximale Linkauslastung im Gleichlast-Szenario

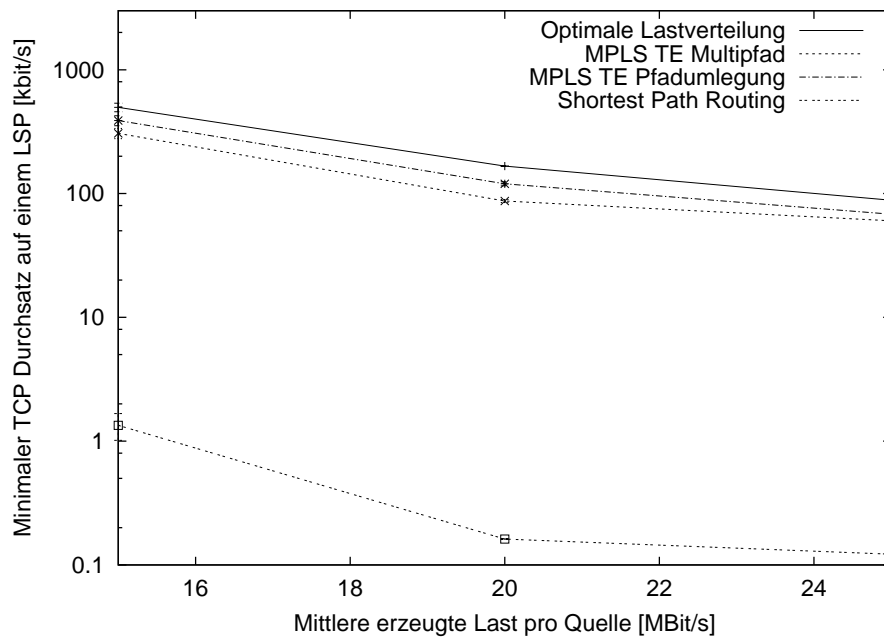


Abbildung 59: Minimaler TCP Durchsatz im Gleichlast-Szenario

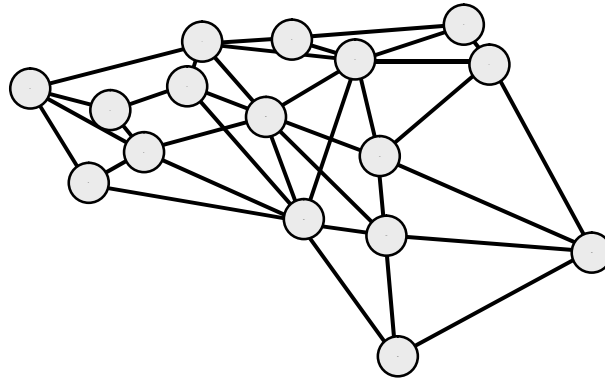


Abbildung 60: Topologie des 16Knoten-Netztes

Die Abbildungen 61 und 62 zeigen die maximale Linkauslastung sowie den minimalen Durchsatz einer TCP Verbindung in dem 16Knoten-Netz. Sowohl der Pfadumlegungsansatz als auch der Mehrpfadansatz erzielen eine mit der optimalen Lastverteilung vergleichbare Netzperformanz. Der minimale Durchsatz einer TCP Verbindung entspricht in der Untersuchung mindestens 55% des Wertes, welcher mit der optimalen Lösung erzielt wird.

Der Vergleich mit dem Shortest Routing zeigt, dass unter Verwendung der reaktiven Verkehrslenkungsverfahren die Eingangslast um 100% gesteigert werden kann, ohne eine geringere Netzperformanz zu erzielen.

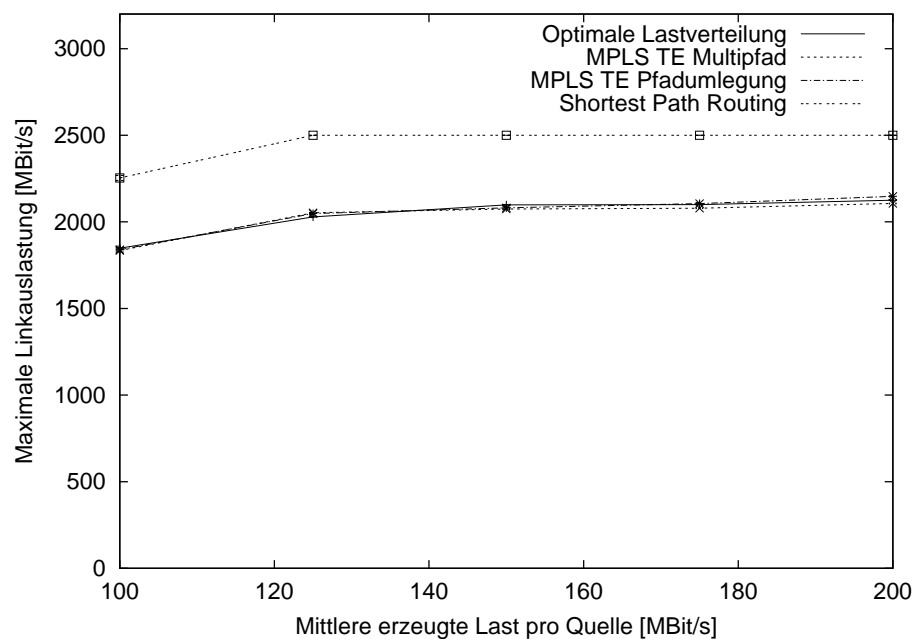


Abbildung 61: Maximale Linkauslastung im Gleichlast-Szenario

5.1.6 Zusammenfassende Bewertung der Netzperformanz

Eine wichtige Erkenntnis der vorgestellten Untersuchungen ist, dass der Mehrpfadansatz unabhängig von dem Netz- und Auslastungsszenario die optimale Netzperformanz sehr gut annähert. Dieses gute Ergebnis erreicht der Mehrpfadansatz bereits mit der geringen Anzahl von zwei bzw. drei LSPs pro LSP-Fächer.

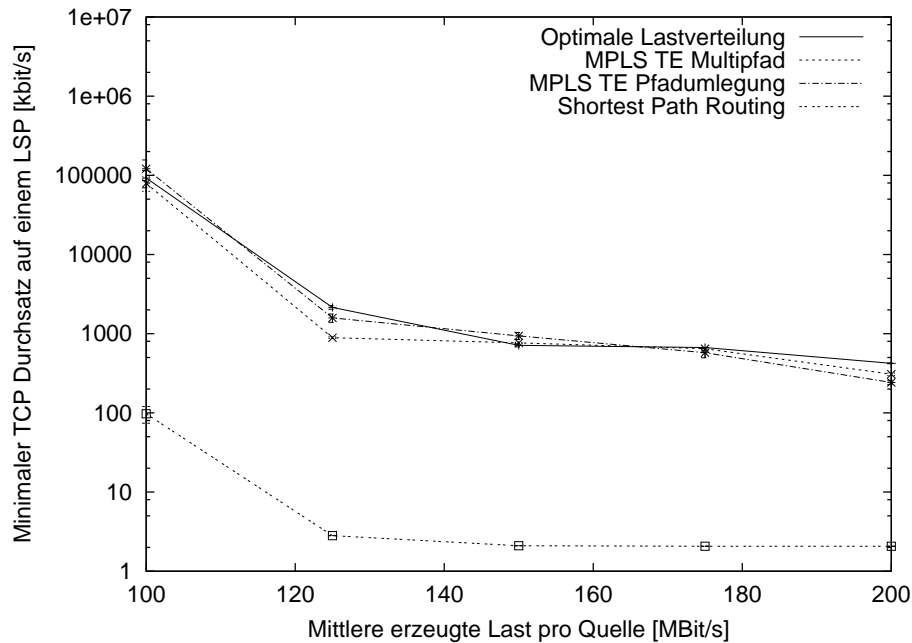


Abbildung 62: Minimaler TCP Durchsatz im Gleichlast-Szenario

Bei der Analyse des Pfadumlegungsansatzes zeigt sich eine Abhängigkeit der Performanz von dem Netz- und Auslastungsszenario. In kleinen Netzen bzw. in Auslastungsszenarien, in welchen sich die Last auf wenige LSPs verteilt, erreicht der Pfadumlegungsansatz nicht die Netzperformanz der optimalen Lastverteilung. Diese Einschränkung lässt sich auf die durch die Anzahl der LSPs vorgegebene Granularität zurückführen, mit der die Last in einem Netz verteilt werden kann. In großen Netzen hingegen, in welchen sich die gesamte Netzlast auf viele LSPs verteilt, nähert auch der Pfadumlegungsansatz die Netzperformanz der optimalen Lastverteilung sehr gut an.

Bei dem Vergleich der reaktiven Verkehrslenkungsverfahren mit dem Shortest Path Routing zeigt sich in allen Fällen eine deutliche Verbesserung der Netzperformanz. Die Anwendung der reaktiven Verkehrslenkungsverfahren lässt in den untersuchten Szenarien eine Erhöhung der gesamten Eingangslast zwischen 25% und 100% zu, ohne eine geringere Netzperformanz als das Shortest Path Routing zu erzielen.

5.2 Analyse der Konvergenzdauer

Der zweite Schwerpunkt der in diesem Kapitel durchgeführten quantitativen Analyse des reaktiven Verkehrslenkungsverfahrens ist die Untersuchung der Konvergenzdauer. Hierbei werden zwei Aspekte berücksichtigt. Zum einen gilt es zu zeigen, dass das reaktive Verkehrslenkungsverfahren schneller konvergiert als das Konkurrenz-Verfahren MATE. Das reaktive Verfahren ist unter anderem mit der Vorgabe entwickelt worden, die geringe Konvergenzgeschwindigkeit des MATE-Verfahrens zu verbessern. Zum anderen hat die qualitative Analyse der reaktiven Verkehrslenkungsverfahren gezeigt, dass die Konvergenzdauer eine Abhängigkeit vom Netz- und Auslastungsszenario aufweist. Diese gilt es im Rahmen dieser Untersuchung zu quantifizieren. Dieser Abschnitt präsentiert die Ergebnisse der quantitativen Analyse der Konvergenzdauer anhand von zwei Netzszenarien.

Zur Koordinierung der verteilten LSRs verwenden die reaktiven Verkehrslenkungsverfah-

ren mehrere Timer. Diese in Abschnitt 3.3.4 definierten Timer hängen vom Netzscenario ab und beeinflussen die Konvergenzdauer maßgeblich. Aus diesem Grund präsentiert der Abschnitt 5.2.1 die Festlegung der Timer-Werte.

Um das MATE-Verfahren bewerten und mit den entwickelten reaktiven Verfahren vergleichen zu können, wird ein ratenbasiertes Simulationsmodell des MATE-Verfahrens benötigt. Die im Rahmen dieser Untersuchung verwendete Modellierung des MATE-Verfahrens beschreibt Abschnitt 5.2.2.

Um den Einfluss der Netzgröße auf die Konvergenzdauer zu evaluieren, wurden zwei Netze unterschiedlicher Größe ausgewählt. Die Untersuchung der Konvergenzdauer erfolgt anhand des 44Knoten-Netzes und des 14Knoten-Netzes, welche in den Abschnitten 5.1.2 und 5.1.3 vorgestellt wurden. Die Ergebnisse der Untersuchungen der Konvergenzdauer werden in den Abschnitten 5.2.3 und 5.2.4 vorgestellt.

5.2.1 Vorgehen zur Bestimmung der Timer-Werte

Die in Kapitel 3 vorgestellten reaktiven Verkehrslenkungsverfahren definieren Timer, welche zur Koordinierung des Lastverteilungsprozesses dienen. Über die Timer wird der Ablauf der einzelnen Phasen eines Lastverteilungsprozesses gesteuert. Zu den Phasen eines Lastverteilungsprozesses zählen der Austausch von LSP-Auslastungsinformationen zwischen den LSRs, der Verbindungsaufbau eines LSPs sowie die Berechnung der Lastumverteilung in den LSRs. Die Festlegung der Timer-Werte ist abhängig von dem Netzscenario, da die maximale Dauer der einzelnen Phasen je nach Größe des Netzes variiert. Dieser Abschnitt erläutert das verwendete Vorgehen zur Festlegung der Timer-Werte für ein Netzscenario.

Bestimmung der maximalen Dauer $T_{last-info}$ zum Austausch von LSP-Auslastungsinformationen:

Gleichung 75 zeigt die Formel zur Bestimmung des Timers $T_{last-info}$, welcher zum Austausch von LSP-Auslastungsinformationen zwischen zwei LSRs im Netz benötigt wird. Der Timer ergibt sich aus der maximalen Signallaufzeit zwischen zwei Netzknoten T_{sl} , der Verarbeitungsdauer der Nachrichten $T_{vd-last-info}$ in den Netzknoten auf dem Pfad sowie der Dauer $T_{sp-last-info}$ zum Speichern der Auslastungsinformationen. Die sich aus diesen Werten ergebende Gesamtdauer wird mit einem Sicherheits-Faktor F multipliziert, um mögliche Abweichungen von der maximalen Gesamtdauer auszugleichen. Der Faktor muss größer gleich eins gewählt werden.

$$T_{last-info} = F \cdot (T_{sl\ max} + (HC_{max} + 1) \cdot T_{vd-last-info} + T_{sp-last-info}) \text{ mit } F \geq 1 \quad (75)$$

Die maximale Signallaufzeit T_{sl} leitet sich aus der maximalen Länge einer Übertragungsstrecke zwischen zwei LSRs und der Signalgeschwindigkeit auf dem Medium ab. Die Gesamt-Verarbeitungsdauer einer Meldung in den LSRs auf dem Pfad berechnet sich durch die Multiplikation des maximalen Hop Counts HC_{max} und der Verarbeitungsdauer $T_{vd-last-info}$ einer Lastinformationsmeldung. Da die Speicherung der Auslastungsinformationen parallel zum Aussenden der Meldung erfolgt, wird diese Dauer nur für den letzten LSR berechnet.

Bestimmung der maximalen Dauer $T_{lsp-sig}$ zum Aufbau eines LSPs:

Gleichung 76 zeigt die Berechnungsvorschrift zur Bestimmung des Timers $T_{lsp-sig}$, welcher bei dem Aufbau eines LSPs verwendet wird. Die maximale Dauer $T_{lsp-sig}$ berechnet sich aus der maximalen Signallaufzeit $T_{sl\ max}$ zwischen zwei Netzknoten und der gesamten Verarbeitungsdauer der Signalisierungs-Nachrichten $T_{vd-lsp-sig}$. Zum Ausgleich möglicher Abweichungen von der maximalen Gesamtdauer wird die gesamte maximale Signalisierungsdauer ebenfalls mit einem Sicherheits-Faktor F multipliziert.

$$T_{lsp-sig} = F \cdot (2 \cdot (T_{sl\ max} + (HC + 1) \cdot T_{vd-lsp-sig})) \quad (76)$$

Die maximale Signallaufzeit T_{sl} sowie die gesamte Verarbeitungsdauer der Nachrichten in den LSRs berechnet sich wie in dem zuvor beschriebenen Fall. Da der Verbindungsaufbau der Label Distribution Protokolle im Downstream On Demand Modus (siehe hierzu Abschnitt 2.1.2.3) erfolgt und sich somit eine Verbindungs-Anfrage von der Quelle an die Senke und eine Verbindungs-Bestätigung von der Senke an die Quelle gesendet wird, wird die maximale Signallaufzeit T_{sl} und sowie die gesamte Dauer zur Verarbeitung der Signalisierungsnachrichten mit dem Faktor zwei multipliziert.

Bestimmung der maximalen Dauer $T_{lastverteilung-cal}$ zur Berechnung der Lastumverteilung eines LSPs:

In den Verkehrlenkungsverfahren nach dem Pfadumlegungsansatz entspricht eine Lastumverteilungsaktion der Durchführung eines Shortest Path Routings. Somit ergibt sich die Dauer einer Lastumverteilungsaktion nach dem Pfadumlegungsansatz aus der Dauer einer Shortest Path Berechnung.

Die Dauer einer Lastumverteilungsaktion bei Verwendung des Mehrpfadansatzes wird vereinfachend der gleiche Wert angesetzt.

Bestimmung der maximalen Dauer $T_{lastverteilung-par-cal}$ zur Berechnung der Lastumverteilung im parallelen Betriebsmodus:

Die Bestimmung der maximalen Dauer $T_{lastverteilung-par-cal}$ zur Berechnung der Lastumverteilung im parallelen Betriebsmodus wird nach Gleichung 9 aus der Berechnungsdauer einer Lastverteilungsaktion $T_{lastverteilung-cal}$ und der maximalen Anzahl der LSPs im Netz abgeleitet.

Bestimmung der maximalen Dauer $T_{netzlast-cal}$ zur Berechnung der Netzauslastung:

Die Bestimmung der maximalen Dauer $T_{netzlast-cal}$ zur Berechnung der Netzauslastung ergibt sich aus der Anzahl der Berechnungsschritte zur Bestimmung der Lastverteilung eines LSP-Fächers sowie den Leistungsparametern der verwendeten Hardware.

5.2.2 Modellierung des MPLS Adaptive Traffic Engineering MATE

Um das MATE-Verfahren mit Hilfe der modifizierten ratenbasierten Simulations-Methode untersuchen zu können, ist eine entsprechende Modellierung des Verfahrens notwendig. Die relevanten Aspekte der Modellierung beschreibt dieser Abschnitt.

MATE verwendet zur Bestimmung der Netzauslastung die Messung der Paketverzögerung und des Paketverlusts von gesendeten Mess-Paketen. Da die paketbasierte Messung nicht

auf das ratenbasierte Simulationskonzept übertragbar ist, basiert in der folgenden Untersuchung die Lastverteilung durch das MATE-Verfahren auf der Kenntnis der tatsächlichen aktuellen Netzauslastung. Das entspricht dem optimalen Fall, dass die paketbasierte Messung die exakte LSP-Auslastung bestimmt. Die Länge einer Monitoring-Periode wird wie in den in [EJS01] beschriebenen Untersuchungen auf 30 Sekunden gesetzt.

Die Schrittgröße des Gradienten-Projektions-Algorithmus, auf dem die Lastverteilung des MATE-Verfahren beruht, ist abhängig von der gewählten Kostenfunktion und von der Netztopologie (siehe Gleichung 7). Die maximale Steigung der Linkkostenfunktion wird mit der Lipschitz-Konstanten $K_{lipschitz}$ abgeschätzt. Bei der in Abschnitt 3.4.1 beschriebenen Kostenfunktion hat die maximale Steigung den Wert 2066.57.

In den folgenden Untersuchungen verwendet das auf dem Mehrpfadansatz basierende Verfahren MATE zwei LSPs pro LSP-Fächer. Das initiale Routing wird mit dem in Abschnitt 2.5.2.3 beschriebenen K-Shortest-Path Algorithmus durchgeführt.

5.2.3 Untersuchung eines 44Knoten-Netzes

In Abschnitt 5.2.3.1 werden zunächst die Werte für die Timer festgelegt. Im anschließenden Abschnitt 5.2.3.2 folgt eine Analyse der Konvergenzdauer.

5.2.3.1 Festlegung der Timer-Werte Die Timer, welche die reaktiven Verkehrslenkungsverfahren zur Koordinierung der verteilten TE-Einheiten definieren, hängen von der Größe des jeweiligen Netzes ab. Dieser Abschnitt legt die Timer-Werte für das 44Knoten-Netz fest.

Die Berechnung der Timer-Werte beruht auf Topologie- sowie Router-spezifischen Parametern. Die Topologie-spezifischen Parameter wie z.B. der maximale Hop Count bzw. der maximale Abstand zwischen zwei Netzknoten lassen sich direkt aus dem 44Knoten-Netz ableiten. Für die Router-spezifischen Parameter wie die Verarbeitungsdauer einer Nachricht bzw. die Berechnungsdauer einer Lastverteilungsaktion müssen Annahmen über die Performanz der Router getroffen werden.

In dieser Untersuchung werden die Router-spezifischen Parameter aus einer Messung des Zeitverhaltens von zwei verschiedenen OSPF-Routern abgeleitet, welche in [SG01] beschrieben sind. Der in [SG01] bestimmte Wert der Verarbeitungsdauer $T_{vd-last-info}$ einer OSPF-Nachricht liegt in der Größenordnung von 40 Millisekunden, wobei dieser Wert unabhängig von der Größe des Netzes ist. Für die Verarbeitungsdauer $T_{vd-lsp-sig}$ einer LSP-Signalisierungsnachricht wird vereinfachend der gleiche Wert verwendet. Für die Dauer $T_{sp-last-info}$ zur Speicherung der Auslastungs-Informationen wurde in [SG01] ein Wert in der Größenordnung von 300 Millisekunden ermittelt. Die Dauer einer Shortest Path Berechnung für ein Netz mit 50 Knoten ist 7.5ms. Für die Dauer einer Lastverteilungsaktion nach dem Mehrpfadansatz wird vereinfachend der gleiche Wert angenommen. Der Sicherheitsfaktor F ist gleich 2.

Aus den Parametern werden nach dem in Abschnitt 5.2.1 beschriebenen Vorgehen die Timer-Werte berechnet. Die folgende Aufzählung gibt die Werte der Timer an:

- Timer zum Austausch von LSP-Auslastungsinformationen $T_{last-info} : 2s$

- Timer zum Aufbau eines LSPs $T_{lsp-sig} : 2.5s$
- Timer für eine LSP-Lastverteilungsaktion $T_{lastverteilung-cal} : 7.5ms$
- Timer für die Lastumverteilung im parallelen Betrieb $T_{lastverteilung-par-cal} : 13s$

5.2.3.2 Auswertung der Konvergenzdauer Zur Analyse des Konvergenzverhaltens wird das Gleichlast-Szenario verwendet, in welchem die Quellen eine mittlere Auslastung von $40 \frac{MBit}{s}$ mit einer Varianz von $120 \frac{MBit}{s}$ erzeugen (siehe hierzu auch Abschnitt 5.1.2.1). Die Eingangslast ist konstant über die Dauer der Simulation. Die initiale Lastverteilung erfolgt mit dem Shortest Path Routing. Die Dauer zur Erkennung der Überlastsituation findet in dieser Untersuchung keine Berücksichtigung und der Lastverteilungsprozess beginnt bei allen Verfahren zum Zeitpunkt 0.5 Sekunden.

Das Konvergenzverhalten wird anhand des minimalen Durchsatzes einer TCP Verbindung in dem Netz über der Zeit bewertet. Die Abbildungen 63 und 64 zeigen den für das Szenario typischen Verlauf des Konvergenzverhaltens der Verfahren über einen Zeitraum von 2500 und 250 Sekunden.

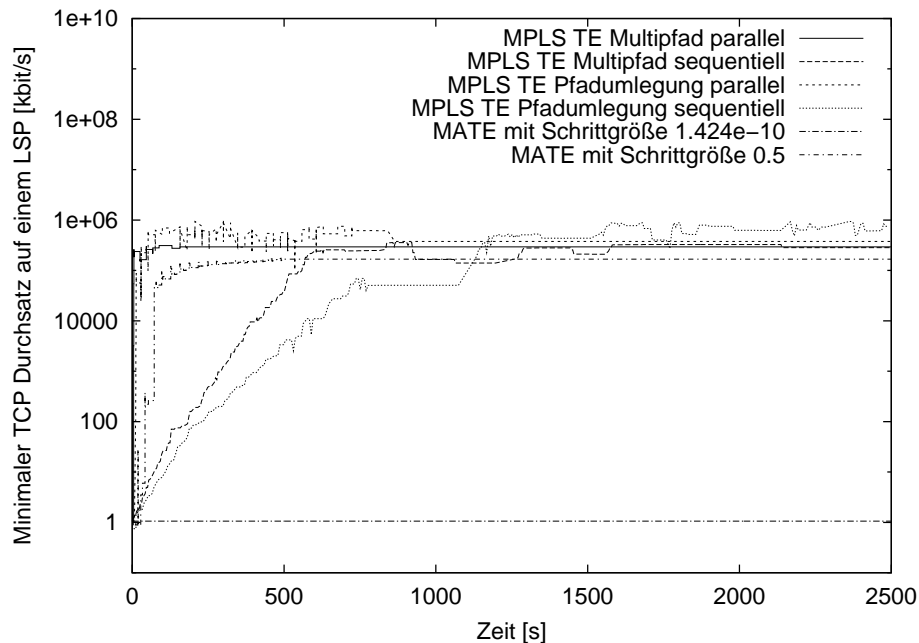


Abbildung 63: Minimaler TCP Durchsatz über der Zeitraum von 2500 Sekunden

Ein wichtiges Ergebnis der Untersuchung ist, dass sowohl der Mehrpfadansatz als auch der Pfadumlegungsansatz wie erwartet erheblich schneller konvergieren als das MATE-Verfahren mit der nach Gleichung 7 berechneten Schrittgröße von $1.424 \cdot 10^{-10}$. Das gilt sowohl für die parallele als auch die sequentielle Durchführung des Lastverteilungsprozesses des Verfahrens. Für das MATE-Verfahren zeigt sich über den Zeitraum von 2500 Sekunden keine merkliche Verbesserung des minimalen TCP Durchsatzes im Netz.

Wird die Schrittgröße des MATE-Verfahrens auf eine für das Szenario ideale Größe von 0.5 gesetzt, konvergiert das Verfahren erheblich schneller. Mit dieser Schrittgröße können jedoch bei anderen Auslastungsszenarien Last-Oszillationen auftreten (siehe hierzu Abschnitt 2.5.3.1). Ein Vergleich mit dem reaktiven verteilten Verkehrslenkungsverfahren

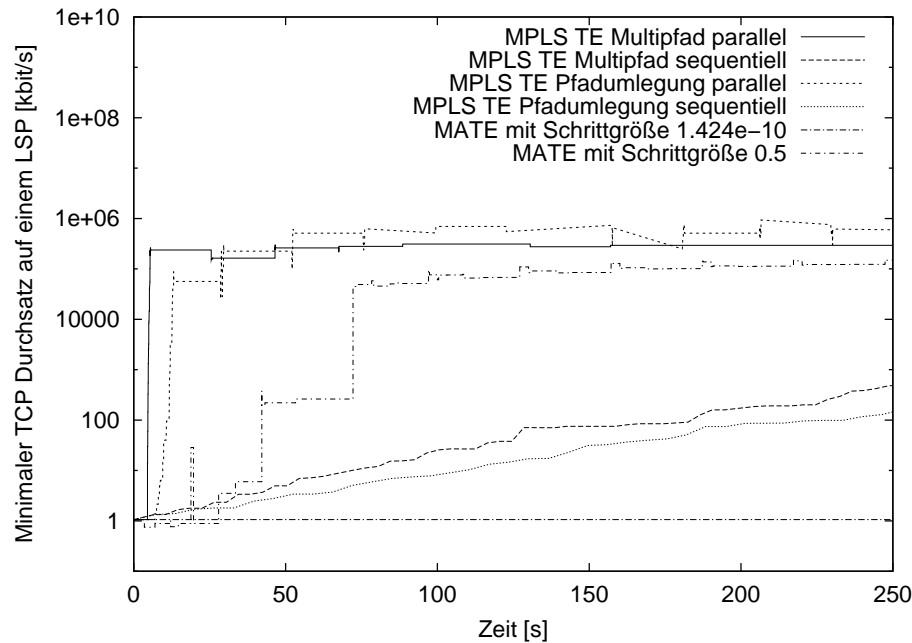


Abbildung 64: Minimaler TCP Durchsatz über der Zeitraum von 250 Sekunden

zeigt, dass der Mehrpfadansatz und der Pfadumlegungsansatz mit paralleler Durchführung des Lastverteilungsprozesses auch eine kürzere Konvergenzdauer als das MATE-Verfahren unter idealen Bedingungen aufweisen.

Beim Vergleich des Mehrpfadansatzes und des Pfadumlegungsansatzes bei paralleler Durchführung der Lastverteilungsaktionen zeigt die Konvergenzdauer keine nennenswerte Unterschiede. Bei sequentieller Durchführung hingegen konvergiert der Mehrpfadansatz schneller als der Pfadumlegungsansatz. Das lässt sich darauf zurückführen, dass im Pfadumlegungsansatz zusätzliche Zeit für den Aufbau der neuen LSPs benötigt wird. In dem 44Knoten-Netz ist der Timer-Wert $T_{lsp-sig}$ zum Abwarten eines Verbindungsaufbaus auf den Wert von 2.5 Sekunden gesetzt.

Ein Vergleich der parallelen und der sequentiellen Durchführung beider Ansätze zeigt eine erhebliche Verkürzung der Konvergenzdauer durch die Parallelisierung des LSP-Lastumverteilungen. Sowohl bei dem Pfadumlegungsansatz als auch bei dem Mehrpfadansatz erreicht der minimale TCP-Durchsatz auf einem LSP in dem Netz bei parallel durchgeführten LSP-Lastumverteilungen bereits nach wenigen Sekunden eine gute Näherung an den maximal erreichbaren TCP-Durchsatz. Wie in Abbildung 64 zu erkennen, benötigt der Mehrpfadansatz ca. 6 Sekunden und der Pfadumlegungsansatz ca. 13 Sekunden. Im sequentiellen Betriebsmodus benötigt der Mehrpfadansatz hingegen ca. 500 Sekunden und der Pfadumlegungsansatz ca. 750 Sekunden (siehe hierzu Abbildung 63), um die Netzperformanz auf einen vergleichbar guten Wert zu steigern. Diese Zahlenwerte sind typisch für das untersuchte Szenario.

5.2.4 Untersuchung eines 14Knoten-Netzes

Dieser Abschnitt führt die zuvor beschriebene Untersuchung der Konvergenzdauer für das 14Knoten-Netz durch. Abschnitt 5.2.4.1 legt die Timer-Werte der reaktiven verteilten Lastverteilungsverfahren fest und Abschnitt 5.2.4.2 präsentiert das Ergebnis der Analyse der Konvergenzdauer.

5.2.4.1 Festlegung der Timer-Werte Dieser Abschnitt beschreibt die Festlegung der Timer-Werte der reaktiven verteilten Verkehrslenkungsverfahren für das 14Knoten-Netz. Das Vorgehen zur Festlegung der Timer-Werte erfolgt analog zu dem in Abschnitt 5.2.3.1 beschriebenen Vorgehen.

Die Topologie-spezifischen Parameter leiten sich aus dem in Abschnitt 5.1.3 beschriebenen 14Knoten-Netz ab. Die verwendeten Werte der Router-spezifischen Parameter entsprechen den in Abschnitt 5.2.3.1 gewählten Werten. Die Ausnahme bildet die Dauer einer Shortest Path Berechnung, welche gemäß der in [SG01] vorgestellten Messung eines 20Knoten-Netzes auf 2.5 Millisekunden gesetzt wird.

Die sich nach dem in Abschnitt 5.2.1 beschriebenen Vorgehen ergebenden Timer-Werte zeigt die folgende Aufzählung:

- Timer zum Austausch von LSP-Auslastungsinformationen $T_{last-info} : 1.5s$
- Timer zum Aufbau eines LSP $T_{lsp-sig} : 1.5s$
- Timer für eine LSP-Lastverteilungsaktion $T_{lastverteilung-cal} : 2.5ms$
- Timer für die Lastumverteilung im parallelen Betrieb $T_{lastverteilung-par-cal} : 0.5s$

5.2.4.2 Auswertung der Konvergenzdauer Die Analyse des Konvergenzverhaltens über der Zeit erfolgt für das in 5.1.3.1 beschriebene Gleichlast-Szenario. Der Mittelwert der Senderaten der Quellen ist $400 \frac{MBit}{s}$ und die Varianz entspricht $1200 \frac{MBit}{s}$. Die Eingangslast ist über den beobachteten Zeitraum konstant.

Die Ergebnisse in den Abbildungen 65 und 66 stellen den für das Szenario typischen Verlauf des minimalen Durchsatzes einer TCP Verbindung über den Zeitraum von 2500 bzw. 250 Sekunden dar.

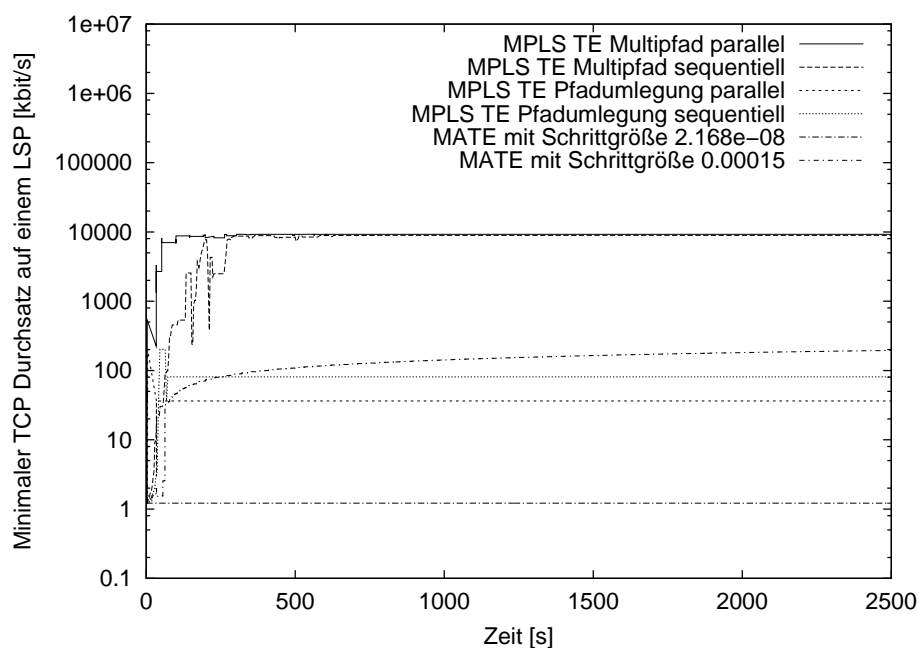


Abbildung 65: Minimaler TCP Durchsatz über der Zeitraum von 2500 Sekunden

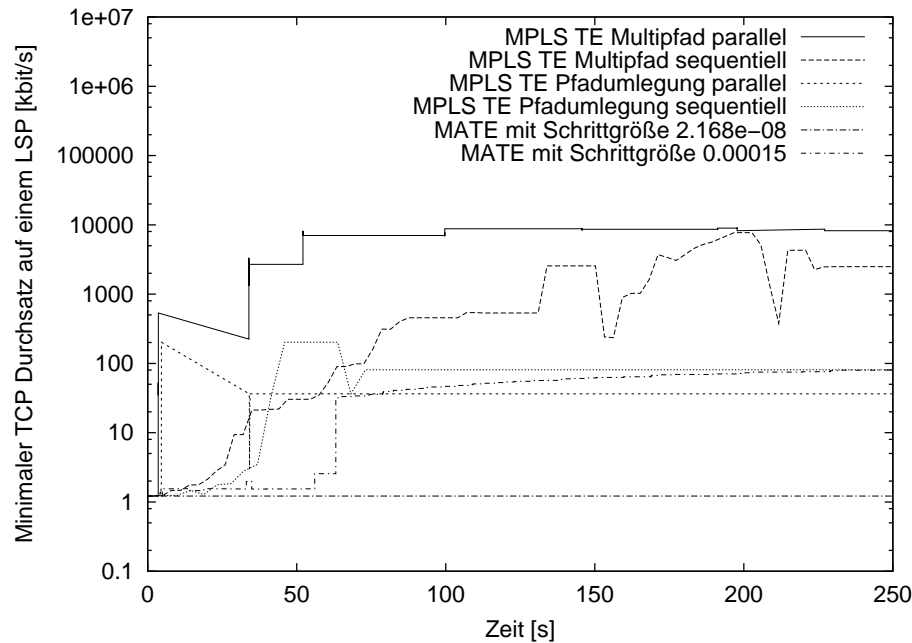


Abbildung 66: Minimaler TCP Durchsatz über der Zeitraum von 250 Sekunden

Zunächst ist festzustellen, dass das Konvergenzverhalten des reaktiven Verkehrslenkungsverfahrens nach dem Pfadumlegungsansatz die in Abschnitt 5.1.3.1 erzielten Ergebnisse bestätigt. Diese erreichen in dem 14Knoten-Netz bedingt durch die zu grobe Granularität der LSP-Last nicht die Netzperformanz des Mehrpfadansatzes.

Der Vergleich mit dem MATE-Verfahren zeigt, dass dieses mit der nach Gleichung 7 berechneten Schrittgröße von $2.168 \cdot 10^{-08}$ erheblich langsamer konvergiert als das in dieser Arbeit entwickelte Verfahren. Über die Dauer von 2500 Sekunden ist keine merkliche Verbesserung des minimalen TCP Durchsatzes im Netz festzustellen. Wird für das MATE-Verfahren die Schrittgröße auf einen für das Szenario idealen Wert von $1.5 \cdot 10^{-4}$ angepasst, konvergiert das Verfahren langsamer als die Verfahren im sequentiellen und parallelen Betriebsmodus. Zudem ist festzustellen, dass die Konvergenzgeschwindigkeit des Verfahrens bei Annäherung an die optimale Lösung erheblich nachlässt. Selbst mit der idealen Schrittgröße nähert das MATE-Verfahren über den Zeitraum von 2500 Sekunden die Performanz des Mehrpfadansatzes nicht an. Zu der idealen Schrittgröße des MATE-Verfahrens ist anzumerken, dass es mit dieser Schrittgröße in anderen Auslastungsszenarien des 14Knoten-Netzes Routing-Oszillationen auslösen kann.

Der Vergleich der Verkehrslenkungsverfahren im parallelen und sequentiellen Betriebsmodus bestätigt die Beobachtung aus Abschnitt 5.2.3.2. Die Verfahren konvergieren im parallelen Betriebsmodus schneller als im sequentiellen Betriebsmodus. Im Unterschied zu dem 44Knoten-Netz ist der Unterschied zwischen den Konvergenzdauern allerdings geringer. Sowohl der Mehrpfadansatz als auch der Pfadumlegungsansatz erzielen bei paralleler Durchführung der Lastverteilung bereits nach 5 Sekunden eine große Steigerung des minimalen TCP Durchsatzes im Netz. Im sequentiellen Betriebsmodus hingegen erreichen der Mehrpfadansatz nach ca. 80 Sekunden und der Pfadumlegungsansatz nach ca. 50 Sekunden einen vergleichbaren minimalen TCP Durchsatz im Netz. In diesem Fall konvergiert der Pfadumlegungsansatz schneller, da die Anzahl der Lastverteilungsaktionen geringer ist.

5.2.5 Zusammenfassende Bewertung der Konvergenzdauer

Der Vergleich mit dem Konkurrenz-Verfahren MATE weist die erheblich geringere Konvergenzdauer des reaktiven Verkehrslenkungsverfahrens nach. Das gilt für alle in dieser Arbeit vorgestellten Betriebsarten des Verfahrens.

Das reaktive Verkehrslenkungsverfahren zeigt im parallelen Betriebsmodus die geringste Konvergenzdauer. Hierbei lassen sich keine nennenswerten Unterschiede zwischen dem Mehrpfad- und dem Pfadumlegungsansatz erkennen. Die Konvergenzdauer hängt im parallelen Betriebsmodus nur schwach von dem jeweiligen Netz- und Auslastungsszenario ab. Die schwache Abhängigkeit ergibt sich durch die in Abhängigkeit des Netzszenarios festzulegenden Timerwerte, welche das Verfahren zur Koordinierung der verteilten TE-Einheiten verwendet.

Die Konvergenzdauer der reaktiven Verfahren im sequentiellen Betriebsmodus hängt von der Anzahl der durchgeführten LSP-Lastverteilungsaktionen ab. Da die Anzahl der LSP-Lastverteilungsaktionen stark mit dem jeweiligen Netz- und Auslastungsszenario variiert, unterliegt auch die Konvergenzdauer starken Schwankungen. Das kann insbesondere in großen Netzen, in welchen eine Vielzahl von Lastverteilungsaktionen notwendig sind, eine erheblich längere Konvergenzdauer verglichen mit dem parallelen Betriebsmodus bedeuten.

5.3 Analyse des Konvergenzverhaltens bei dynamischem Eingangsverkehr

Dieser Abschnitt bewertet das Konvergenzverhalten des reaktiven Verkehrslenkungsverfahrens bei dynamischem Eingangsverkehr. Hierzu vergleicht dieser Abschnitt das reaktive Verkehrslenkungsverfahren mit dem Shortest Path Routing sowie einem proaktiven Verkehrslenkungsverfahren.

Die zur Durchführung der Untersuchung benötigte Modellierung des dynamischen Eingangsverkehrs zeigt Abschnitt 5.3.1. Im Anschluss erläutert Abschnitt 5.3.2 das proaktive Verkehrslenkungsverfahren, mit welchem das reaktive Verkehrslenkungsverfahren verglichen wird. Die Durchführung der Untersuchung sowie die erzielten Ergebnisse werden in Abschnitt 5.3.3 präsentiert.

5.3.1 Verkehrsmodellierung

Das zur Untersuchung des Konvergenzverhaltens des reaktiven Verkehrslenkungsverfahrens bei dynamischem Eingangsverkehr verwendete Quellenmodell beschreibt dieser Abschnitt.

Wie in Abbildung 67 dargestellt, variiert die mittlere Senderate einer Quelle mit einer Normalverteilung. Die Dauer zwischen zwei Ratenänderungen schwankt mit einer negativ exponentiellen Verteilung. Der Mittelwert und die Varianz der Normalverteilung sowie die mittlere Dauer zwischen zwei Ratenänderungen werden je nach Untersuchung angepasst.

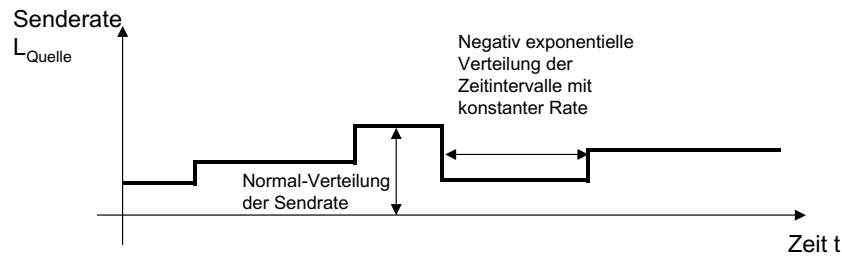


Abbildung 67: Modellierung des dynamischen Eingangsverkehrs

5.3.2 Modellierung eines proaktiven Verkehrslenkungsverfahrens

Dieser Abschnitt beschreibt das proaktive Verkehrslenkungsverfahren, mit welchem die reaktiven Verkehrslenkungsverfahren bezüglich des Konvergenzverhaltens bei dynamischem Eingangsverkehr verglichen werden.

Das proaktive Verfahren basiert auf der Idee des in Kapitel 2.4.3.2 beschriebenen Verfahrens von Wang und Wang [WW99]. Mit einem linearen mathematischen Programm wird für eine gegebene Eingangslast die optimale Lastverteilung berechnet. Die für die Berechnung verwendete Eingangslast entspricht der mittleren erzeugten Last der Quellen. Im Gegensatz zu dem Verfahren von Wang und Wang optimiert das lineare mathematische Programm die Lastverteilung hinsichtlich minimaler Gesamtkosten. Hierfür wird das in Abschnitt 5.1.1 beschriebene lineare mathematische Programm verwendet. Aus der berechneten optimalen Lastverteilung werden die Pfade der LSPs zwischen den Eingangs- und Ausgangs-LSRs abgeleitet. Je nach der berechneten Lastverteilung ist ein Eingangs-LSR über einen oder über mehrere LSPs mit einem Ausgangs-LSR verbunden.

5.3.3 Untersuchung eines 44Knoten-Netzes

Dieser Abschnitt beschreibt die Untersuchung des Konvergenzverhaltens des reaktiven Verkehrslenkungsverfahrens bei dynamischen Eingangsverkehr im 44Knoten-Netz. Die Untersuchung erfolgt in zwei Stufen. Die erste Stufe, welche Abschnitt 5.3.3.1 beschreibt, zeigt die Reaktion der Verkehrslenkungsverfahren auf unvorgesehene Laständerungen an zwei Beispielszenarien. Die zweite Stufe quantifiziert den Einfluss der Verkehrsdynamik auf das Konvergenzverhalten der Verfahren. Die Dynamik des Eingangsverkehrs wird über den mittleren zeitlichen Abstand zwischen zwei Ratenänderungen sowie der Größe der Ratenänderungen variiert. Die Ergebnisse dieser Untersuchung beschreibt Abschnitt 5.3.3.2.

5.3.3.1 Konvergenzverhalten bei großen Laständerungen in zwei Beispielszenarien Die in diesem Abschnitt präsentierte Untersuchung evaluiert die Reaktion der reaktiven Verkehrslenkungsverfahren auf unvorhergesehene Änderungen der Eingangslast an zwei Beispielszenarien.

Beide gehen zunächst von einem Auslastungszenario aus, bei dem die Senderaten aller Quellen einen Mittelwert von $28 \frac{MBit}{s}$ und eine Varianz von $120 \frac{MBit}{s}$ haben. Die mittlere Dauer zwischen zwei Ratenänderungen entspricht 1500 Sekunden. Im ersten Beispielszenario erhöht ein Rand-LSR des 44Knoten-Netzes die mittlere Eingangslast von jedem

ausgehenden LSP um den Faktor 3.5. In zweiten Beispielszenario erhöht ein Rand-LSR die mittlere Eingangslast eines einzelnen LSPs um den Faktor 50. In beiden Fällen tritt die unvorhergesehene Änderung der Eingangslast nach 2000 Sekunden auf. Das Mess-Intervall der reaktiven Verkehrslenkungsverfahren zur Bestimmung der mittleren Auslastung der LSPs ist in den Untersuchungen auf den Wert von 300 Sekunden gesetzt.

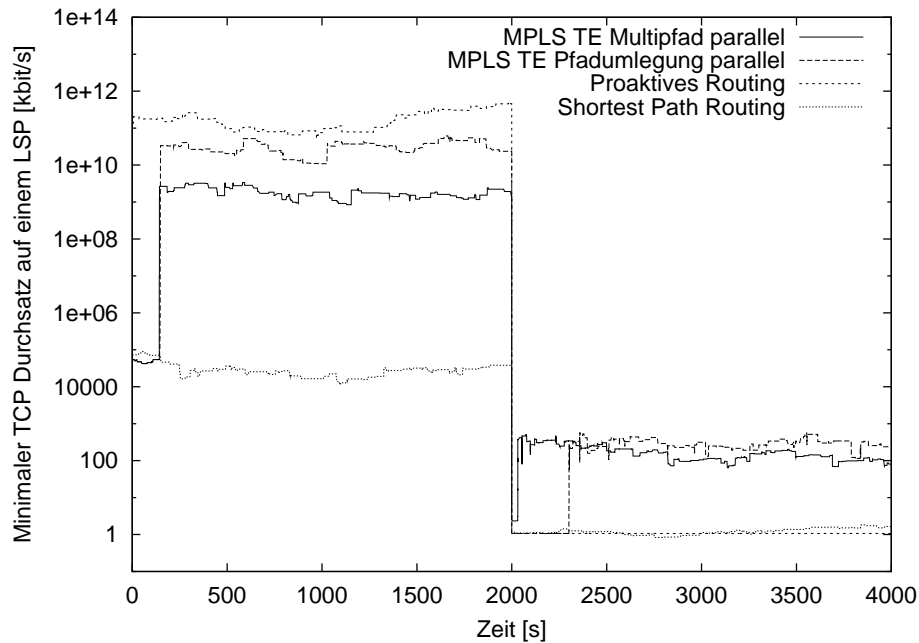


Abbildung 68: Minimaler TCP Durchsatz bei unerwarteter Laständerung nach Fall 1

Die Ergebnisse der Untersuchungen des Konvergenzverhaltens in beiden Fällen sind in den Abbildungen 68 und 69 dargestellt. Es ist zu erkennen, dass sowohl der Multipfadansatz als auch der Pfadumlegungsansatz in beiden Fällen den minimalen TCP Durchsatz innerhalb der Dauer eines Mess-Intervalls erheblich verbessern. In diesem Fall setzt sich die Konvergenzdauer aus der Erkennung der Überlast sowie der Durchführung der Lastumverteilung zusammen.

Die Analyse des proaktiven Verkehrslenkungsverfahrens zeigt, dass dies für die erwartete Eingangslast die beste Netzperformanz erzielt. Treten jedoch große unerwartete Laständerungen auf, bricht die Performanz der proaktiven Verkehrslenkungsverfahren ein. In diesem Fall erreicht das proaktive Verfahren erheblich eine schlechtere Netzperformanz als das reaktive Verfahren.

5.3.3.2 Auswertung des Konvergenzverhaltens in Abhängigkeit von der Verkehrsdynamik Dieser Abschnitt quantifiziert die Abhängigkeit des Konvergenzverhaltens des reaktiven Verkehrslenkungsverfahrens von der Verkehrsdynamik. Dabei wird das Konvergenzverhalten des reaktiven Verkehrslenkungsverfahrens in Abhängigkeit von der mittleren Dauer zwischen zwei Ratenänderungen sowie der Varianz der Ratenänderungen analysiert.

In dem Szenario sendet jede Quelle in dem 44Knoten-Netz mit einer mittleren Senderate von $40 \frac{MBit}{s}$. Zur Auswertung wird das 95% Quantil des minimalen TCP Durchsatzes über der Zeit bestimmt. Dieser Wert gibt die untere Grenze des minimalen TCP Durchsatzes im Netz über 95% des untersuchten Zeitintervalls an. Das Mess-Intervall des reaktiven

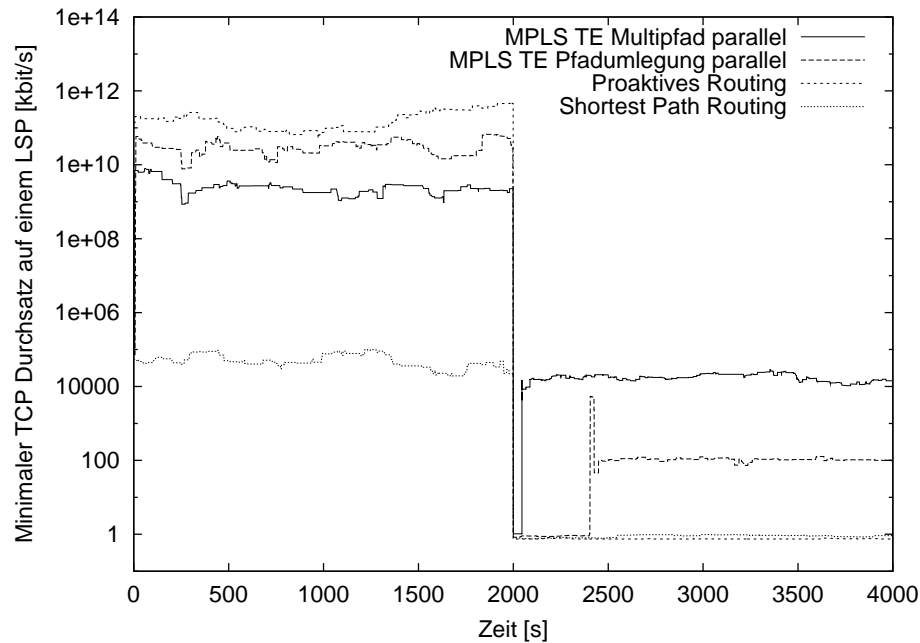


Abbildung 69: Minimaler TCP Durchsatz bei unerwarteter Laständerung nach Fall 2

Verkehrslenkungsverfahren zur Bestimmung der mittleren Auslastung der LSPs ist in den Untersuchungen auf den Wert von 300 Sekunden gesetzt.

Die Ergebnisse der Untersuchung sind in den Abbildungen 70 und 71 dargestellt. Die Analyse der Abhängigkeit von der mittleren Dauer zwischen zwei Ratenänderungen zeigt, dass die reaktiven Verkehrslenkungsverfahren bei Zeitabständen in der Größenordnung des Mess-Intervalls und größer eine erhebliche Steigerung der Performanz gegenüber dem proaktiven Verkehrslenkungsverfahren aufweisen. Bei geringer mittlerer Dauer zwischen zwei Ratenänderungen hingegen erzielen proaktive Verfahren eine bessere Netzperformanz. Das lässt sich auf die beiden in Abschnitt 3.5.2 getroffenen Feststellungen zurückführen. Erstens ist bei Verkehrsschwankungen mit geringem zeitlichen Abstand keine genaue Messung des Verkehrs möglich. Zweitens ist die Dauer zur Anpassung der Lastverteilung im Verhältnis zu den Verkehrsschwankungen zu groß. Für die Untersuchung der Abhängigkeit der Performanz der reaktiven Verfahren von der mittleren Dauer zwischen zwei Ratenänderungen wurde die Varianz der Quellen auf den Wert $600 \frac{MBit}{s}$ gesetzt.

Die Analyse der Abhängigkeit von der Varianz der Ratenänderungen zeigt, dass die Vorteile der reaktiven Verkehrslenkungsverfahren gegenüber proaktiven Verfahren mit steigender Varianz zunehmen. Während bei geringer Varianz und somit einer geringen Abweichung von dem geschätzten Mittelwert kein Unterschied zwischen den Verfahren zu erkennen sind, verbessern die reaktiven Verkehrslenkungsverfahren die Netzperformanz bei einer großen Varianz und somit großen Verkehrsschwankungen gegenüber proaktiven Verfahren erheblich. Bei den Untersuchungen der Abhängigkeit von der Varianz hat die mittlere Dauer zwischen zwei Ratenänderungen dem Wert 3000 Sekunden entsprochen.

5.4 Zusammenfassung der Ergebnisse der Performanz-Analyse

Die quantitative Untersuchung der Performanz des reaktiven Verkehrslenkungsverfahrens wurde in drei Teilen durchgeführt. Der erste Teil hat die Netzperformanz evaluiert, welche

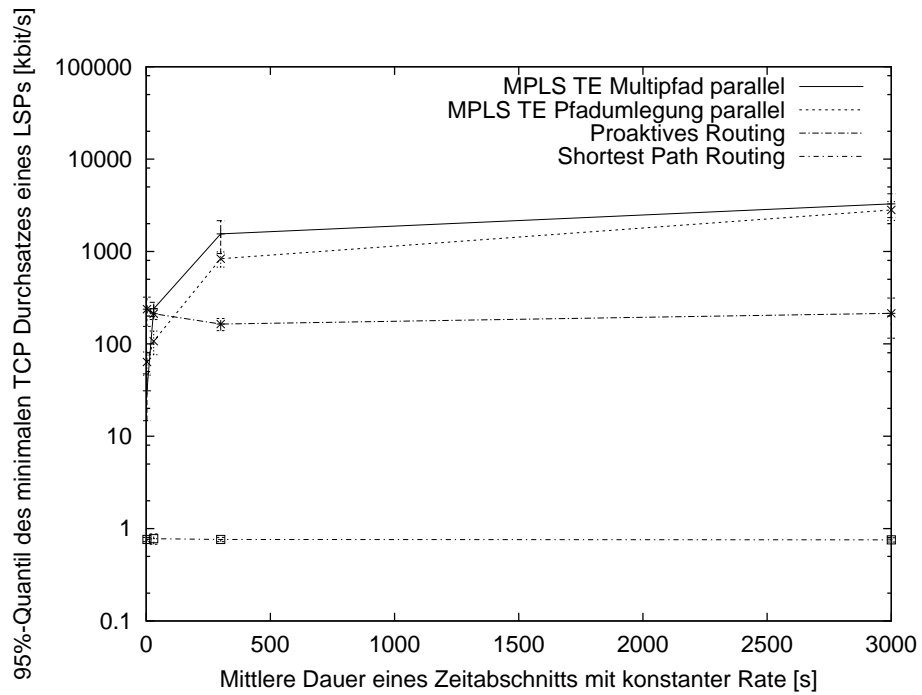


Abbildung 70: Abhängigkeit des TCP Durchsatzes vom mittleren Abstand zwischen zwei Ratenänderungen

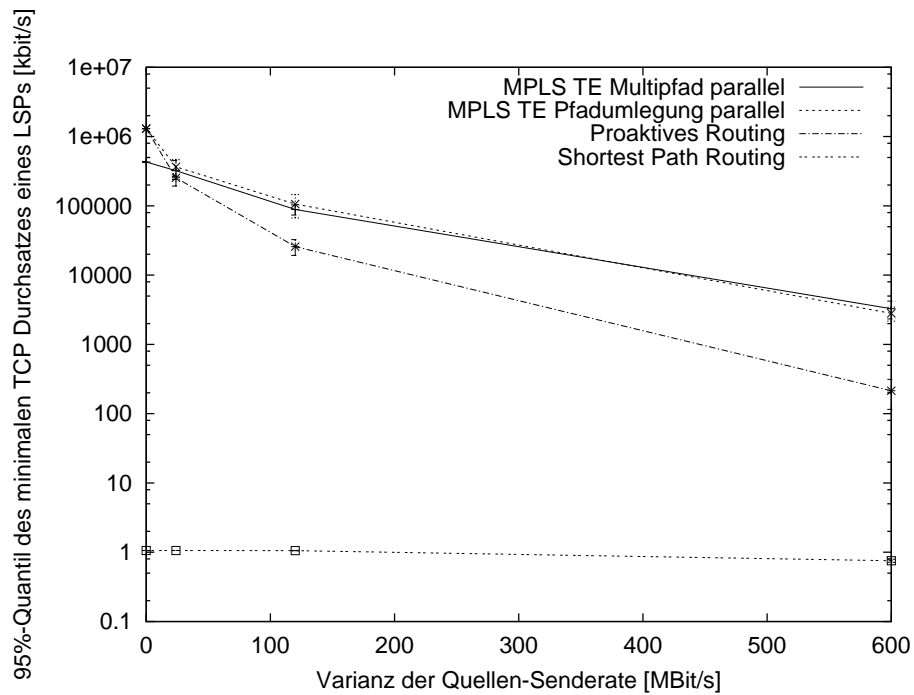


Abbildung 71: Abhängigkeit des TCP Durchsatzes von der Varianz der Ratenänderungen

das reaktive Verkehrslenkungsverfahren nach Durchführung der Lastumverteilung erzielt. Der zweite Teil der Untersuchungen hat sich mit der Analyse der Konvergenzdauer auseinandergesetzt und der dritte Teil hat das Konvergenzverhalten bei dynamischen Eingangsverkehr bewertet.

Die Ergebnisse des ersten Teils der Untersuchungen zeigen, dass das reaktive Verkehrslenkungsverfahren bei Verwendung des Mehrpfadansatzes unabhängig vom Netz- und Auslastungsszenario die optimale Lastverteilung sehr gut annähert. Diese Netzperformanz erreicht das Verfahren nach dem Mehrpfadansatz bereits mit der geringen Anzahl von zwei bzw. drei LSPs pro LSP-Fächer. Bei Verwendung des Pfadumlegungsansatzes zeigt das Verfahren ein vergleichbares Ergebnis, wenn das Netz ausreichend groß ist und sich die gesamte Netzlast auf viele LSPs verteilt. In kleinen Netzen bzw. in Netzen, in welchen sich die Last nur auf wenige LSPs aufteilt, erreicht der Pfadumlegungsansatz hingegen nicht die Performanz des Mehrpfadansatzes. Das lässt sich mit Granularität der Last begründen, welche für den Pfadumlegungsansatz durch die Anzahl der LSPs im Netz vorgegeben ist. Die Performanz des Pfadumlegungsansatzes ist im Gegensatz zum Mehrpfadansatz abhängig vom Netz- und Auslastungsszenario.

Der Vergleich des reaktiven Verkehrslenkungsverfahrens zum Shortest Path Routing weist eine erhebliche Steigerung der Netzperformanz nach. Somit hat in den untersuchten Szenarien der Einsatz des Verfahrens bei gleicher Netzperformanz eine Erhöhung der Eingangslast zwischen 25% und 100% zugelassen.

Die quantitative Analyse der Konvergenzdauer im zweiten Teil der Untersuchung hat gezeigt, dass das reaktive Verkehrslenkungsverfahren erheblich schneller konvergiert als das Konkurrenz-Verfahren MATE. Diese Aussage gilt für alle vorgestellten Betriebsarten des Verfahrens.

Die geringste Konvergenzdauer hat das reaktiven Verfahren im parallelen Durchführungsmodus gezeigt, wobei keine Unterschiede zwischen dem Mehrpfad- und dem Pfadumlegungsansatz festgestellt wurden. Die Konvergenzdauer ohne Berücksichtigung der Überlasterkennung lag in der Größenordnung einiger Sekunden. Ein weiterer Vorteil des Verfahrens in diesem Betriebsmodus ist, dass lediglich eine schwache Abhängigkeit vom Netz- und Auslastungsszenario vorliegt.

Die Konvergenzdauer im sequentiellen Durchführungsmodus hängt hingegen stark von dem Netz- und Auslastungsszenario ab. Das ist dadurch bedingt, dass diese mit der Anzahl der durchgeführten LSP-Lastverteilungsaktionen variiert. In den untersuchten Szenarien lag die Konvergenzdauer in der Größenordnung mehrere Minuten.

Die Analyse des Konvergenzverhaltens bei dynamischem Eingangsverkehr im dritten Teil der Untersuchung unterstreicht die Vorteile des reaktiven Verfahrens gegenüber den proaktiven Verkehrslenkungsverfahren bei unerwarteten Lastschwankungen. In den gezeigten Beispielszenarien hat die Verwendung des reaktiven Verkehrslenkungsverfahrens eine erhebliche Steigerung der Netzperformanz gegenüber dem proaktivem Verfahren erzielt. Aus den durchgeführten Untersuchungen lässt sich zudem ableiten, dass mit zunehmender Größe der Lastschwankungen die Vorteile des reaktiven Verfahrens gegenüber proaktiven Verfahren zunehmen. Hierbei muss allerdings die Voraussetzung gelten, dass der zeitliche Abstand zwischen den Schwankungen der Auslastung groß gegenüber dem Mess-Intervall des reaktiven Verfahrens sein muss.

6 Zusammenfassung

Der wirtschaftlicher Betrieb von IP-Netzen rückt für die Betreiber immer mehr in den Mittelpunkt. Diese Anforderung gilt es unter den Aspekten eines stetig wachsenden Datenvolumens sowie einer zunehmenden Verkehrsdynamik zu bewältigen. Um eine optimale Ausnutzung der Netz-Ressourcen zu erzielen, kommt lastabhängigen Verkehrslenkungsverfahren in IP-Netzen eine immer größere Bedeutung zu.

In den letzten Jahren lag der Fokus auf der Untersuchung proaktiver Verkehrslenkungsverfahren. Diese Verfahren optimieren die Lastverteilung in periodischen Zeitabständen von Tagen bzw. Wochen basierend auf der Kenntnis der Eingangsverkehrs. Der Nachteil der proaktiven Verfahren ist, dass diese keine schnelle Reaktion auf unvorhersehbare Lastschwankungen zulassen. Da bedingt durch die Zunahme der Verkehrsdynamik von einem vermehrten Auftreten dieser auszugehen ist, liegt der Schwerpunkt der Arbeit auf der Untersuchung reaktiver Verkehrslenkungsverfahren. Reaktive Verfahren basieren auf der Messung der aktuellen Netzauslastung in Zeitabständen von Sekunden bis Minuten und erlauben damit eine schnelle Anpassung der Lastverteilung in Überlastsituationen. Bei statischem Eingangsverkehr lässt sich zudem mit reaktiven Verfahren die optimale Lastverteilung annähern und somit eine zu proaktiven Verfahren vergleichbare Netzperformanz erzielen.

Die Analyse des Stands der Technik hat gezeigt, dass existierende reaktive Verkehrslenkungsverfahren Schwächen aufweisen. Das bekannteste Verfahren MPLS Adaptive Traffic Engineering MATE zeigt eine geringe Konvergenzgeschwindigkeit zur Anpassung der Lastverteilung nach dem Auftreten einer Überlastsituation. Dieser Effekt zeigt sich insbesondere bei Anwendung der Verfahren in großen Netzen. Das Verfahren Optimization Integrated Adaptive Traffic Engineering OpIATE erzeugt bei der Bestimmung der neuen Lastverteilung einen hohen Berechnungsaufwand, welcher insbesondere bei einer verteilten Realisierung zu einer hohen zusätzlichen Prozessorlast in den Netzknoten führen kann. Aus diesem Grund wurde im Rahmen dieser Arbeit ein neues verteiltes reaktives Verkehrslenkungsverfahren entwickelt. Dieses Verfahren basiert auf der Verwendung von Multiprotocol Label Switching MPLS und ist für die Optimierung der Lastverteilung des Best Effort IP-Verkehr konzipiert.

Die prinzipielle Funktionsweise des entwickelten Verfahrens beruht auf einer Vollvermaschung der Randknoten über MPLS-Verbindungen. Da das Verfahren sowohl im Pfadumlegungsansatz als auch im Mehrpfadansatz betrieben werden kann, sind die Randknoten entweder einfach oder mehrfach vollvermascht. Jeder Randknoten bestimmt periodisch die mittleren Auslastungen der abgehenden MPLS-Verbindungen über eine Dauer von Sekunden bis Minuten. Die gemessenen Auslastungen der MPLS-Verbindungen verteilen die Randknoten über einen Flutungs-Mechanismus an alle Knoten im Netz. Aus der Kenntnis der Auslastungen sowie der Pfade aller MPLS-Verbindungen bestimmt jeder Knoten im Netz die aktuelle Netzauslastung. Erkennen die Netzknoten Link-Auslastungen, welche über einer Überlastschwelle liegen, wird die Umverteilung der Netzlast angestoßen. Bei Verwendung des Pfadumlegungsansatzes basiert diese auf dem Rerouting einzelner MPLS-Verbindungen. Bei Verwendung des Mehrpfadansatzes erfolgt diese durch die Umverteilung der Last auf die bestehenden MPLS-Verbindungen. Die Lastumverteilung endet nach der Auflösung der Überlastsituationen auf den kritischen Links. Die Koordination der Lastverteilungsaktionen einzelner MPLS-Verbindungen, welche durch die verteilten Randknoten ausgeführt werden, basiert auf der Kenntnis einer identischen Netzausla-

stung. Unter Verwendung des gleichen Algorithmus legen die Randknoten die gleiche Reihenfolge der notwendigen Lastverteilungsaktionen einzelner MPLS-Verbindungen fest. Die Ausführung der Lastumverteilungsaktionen kann sequentiell oder parallel erfolgen.

Die durchgeführte Analyse des Verfahrens zeigt die Vorteile des entwickelten reaktiven Verkehrslenkungsverfahrens gegenüber den Konkurrenzverfahren. Verglichen mit dem OpIATE-Verfahren konnte die Berechnungskomplexität erheblich reduziert werden. Gegenüber dem Verfahren MATE wurde eine deutliche Steigerung der Konvergenzgeschwindigkeit erzielt. Zudem konnten die Performanzvorteile des reaktiven Verfahrens bei dem Auftreten von unvorhersehbaren Lastschwankungen im Vergleich zu den proaktiven Verfahren nachgewiesen werden.

Zur quantitativen Bewertung des Konvergenzverhaltens des Verfahrens wurde eine simulative Untersuchung anhand von Fallstudien durchgeführt. Die Ergebnisse der Untersuchung zeigen, dass der Einsatz des Verfahrens eine erhebliche Steigerung der Netzperformanz gegenüber dem klassischen Shortest Path Routing bedeutet. Unter Verwendung des Verfahrens konnte die Eingangslast ohne eine Reduktion der Netzperformanz je nach untersuchten Szenario zwischen 25% und 100% gesteigert werden.

Die Ergebnisse zeigen ferner, dass das Verfahren bei Verwendung des Mehrpfadansatzes in allen Fällen die Netzperformanz der optimalen Lastverteilung sehr gut annähert. Diese Performanz erreicht der Mehrpfadansatz bereits mit der geringen Anzahl von zwei bzw. drei MPLS-Verbindungen pro Eingangs-Ausgangsknotenpaar. Bei Verwendung des Pfadumlegungsansatzes konnte das Verfahren lediglich in großen Netzen mit fein-granularer Verteilung der Last auf die MPLS-Verbindungen eine vergleichbare Performanz erzielen. In kleinen Netzen hingegen verschlechtert sich bei Verwendung des Pfadumlegungsansatzes die Performanz deutlich gegenüber dem Mehrpfadansatz. Das ist bedingt durch die geringere Anzahl an MPLS-Verbindungen, welche bei dem Pfadumlegungsansatz die Granularität der Lastverteilung vorgeben. Daher ist der Mehrpfadansatz die zu empfehlende Betriebsart bei der Verwendung des Verfahrens. Der Einsatz des Pfadumlegungsansatzes ist nur in Netzen sinnvoll, in welchen auf Grund der begrenzten Anzahl möglicher MPLS-Verbindungen der Mehrpfadansatz nicht betrieben werden kann.

Die quantitative Analyse der Konvergenzdauer zeigt, dass das entwickelte Verfahren nach der Überlasterkennung eine Annäherung an die optimale Lastverteilung innerhalb weniger Sekunden erreicht. Hierzu muss das Verfahren im parallelen Durchführungsmodus betrieben werden, in welchem die Randknoten die Lastverteilungsaktionen gleichzeitig durchführen. Die Konvergenzdauer hängt dabei nur schwach von dem Netz- und Auslastungsszenario ab. Im sequentiellen Durchführungsmodus hingegen, in welchem die Lastverteilungsaktionen hintereinander ausgeführt werden, können abhängig vom Netz- und Auslastungsszenario erheblich längere Konvergenzdauer im Bereich einiger Minuten entstehen. Zusammenfassend ist festzustellen, dass die parallele Ausführung der Lastverteilungsaktionen in Situationen mit starker Überlast sinnvoll ist, in welchen eine nachhaltige Störung der Datenübertragung im Netz vorliegt und eine schnelle Auflösung der Überlastsituation erforderlich ist. In Situationen mit schwacher Überlast, in welchen die Lastumverteilung einen vorbeugenden Charakter hat, ist die sequentielle Durchführung der Lastverteilungsaktionen zu bevorzugen. Hier ist der Vorteil des sequentiellen Durchführungsmodus relevant, dass dieser die Einhaltung der Reihenfolge der Lastverteilungsaktionen garantiert und somit mögliche Performanzeinbrüche während der Lastumverteilung verhindert.

Zur simulativen Untersuchung des reaktiven Verkehrslenkungsverfahrens wurde im Rah-

men dieser Arbeit eine skalierbare Simulations-Methode entwickelt. Diese erlaubt die Untersuchung der Verfahren auch in großen Netzen. Die Grundlage der Methode ist die ratenbasierten Modellierung des Eingangsverkehrs. Unter Verwendung dieser Mittelwerte der Eingangsraten, welche über Zeiträume variabler Länge konstant sind, erlaubt die Methode die Bestimmung der Lastverteilung im Netz. Hieraus lässt die Methode eine Abschätzung der Netzperformanz unter Berücksichtigung der Eigenschaften des Transport Control Protocol TCP zu. Damit ermöglicht diese Methode erstmals eine Bewertung reaktiver Verkehrslenkungsverfahren in realistischen Netzszenarien.

Literatur

- [A99] D. Awduche. *MPLS and Traffic Engineering in IP Networks*. IEEE Communications Magazine, Dezember 1999
- [A97] ATM Forum Technical Committee. Multiprotocol over ATM - Version 1.0. ATM Forum, af-mpoa-0087.000, Juli 1997
- [A95] ATM Forum Technical Committee. LAN Emulation over ATM - Version 1.0. ATM Forum, af-lane-0021.000, Januar 1995
- [ABG01] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, G. Swallow. *RSVP-TE: Extensions to RSVP for LSP Tunnels*. RFC3209, Dezember 2001
- [ACE02] D. Awduche, A. Chiu, A. Elwalid, I. Widjaja, X. Xiao. *Overview and Principles of Internet Traffic Engineering*. RFC3272, Mai 2002
- [ADF01] L. Andersson, P. Doolan, N. Feldman, A. Fredette, B. Thomas. *LDP Specification*. RFC3036, Januar 2001
- [AGB94] R. Albrightson, J.J. Garcia-Luna-Aceves, J. Boyle. *EIGRP - a fast routing protocol based on distance vectors*. Network/Interop, Las Vegas, USA, Mai 1994
- [AMA99] D. Awduche, J. Malcolm, J. Agogbua, M. O'Dell, J. McManus. *Requirements for Traffic Engineering Over MPLS*. RFC2702, September 1999
- [AT02] D. Applegate, M. Thorup. *Load Optimal MPLS routing with N+M labels*. IEEE INFOCOM, New York, USA, Juni 2002
- [AWK99] G. Apostolopoulos, D. Williams, S. Kamat, R. Guerin, A. Orda, T. Przygienda. *QoS Routing Mechanisms and OSPF Extensions*. RFC2676, August 1999
- [B03a] L. Berger. *Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description*. RFC3471, Januar 2003
- [B03b] S. Butenweg. *Two distributed reactive MPLS Traffic Engineering mechanisms for throughput optimization in Best Effort MPLS networks*. 8. IEEE Symposium on Computers and Communications, Kemer, Türkei, Juni/Juli 2003
- [B01] S. Butenweg. *A simulation environment for the performance evaluation of MPLS Traffic Engineering solutions in IP networks*. 5. World Multiconference on Systemics, Cybernetics and Informatics, (SCI), Orlando, USA, Juli 2001
- [B99] R. Bhandari. *Survivable Networks: Algorithms for Diverse Routing*. Kluwer Academic Publishers, November 1999
- [B97] M. Bahr. *Entwicklung eines Simulatorkerns für einen ratenbasierten ATM-Traffic-Simulator*. Diplomarbeit, Universität Rostock, Fachbereich Informatik, Rostock, November 1997
- [BB03] M. Bahr, S. Butenweg. *On Rate-based Simulation of Communication Networks*. Conference on Design, Analysis, and Simulation of Distributed Systems (DASD), Orlando, USA, April 2003

- [BBC98] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss. *An Architecture for Differentiated Services*. RFC2475, Dezember 1998
- [BFO98] S. Bhatt, R. Fujimoto, A. Ogielski, K. Perumalla. *Parallel Simulation Techniques for Large-Scale Networks*. IEEE Communications Magazine, August 1998
- [BG87] D. Bertsekas, R. Gallager. *Data Networks*. Prentice-Hall, 1987
- [BS79] I.N. Bronstein, K.A. Semendjajew. *Taschenbuch der Mathematik*. Verlag Harri Deutsch, 1979
- [BT00] S. Butenweg, M. Tichatschek. *Comparison of the Rate Based Simulation and the Packet Based Simulation*. Advanced Simulation Technologies Conference (ASTC), Washington, USA, April 2000
- [BZB97] R. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin. *Resource ReSerVation Protocol (RSVP) - Version 1, Functional Specification*, RFC2205, September 1997
- [C90] R. Callon. *Use of OSI IS-IS for Routing in TCP/IP and Dual Environments*. RFC1195, Dezember 1990
- [CC00] S. McCreary, K. Claffy. *Trends in Wide Area IP Traffic Patterns: A View from Ames Internet Exchange*. ITC Specialist Seminar on IP Traffic Measurement, Modeling, and Management, Monterey, USA, September 2000
- [CLR90] T. H. Cormen, C. E. Leiserson, R. L. Rivest. *Introduction to Algorithms*. MIT Press, 1990
- [CDM01] A. Conta, P. Doolan, A. Malis. *Use of Label Switching on Frame Relay Networks Specification*. RFC3034, Januar 2001
- [CMT98] K. Claffy, G. Miller, and K. Thompson. *The nature of the beast: recent traffic measurement from an Internet backbone*. INET, Geneva, Switzerland, July 1998
- [CN98] S. Chen, K. Nahrstedt. *An overview of quality of service routing for next-generation high-speed networks: Problems and solutions*. IEEE Network Magazine, Vol. 12, Nr. 6, November/Dezember 1998
- [DAJ00a] E. Dinan, D.O. Awduche, B. Jabbari. *Analytical Framework for Dynamic Traffic Partitioning in MPLS Networks*. IEEE International Conference on Communications, ICC 2000, New Orleans, USA, Juni, 2000
- [DAJ00b] E. Dinan, D.O. Awduche, B. Jabbari. *Optimal Traffic Partitioning in MPLS Networks*. IFIP Networking 2000, Paris, France, Mai 2000
- [DLM01] B. Davie, J. Lawrence, K. McCloghrie, E. Rosen, G. Swallow, Y. Rekhter, P. Doolan. *MPLS using LDP and ATM VC Switching*. RFC3035, Januar 2001
- [E01] J. Eberspächer. *Die Zukunft des Internet im Zeichen der Konvergenz*. In: Kubicek et al.: Internet@Future; Jahrbuch Telekommunikation und Gesellschaft 2001, Hüthig Verlag, 2001
- [E02] J. Eberspächer. *Systemtechnische Grundlagen*. In: V. Jung, H. Warnecke (Hrsg.): Handbuch für die Telekommunikation, Springer-Verlag, 2002

- [EJS01] A. Elwalid, C. Jin, S. Low, I. Widjaja. *MATE: MPLS Adaptive Traffic Engineering*. IEEE INFOCOM, Anchorage, Alaska, April 2001
- [EJS02] A. Elwalid, C. Jin, S. Low, I. Widjaja. *MATE: MPLS Adaptive Traffic Engineering*. Computer Networks - The International Journal of Computer and Telecommunications Networking, vol. 40,no. 6, pp. 695-709, Dezember 2002
- [ERP02] M. Ericsson, M.G.C. Resende, P.M. Pardalos. *A Genetic Algorithm for the Weight Settings Problem in OSPF Routing*. Journal of Combinatorial Optimization, 2002
- [FDD02] F. Le Faucheur, L. Wu, B. Davie, S. Davari, P. Vaananen, R. Krishnan, P. Cheval, J. Heinanen. *Multiprotocol Label Switching (MPLS) Support of Differentiated Services*. RFC3270, Mai 2002
- [FGk93] R. Fourer, D.M. Gay, B.W. Kernighan. *AMPL: A Modeling Language for Mathematical Programming*. The Scientific Press Series, 1993
- [FGL01] A. Feldmann, A. Greenberg, C. Lund, N. Reingold, J. Rexford, F. True. *Deriving traffic demands for operational IP networks: Methodology and experience*. IEEE/ACM Transactions on Networking, Vol.9, No. 3, pp. 265-279, Juni 2001
- [FGL00] A. Feldmann, A. Greenberg, C. Lund, N. Reingold, J. Rexford. *NetScope: Traffic Engineering for IP Networks*. IEEE Networks März/April2000
- [FT00] B. Fortz, M. Thorup. *Internet traffic engineering by optimizing OSPF weights*. IEEE INFOCOM, Tel-Aviv, Israel, März 2000
- [G90] A. Girard. *Routing and Dimensioning in Circuit-Switched Networks*. Addison-Wesley, 1990
- [GKK88] R.J. Gibbens, F.P. Kelly, P.B. Key. *Dynamic Alternative Routing - Modelling and Behaviour*. 12. International Teletraffic Congress, Juni 1988
- [GSE00] R.J. Gibbens, S.K. Sargood, C. Van Eijl, F.P. Kelly, H. Azmoodeh, R.N. Macfadyen, N.W. Macfadyen. *Fixed-point models for the end-to-end performance analysis of IP networks*. 13th ITC Specialist Seminar: IP Traffic Measurement, Modeling and Management, Monterey, USA, September 2000
- [GKL02] T. Güven, C. Kommareddy, R. J. La, M. A. Shayman, B. Bhattacharjee. *Measurement Based Optimal Multi-path Routing*. UM Institute of Advanced Computing Sciences, Tech. Rep. UMIACS-TR-2003-69, July 2003.
- [GKR95] R. Gawlick, C. Kalmanek, K.G. Ramakrishnan. *On-line Routing for Permanent Virtual Circuits*. IEEE Infocom, Bosten, USA, 1995
- [GM02] J. Glasmann, H. Müller. *Resource Management Architecture for Realtime Traffic in Intranets*. Networks 2002, Joint IEEE International Conferences ICN and ICWLHN, Atlanta, USA, August 2002
- [GME02] J. Glasmann, H. Müller, J. Eberspächer. *Ressourcenmanagement-Architektur für Echtzeitverkehre in Intranets*. In: Praxis der Informationsverarbeitung und Kommunikation (PIK), 2002

- [GOW97] R. Guerin, A. Orda, D. Williams. *QoS Routing Mechanisms and OSPF Extensions*. 2. IEEE Global Internet Mini-Conference, pages 1903-1908, Phoenix, USA, November 1997
- [H96] C. Huitema. *Routing in the Internet*. Prentice Hall, 1996
- [H91] C. Hedrick. *An Introduction to IGRP*. The State University of New Jersey Technical Publication, Laboratory for Computer Science Research, August 1991
- [H88] C. Hedrick. *Routing Information Protocol*. RFC1058, Juni 1988
- [JAC02] B. Jamoussi, L. Andersson, R. Callon, R. Dantu, L. Wu, P. Doolan, T. Worster, N. Feldman, A. Fredette, M. Girish, E. Gray, J. Heinanen, T. Kilty, A. Malis. *Constraint-Based LSP Setup using LDP*. RFC3212, Januar 2002
- [KB03] S. Köhler, A. Binzenhöfer. *MPLS Traffic Engineering in OSPF Networks - a combined Approach*. Technischer Report No. 304, Universität Würzburg, Deutschland, Februar 2003
- [KNE97] Y. Katsube, K. Nagami, H. Esaki. *Toshiba's Router Architecture Extensions for ATM : Overview*. RFC2098, Februar 1997
- [L99] K. Lindberger. *Balancing Quality of Service, Pricing and Utilisation in Multiservice Networks with Stream and Elastic Traffic*. International Teletraffic Congress, Edinburgh, UK, Juni 1999
- [L94] M. Laubach. *Classical IP and ARP over ATM*. RFC1577, Januar 1994
- [LKP98] J. Luciani, D. Katz, D. Piscitello, B. Cole, N. Doraswamy. *NBMA Next Hop Resolution Protocol*. RFC2332, April 1998
- [M93] G. Malkin. *RIP Version 2: Carrying Additional Information*. RFC1388, January 1993
- [M98] J. Moy. *OSPF Version 2*. RFC2328, April 1998
- [MKS02] J. Milbrandt, S. Köhler, D. Stähle, L. Berry. *Decomposition of Large IP Networks for Routing Optimization*. Technischer Report No. 293, Universität Würzburg, Deutschland, Februar 2002
- [MR01] D. Mitra, K.G. Ramakrishnan. *Techniques for Traffic Engineering of Multiservice, Multipriority Networks*. Bell Labs Technical Journal, Januar-Juni 2001
- [MR99] D. Mitra, K.G. Ramakrishnan. *A Case Study of Multiservice, Multipriority Traffic Engineering Design for Data Networks*. IEEE GLOBECOM 99, Rio de Janeiro, Brasilien, Dezember 1999
- [MS98] Q. Ma, P. Steenkiste. *Routing Traffic with Quality-of-Service Guarantees in Integrated Services Networks*. Workshop on Network and Operating Systems Support for Digital Audio and Video, Juli 1998
- [MS97a] Q. Ma, P. Steenkiste. *Quality-of-Service Routing for Traffic with Performance Guarantees*. IFIP Fifth International Workshop on Quality of Service, New York, USA, Mai 1997

- [MS97b] Q. Ma, P. Steenkiste. *On Path Selection for Traffic with Bandwidth Guarantees*. IEEE International Conference on Network Protocols, Atlanta, USA, Oktober 1997
- [MSM97] M. Mathis, J. Semke, J. Mahdavi. *The Macroscopic Behavior of the Congestion Avoidance Algorithm*. Computer Communication Review, Vol. 27, No. 3, Juli 1997
- [MSZ96] Q. Ma, P. Steenkiste, H. Zhang. *Routing High-bandwidth Traffic in Max-min Fair Share Networks*. ACM SIGCOMM'96, pp 206-217, Stanford, USA, August 1996
- [NEH96a] P. Newman, W. L. Edwards, R. Hinden, E. Hoffman, F. Ching Liaw, T. Lyon, G. Minshall. *Ipsilon Flow Management Protocol Specification for IPv4 Version 1.0*, RFC1953, Mai 1996
- [NEH96b] P. Newman, W. L. Edwards, R. Hinden, E. Hoffman, F. Ching Liaw, T. Lyon, G. Minshall. *Transmission of Flow Labelled IPv4 on ATM Data Links Ipsilon Version 1.0*. RFC1954, Mai 1996
- [P95] J.M. Pitts. *Cell-rate modelling for accelerated simulation of ATM at the burst level*. IEE Proceedings-Communications, Vol. 142, No. 6, Dezember 1995
- [P93] J.M. Pitts. *Cell-rate Simulation Modelling of Asynchronous Transfer Mode Telecommunications Networks*. Doktorarbeit, Universität von London, Juli 1993
- [P81a] J. Postel. *Transmission Control Protocol*. RFC793, September 1981
- [P81b] J. Postel. *Internet Protocol*. RFC791, September 1981
- [P80] J. Postel *User Datagram Protocol*. RFC 768, August 1980
- [PF97] V. Paxson, S. Floyd. *Why We Don't Know How To Simulate The Internet*. Winter Simulation Conference, Atlanta, USA, 1997
- [PFT98] Jitedra Padhye, Victor Firoiu, Don Towsley, and Jim Krusoe. *Modeling TCP throughput: A simple model and its empirical validation*. ACM SIGCOMM, pages 303-314, Vancouver, CA, September 1998
- [R03] A. Riedl. *Optimized Routing Adaptation in IP Networks Utilizing OSPF and MPLS*. IEEE International Conference On Communication, Anchorage, USA, Mai 2003
- [R02] A. Riedl. *A Hybrid Genetic Algorithm for Routing Optimization in IP Networks Utilizing Bandwidth and Delay Metrics*. IEEE Workshop on IP Operations and Management (IPOM), Dallas, USA, ?? 2002.
- [RBP00] Anton Riedl, Maren Perske, Thomas Bauschert, Andreas Probst. *Investigation of the M/G/R Processor Sharing Model for Dimensioning of IP Access Networks with Elastic Traffic*. First Polish-German Teletraffic Symposium PGTS 2000, Dresden, September 2000
- [RDK97] Y. Rekhter, B. Davie, D. Katz, E. Rosen, G. Swallow. *Cisco Systems Tag Switching Architecture Overview*. RFC2105, Februar 1997
- [RWL95] Y. Rekhter, T.J. Watson, T. Li. *A Border Gateway Protocol 4 (BGP-4)*. RFC1771, März 1995

- [RVC01] E.C. Rosen, A. Viswanathan, R. Callon. *Multiprotocol Label Switching Architecture*. RFC3031, Januar 2001
- [S99a] G. Swallow. *MPLS Advantages for Traffic Engineering*. IEEE Communications Magazine, Dezember 1999
- [S99b] G. Sigmund. *Technik der Netze*. Hüthig Verlag, 1999
- [S97] S. Sainio. *Comparison of a rate-based ATM simulator with a cell-based ATM simulator*. Master Thesis, Helsinki University of Technology, Department of Electrical and Communications Engineering, November 1997
- [S96] R. Siebenhaar. *Verkehrslenkung und Kapazitätsanpassung in ATM-Netzen mit virtuellen Pfaden*. Dissertation, Technische Universität München, Lehrstuhl für Kommunikationsnetze, Herbert Utz Verlag, Juni 1996
- [S94] W. Stevens. *TCP/IP Illustrated, Vol.1 The Protocols*. Addison-Wesley, 1994
- [SG01] A. Shaikh, A. Greenberg. *Experience in Black-box OSPF Measurement*. ACM SIGCOMM Internet Measurement Workshop (IMW), San Francisco, USA, November 2001
- [SH02] S. Schnitter, G. Haßlinger. *Heuristic Solutions to the LSP-Design for MPLS Traffic Engineering*. 10th International Telecommunication Network Strategy and Planning Symposium (NETWORKS), München, Juni 2002.
- [SP02] K. Sinha, S. Patek. *OpIATE: Optimization Integrated Adaptive Traffic Engineering*. Systems and Information Engineering Report: SIE-020001, University of Virginia, USA, November 2002
- [TAP01] P. Trimintzios, I. Andrikopoulos, G. Pavlou, P. Flegkas, D. Griffin, P. Georgatos, D. Goderis, Y. T'Joens, L. Georgiadis, C. Jacquennet, R. Egan. *A Management and Control Architecture for Providing IP Differentiated Services in MPLS-Based Networks*. IEEE Communications Magazine, Mai 2002
- [TGP01] P. Trimintzios, L. Georgiadis, G. Pavlou, D. Griffin, C.F. Cavalcanti, P. Georgatos, C. Jacquennet. *Engineering the Multi-Service Internet: MPLS and IP-based Techniques*. IEEE International Conference on Telecommunications, Bukarest, Rumänien, Juni 2001
- [TMW97] K. Thompson, G.J. Miller, R. Wilder. *Wide-Area Internet Traffic Patterns and Characteristics*. IEEE Network, November/Dezember 1997
- [W97] J. Wroclawski. *The Use of RSVP with IETF Integrated Services*. RFC2210, September 1997
- [WC96] Z. Wang, J. Crowcroft. *QoS Routing for Supporting Resource Reservation*. IEEE Journal on Selected Areas in Communications, September 1996
- [WW99] Y. Wang, Z. Wang. *Explicit Routing Algorithms for Internet Traffic Engineering*. 8th International Conference on Computer Communications and Networks (ICCCN), Boston, USA, October 1999

- [WWZ01] Y. Wang, Z. Wang, L. Zhang. *Internet Traffic Engineering without Full Mesh Overlaying*. IEEE INFOCOM, Anchorage, Alaska, April 2001
- [XHB00] X. Xiao, A. Hannan, B. Bailey, L. M. Ni. *Traffic Engineering with MPLS in the Internet*. IEEE Network, März/April2000

